Towards a functional neuroanatomy of speech perception

Gregory Hickok and David Poeppel

The functional neuroanatomy of speech perception has been difficult to characterize. Part of the difficulty, we suggest, stems from the fact that the neural systems supporting 'speech perception' vary as a function of the task. Specifically, the set of cognitive and neural systems involved in performing traditional laboratory speech perception tasks, such as syllable discrimination or identification, only partially overlap those involved in speech perception as it occurs during natural language comprehension. In this review, we argue that cortical fields in the posterior–superior temporal lobe, bilaterally, constitute the primary substrate for constructing sound-based representations of speech, and that these sound-based representations interface with different supramodal systems in a task-dependent manner. Tasks that require access to the mental lexicon (i.e. accessing meaning-based representations) rely on auditory-to-meaning interface systems in the cortex in the vicinity of the left temporal–parietal–occipital junction. Tasks that require explicit access to speech segments rely on auditory–motor interface systems in the left frontal and parietal lobes. This auditory–motor interface system also appears to be recruited in phonological working memory.

Apriori, one might imagine that identifying the cortical fields that support speech perception would be a relatively straightforward task. After all, speech perception is tightly coupled to a single sensory modality, unlike higher-level lexical or syntactic processes in language, and therefore should have a relatively constrained neural organization. But it has been a century and a quarter since the first hypothesis concerning the neurology of speech perception was put forward[1], and still there is no consensus. Early authors proposed that auditory cortical fields in the left superior temporal lobe comprise the primary substrate of speech perception[1,2], others have pointed to the left inferior parietal lobe[3], or have argued that the left inferior frontal cortex is a critical component of the system[4], still others have emphasized the role of auditory cortices bilaterally[5].

Part of this confusion stems from differences in what one means by 'speech perception' and how one tests it behaviorally. Psychological research on speech perception typically utilizes tasks that involve the identification and/or discrimination of 'sub-lexical' segments of speech, such as meaningless syllables, and many neuropsychological and functional imaging studies have borrowed from this rich literature. The tacit assumption in this work, of course, is that these tasks tap much of the same set of processes involved in the perception of speech sounds embedded in 'normal' (e.g. conversational) speech. Whether this assumption is valid is an empirical
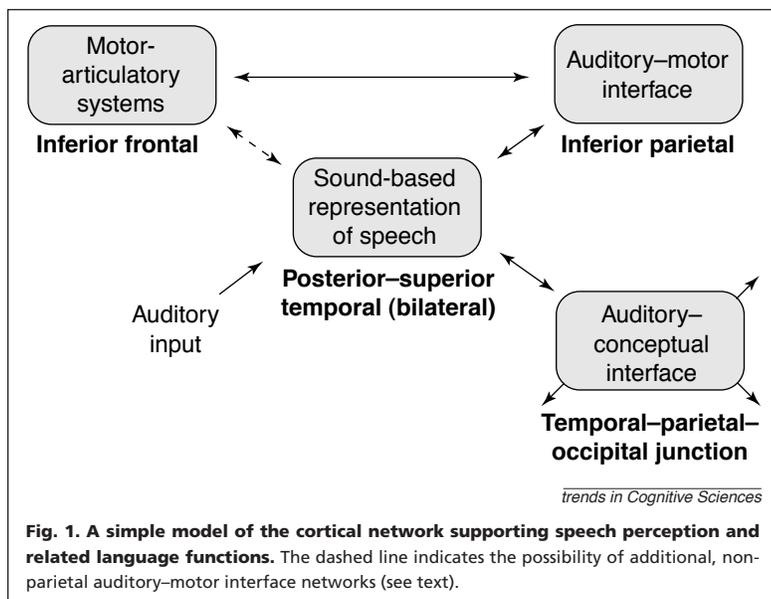
issue, and in fact, we will argue on the basis of neuropsychological and neurophysiological data, that it is *not* a fully valid assumption.

We would like to propose the following functional neuroanatomical model of speech perception. We suggest that auditory related cortical fields in the posterior half of the superior temporal lobe, *bilaterally*, constitute the primary substrate for constructing sound-based representations of speech. From this point, however, we will argue that there are at least two distinct pathways that participate in speech perception in a task-dependent manner, and that they are more strongly lateralized to the left hemisphere. The first is a ventral pathway, which probably involves cortex in the vicinity of the temporal-parietal-occipital junction. This pathway appears to be important for interfacing sound-based representations of speech with widely distributed conceptual representations[6], and therefore is involved in tasks that require access to the mental lexicon. The second is a dorsal pathway, which involves inferior parietal and frontal systems. This pathway appears to play a greater role in tasks that require explicit access to certain sub-lexical speech segments. We suggest below that these two pathways can be viewed in a way analogous to recent hypotheses concerning ventral and dorsal pathways in audition[7] and vision, with the ventral pathway subserving perception of the semantic content of sensory information (e.g. object recognition), and the dorsal pathway

G. Hickok is at the Department of Cognitive Sciences University of California, Irvine, CA 92697, USA.

tel: +1 949 824 1409
fax: +1 949 824 2307
e-mail: gshickok@ uci.edu

D. Poeppel is at the Department of Linguistics and Department of Biology, University of Maryland, College Park, MD, USA.

Fig. 1. A simple model of the cortical network supporting speech perception and related language functions. The dashed line indicates the possibility of additional, non-parietal auditory–motor interface networks (see text).

serving as an interface system between sensory and motor processes[8,9] (see Fig. 1). We will also discuss how verbal working memory, which also appears to involve the dorsal system, may fit into the proposed model.

The development of the present framework was motivated, in part, by a 20-year-old observation: In an effort to determine whether auditory comprehension deficits in aphasia resulted from phonemic perception deficits, Blumstein and colleagues[10] studied the ability of aphasic patients to identify and discriminate CV syllables in relation to their ability to comprehend auditorily presented speech. They found that the level of performance in the CV tasks was a poor predictor of auditory comprehension ability as measured by standard clinical diagnostic tests. In fact, in one study[11], the patient with the best score on the auditory comprehension subtest of the Boston Diagnostic Aphasia Examination[12] (BDAE), a Broca's aphasic, could not reliably label or discriminate the CV stimuli; and the patient with the worst outcome on the BDAE auditory comprehension subtest, a Wernicke's aphasic, performed normally on both of the CV perception tasks. This double dissociation demonstrates that the two types of tasks, sub-lexical tasks versus auditory comprehension tasks, involve different processes at some level of analysis. Of course, because both tasks involve similar sensory inputs they must also involve common neural systems up to some level in the processing stream. A central aim of this review is to clarify the point at which the processing streams supporting these two classes of tasks diverge. We will therefore review evidence concerning the functional neuroanatomy of speech perception first from the more ecologically valid perspective of auditory comprehension tasks, and then turn our attention to studies of speech perception involving tasks that require explicit attention to sub-lexical speech segments. Figure 2 illustrates the anatomy to which we will refer.

Speech perception subserving access to the mental lexicon
This section examines the functional neuroanatomy of speech perception primarily from the perspective of auditory comprehension tasks. Data from auditory comprehension tasks are relevant to the study of speech perception because a pre-

requisite of auditory comprehension is the construction of a sound-based representation of the speech input which is suitable for making contact with the mental lexicon. Thus, speech perception is one stage in the process of auditory comprehension. Although it is often difficult in any given study to tease apart the contribution of the various levels of processing involved in auditory comprehension, by looking at patterns of data across studies and methods, a clear picture begins to emerge. In this context, several lines of evidence converge on the view that the posterior superior temporal lobe, bilaterally, constitutes the primary substrate for constructing sound-based representations of speech.

Aphasia
Two types of aphasia are associated with substantial auditory comprehension deficits, Wernicke's aphasia and transcortical sensory aphasia (TSA)[13]. Comprehension deficits in Wernicke's aphasia are almost invariably associated with damage to the pSTP/STG, with extension either inferiorly into MTG or posteriorly into temporal-parietal-occipital junction structures, or both[13–15]. Comprehension deficits in TSA are associated with damage to temporal-parietal-occipital junction structures, with all or most of the pSTP/STG spared[16]. Does a defect in speech perception ability contribute to auditory comprehension deficits in either of these syndromes? The answer appears to be 'no' for TSA, and 'yes, but to a relatively small degree' for Wernicke's aphasia. Patients with TSA, by definition, have an intact ability to repeat heard speech[12,13], which shows that phonemic-perception processes are intact in that syndrome, and that the source of the comprehension deficit is post-phonemic. This in turn suggests that cortical fields in the temporal-parietal-occipital junction are involved in auditory comprehension at a post-phonemic level. The source of Wernicke's aphasics' comprehension impairment is less transparent since repetition ability is also impaired. A study by Baker *et al.*[17] addressed this issue by asking Wernicke's aphasics to perform a word-to-picture matching task in which the target picture (e.g. *bear*) was presented along with phonemic (*pear*), semantic (*wolf*), and unrelated (*grapes*) distractor pictures. Wernicke's aphasics rarely chose the unrelated foil (proportion of errors = 0.04), but did make both phonemic (0.12) and semantic errors (0.14), suggesting that impairments at both phonemic and post-phonemic processing levels contribute to auditory comprehension deficits in that syndrome. Overall, however, the Wernicke's aphasics accessed the correct phonological representation of the target more than 80% of the time (the sum of the correct and semantically related responses). This observation suggests that while a phonemic perception impairment contributes to auditory comprehension problems in Wernicke's aphasia, the deficit is not severe. The cortical substrate for this contribution is likely to be within the posterior–superior temporal lobe, which is a site of minimal overlap in the distribution of lesions in Wernicke's aphasia and TSA[16]. In summary, speech perception is not profoundly impaired in either Wernicke's aphasia or TSA (or in any other classical aphasia subtype). On the basis of these data alone, it is debatable whether the relatively preserved speech perception capacity in Wernicke's aphasia reflects residual abilities of the right hemisphere[18,19], or those of intact left hemisphere regions.

Data from other sources, however (reviewed below), suggests that it is the right hemisphere's speech perception abilities that are responsible for the relatively preserved speech perception in such cases.

Word deafness

If speech perception systems are organized bilaterally, then we should expect that bilateral superior temporal lobe lesions will produce profound speech perception deficits. This is exactly the case in 'pure word deafness' (henceforth, word deafness). Word deafness is a form of auditory agnosia in which the ability to comprehend auditorily presented speech is profoundly impaired. Patients are not deaf, and may have normal pure tone thresholds. Importantly, the auditory comprehension deficit in word deafness, unlike in Wernicke's aphasia, appears to reflect auditory speech perception difficulties, at least in those cases that were carefully tested[20,21]. Thus, while Bachman and Albert[22], in their review of auditory comprehension in aphasia, conclude that 'although deficits in phonemic discrimination can be detected in some [unilaterally lesioned] aphasic subjects, disordered phonemic perception does not appear to be the principal underlying deficit in impaired auditory comprehension', they go on to say, 'Pure word deafness, however, may represent an exception to this statement.' (p. 285). Word deafness, then, exhibits the profound speech-perception-based deficit not found among the unilateral aphasias.

The present hypothesis of bilateral organization of speech perception systems in the posterior–superior temporal lobe predicts that the lesions associated with such a condition should be bilateral in the superior temporal lobe, and indeed, the vast majority of word deaf cases present with this lesion pattern[23]. Bilateral posterior MTG lesions do not appear to have the same effect: Nielsen[24] sites such a case in which the lesions affected the posterior half of the MTG, with only partial extension into the white matter of the STG. Nielsen reports that 'after recovery [the patient] was able at all times to understand spoken language perfectly.' (p. 123).

The isolated right hemisphere

If the right hemisphere has the capacity to carry out speech perception processes, it should be possible to demonstrate speech perception abilities in the isolated right hemisphere. Studies of split brain patients and carotid amobarbital injection studies both indicate that in some cases at least, the isolated right hemisphere has the ability to understand simple (i.e. not syntactically complex) speech[25–27]. Even where comprehension errors do occur, there is evidence[27] to indicate that there is not a profound phonemic-based deficit in the right hemisphere. Further evidence comes from a recent amobarbital study, which has demonstrated that the isolated right hemisphere can perform speech discrimination tasks[5].

Physiological evidence

Physiological evidence from a variety of sources supports the view that speech perception involves the superior temporal lobe bilaterally. For example, a range of functional imaging studies of passive perception of speech stimuli – including PET[28–31], fMRI[32–34], and MEG[35–37] – have consistently found bilateral activation in the superior temporal lobe. Less con-
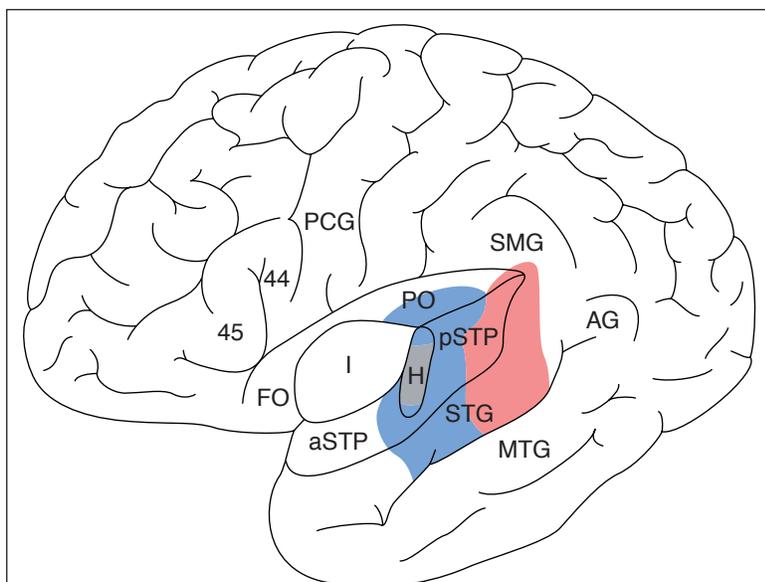
Fig. 2. Lateral view of the left hemisphere with structures inside the Sylvian fossa exposed. AG, angular gyrus; aSTP, pSTP, anterior and posterior supratemporal planes; FO, frontal operculum; H, transverse temporal gyrus of Heschl (which contains primary auditory cortex); I, insula; MTG, middle temporal gyrus; PCG, pre-central gyrus (primary motor cortex); PO, parietal operculum; SMG, supramarginal gyrus; STG, superior temporal gyrus; 44 and 45 refer to Brodmann's designations and together comprise Broca's area. Colors correspond to a rough cytoarchitectonic parcellation[56] into primary, koniocortical fields (gray), secondary, parakoniocortex fields (blue), and fields that are 'transitional' between unimodal sensory cortex and multimodal 'integration' cortex (pink). Note that we use the term STG to refer to the lateral (exposed) portion of that gyrus, and the terms pSTP, aSTP, and Heschl's gyrus to refer to various regions of temporal cortex buried inside the Sylvian fossa. The term 'superior temporal lobe' will be used generically to refer to these structures as a whole.

sistent across studies are activation sites that include Broca's area, MTG, and left temporal-parietal junction (posterior-inferior SMG). It is impossible to know precisely which aspect of the speech stimulus is producing each of these activations based on the functional imaging data alone. Nonetheless, it is clear that bilateral posterior–superior temporal lobe activation is the most consistent finding across studies involving heard speech. Putting this fact together with evidence from lesion studies indicating that focal lesions to Broca's area, the left SMG, or MTG are not associated with substantial auditory comprehension deficits, suggests that the posterior superior temporal lobe is the critical structure supporting speech perception.

A different line of physiological data provides further support for this view. Intraoperative recordings from single units in awake patients undergoing surgical treatment for epilepsy have suggested the existence of cells that are responsive to different aspects of the speech signal such as a particular class of phonemic clusters, monosyllabic versus multisyllabic words, task-relevant versus task-irrelevant speech, and natural versus distorted or backwards speech[38]. The vast majority of speech-responsive cells were found in the middle portion of the STG (recordings from more posterior sites were not made for clinical reasons); very few sites in anterior STG, middle or inferior temporal gyrus yielded detectable responses. No hemispheric asymmetries were found in the distribution of these units. Some caution is warranted, however, in interpreting these findings because clinical constraints precluded detailed and fully controlled experimentation.

Evidence from electrical interference

A study by Boatman and colleagues[39] provided an important demonstration of the involvement of the posterior–superior temporal lobe in speech perception processes. These authors used electrical interference to study the functional anatomy of speech perception in a series of three patients undergoing clinical evaluation for surgical treatment of epilepsy. Across patients, stimulation sites along the left pSTG consistently compromised performance on speech perception tasks, including at various sites syllable discrimination/identification and/or auditory comprehension. A variety of sites in anterior–superior temporal cortex as well as more inferior temporal lobe sites were tested but only pSTG sites were associated in every patient with deficits on one or more of these tasks. This finding provides strong support for the claim that the posterior superior temporal lobe is a critical structure in speech perception processes, no matter how it is measured. However, because perceptual deficits were observed with unilateral stimulation, this result does not seem to support our hypothesis that speech perception processes are bilaterally organized.

A more recent study[5] suggests the need for caution in making strong lateralization claims on the basis of these results. The syllable discrimination ability of a single patient was studied using both Wada and electrical interference procedures. Syllable discrimination remained intact during both left and right intracarotid amobarbital injections indicating that both hemispheres have the capacity to perform the task. When the same task was administered during focal left cortical electrical stimulation, however, a site was found that when stimulated, impaired performance on the syllable discrimination task. The mechanism responsible for these apparently contradictory findings is unclear. Boatman *et al.*[5] suggest that both hemispheres possess a possible route to speech perception, but that the left hemisphere route is preferred. Under conditions of partial deactivation of the left hemisphere system, processing via the preferred route is attempted unsuccessfully; when the entire left hemisphere is deactivated, the preferred route is completely unavailable and the less preferred right hemisphere route is used. We agree with their two-route analysis, but would like to suggest that under normal conditions the two routes operate in a coordinated fashion, each making a different, but important, contribution to the speech perception process. Determining what these contributions are, will be an important next step in this line of research (Box 1).

In summary, the data reviewed in this section make the case for the hypothesis that the posterior–superior temporal lobe, bilaterally, is the primary substrate for systems involved in constructing sound-based representations of heard speech. While each hemisphere appears capable of constructing such representations sufficiently well to allow contact with the mental lexicon, the computational mechanisms involved are probably different in the two hemispheres. Beyond this initial elaboration of acoustic input, sound-based representations of speech must interface with conceptual–semantic information, which is typically assumed to be stored in a widely distributed network throughout the cerebral cortex[6].

The multimodal cortical fields in the vicinity of the left temporal-parietal-occipital junction are a good candidate for networks important for interfacing sound-based representations in auditory cortex with widely distributed conceptual–semantic representations[13,40]: damage to this region is associated with transcortical sensory aphasia, which appears to reflect a post-phonemic deficit. This region is also typically involved to some extent in Wernicke's aphasia, which, as noted above, has a prominent post-phonemic component to the deficit profile. This proposed temporal-parietal-occipital sound–meaning interface network may not be a single focal network, but rather a system of networks whose spatial distribution of activation varies as a function of the cortical distribution of conceptual representations onto which a particular speech input is being mapped. This interface system may correspond to the 'lemma' level of representation in psycholinguistic models[41] in the sense that it serves to bind together different types of information, such as phonemic, semantic, and although not discussed here, perhaps morpho-syntactic information, all of which together define the content of individual entries in the mental lexicon.

Speech perception in tasks requiring explicit attention to phonetic structure

We have argued that the primary substrate for constructing sound-based representations of speech comprises auditory related cortices in the posterior superior temporal lobe bilaterally. Much of this argument was based on evidence from tasks that involve auditory comprehension. A glance at the literature on the functional neuroanatomy of speech perception, however, will reveal evidence suggesting that extra-auditory cortical regions play an important role in receptive phonetic processing. For example, performance on syllable discrimination and identification tasks can be impaired in many types of aphasia[10] including Broca's aphasia wherein the lesions involve left frontal cortex and largely spare temporal lobe structures. A recent lesion study[3] has identified inferior parietal lobe structures (parietal operculum and the SMG) as sites that may support the ability to perform syllable discrimination and identification tasks. And finally, functional imaging studies have identified activations in Broca's area (in addition to activation in superior temporal lobe) during the performance of sub-lexical tasks such as phoneme monitoring or deciding whether the last phoneme of a pair of syllables is the same or not[4,42–44]. Tasks that regularly involve these extra-auditory left hemisphere structures all seem to require explicit attention to segmental information[43,44]. Note that such tasks are fundamentally different from tasks that involve auditory comprehension: when one listens to an utterance in normal conversation, there is no conscious knowledge of the occurrence of specific phonemic segments, only the content of the message is consciously retained.

We hypothesize that the performance of these sub-lexical tasks involves neural systems that differ to some extent from those involved in auditory comprehension tasks. The evidence (reviewed above) supports this view: in neuropsychological studies the ability to perform sub-lexical tasks dissociates from the ability to comprehend auditory speech, and in functional imaging studies, auditory comprehension tasks do not regularly activate frontal and/or parietal structures. What appears to be a common denominator across tasks in functional imaging studies, is the involvement of the posterior–superior temporal lobe. This is consistent with our contention that the

Box 1. Differing contributions to speech perception by left and right temporal cortex

Although both left and right superior temporal cortices are implicated in speech perception, the evidence suggests that the left and right systems make different contributions to perceptual analysis. For example, a lesion study by Robin *et al.* showed that patients with left auditory cortex lesions were impaired on tasks involving the perception of temporal information, but unimpaired on tasks involving the perception of spectral information, whereas patients with right auditory cortex lesions showed the reverse effect (Ref. a). Zaidel reported that the perception of speech sounds by the isolated left or right hemispheres of split brain subjects was differentially affected by the addition of noise to the speech signal: noise adversely affected the performance of the right hemisphere to a greater degree than the performance of the left hemisphere (Ref. b). Differential effects of background noise on the cortical response in the two hemispheres have also been seen using MEG (Ref. c).

Other MEG studies have observed response differences in the two hemispheres that vary as a function of differences in speech stimuli. For example, the latency of the major response of auditory cortex elicited by words varies in the two hemispheres as a function of onset properties of the word (Ref. d). Using PET, Belin and colleagues investigated the response to rapidly (40 ms) versus slowly (200 ms) changing acoustic transitions in speech-like stimuli (Ref. e). They observed significantly different responses in the auditory cortices, with the left auditory cortex activation being more extensive in the rapid-transition condition (versus bilaterally symmetric in the slow-transition condition). A recent review of psychophysical and clinical research on processing asymmetries also highlights hemispheric differences in the degree to which the temporal structure of events can be discriminated, with the left hemisphere being selectively better at the discrimination of fine temporal events (Ref. f).

Although the computational basis of these hemispheric differences remains to be determined, several models have been proposed. Ivry and Robertson have developed the double-filtering-by-frequency model (DFF) to account for lateralization in auditory and visual perception (Ref. g). In DFF, all perceptual analysis begins with a spectral representation of the stimulus. An attentional filter then determines that point in the frequency domain that is relevant for the analysis of the signal. Around this anchoring point, the percept is processed asymmetrically, with the high-frequency part of the signal (with respect to the attentionally defined anchoring point) being processed in the left hemisphere, the low-frequency part in the right hemisphere. An alternative position is held by Zatorre who suggested that the left hemisphere is selectively better at temporal analysis and therefore has to sacrifice on spectral resolving power, whereas the right hemisphere is better at spectral analysis (Ref. h). On a third view, the asymmetric sampling in time model, the left and right temporal cortices analyse the same input signal on differing time scales (Ref. i). In particular, the temporal window of integration is argued to be short for the left hemisphere (on the order of 25 ms) and long for the right hemisphere (150–250 ms). The observed processing asymmetries for speech and non-speech follow from the asymmetric sampling of the waveform in the time domain.

References
a Robin, D.A. *et al.* (1990) Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain Lang.* 39, 539–555
b Zaidel, E. (1985) Language in the right hemisphere. In *The Dual Brain: Hemispheric Specialization in Humans* (Benson, D.F. and Zaidel, E., eds), pp. 205–231, Guilford Press
c Shtyrov, Y. *et al.* (1998) Background acoustic noise and the hemispheric lateralization of speech processing in the human brain: magnetic mismatch negativity study. *Neurosci. Lett.* 251, 141–144
d Gage, N. *et al.* (1998) Auditory evoked M100 reflects onset acoustics of speech sounds. *Brain Res.* 814, 236–239
e Belin, P. *et al.* (1998) Lateralization of speech and auditory temporal processing. *J. Cogn. Neurosci.* 10, 536–540
f Nicholls, M.E.R. (1996) Temporal processing asymmetries between the cerebral hemispheres: evidence and implications. *Laterality* 1, 97–137
g Ivry, R. and Robertson, L. (1998) *The Two Sides of Perception*, MIT Press
h Zatorre, R.J. (1997) Cerebral correlates of human auditory processing: perception of speech and musical sounds. In *Acoustical Signal Processing in the Central Auditory System* (Syka, J., ed.), pp. 453–468, Plenum Press
i Allard, F. and Scott, B.L. (1975) Burst cues, transition cues, and hemispheric specialization with real speech sounds. *Q. J. Exp. Psychol.* 27, 487–497

posterior–superior temporal lobe bilaterally is the *primary* substrate for constructing sound-based representations of speech. From this point, there is a divergence in the processing stream, with one stream involving more ventral–posterior structures, which play a role in auditory comprehension, and the other stream involving more dorsal–anterior systems, which play a greater role in sub-lexical speech tasks.

Why should inferior frontal and inferior parietal structures play a greater role in sub-lexical speech perception tasks? We hypothesize that these structures are part of a network important for auditory–motor integration. The existence of a sensory–motor interface system has been established in the visual domain where it is thought to be important for visually guided action[8,9,45]. In the spoken-language domain, the idea of sensory participation in production is an old one dating back at least to Wernicke, and garners support from the repeated observation that aphasics with posterior lesions regularly present with disordered speech output[13]. There is also evidence that unimodal auditory cortex plays a role in phonemic aspects of speech production[46] (Box 2). Furthermore, the ability to repeat heard pseudowords (where semantic mediation is impossible) clearly demonstrates the existence of an auditory–motor interface system.

Assuming that the frontal cortical fields implicated in speech are involved in coding speech in an articulatory context (not necessarily for overt speech only)[44], a reasonable hypothesis is that the inferior parietal lobe contains an important (but not necessarily the only) interface system mediating between auditory and articulatory representations of speech[47]. Such a network might be established (or tuned) in development as part of the process of learning to articulate speech where the young child must compare the sound of words uttered by others with those that he or she produces as a mechanism for tuning articulatory processes to the target language[48,49]. Attention to phonemic segments would seem

Box 2. A role for left auditory cortex in speech production

Early models of the neurology of language hypothesized the involvement of auditory cortex in speech production (Ref. a). Specifically, the same auditory networks involved in speech perception were thought to be activated during speech production. This claim was based on the observation that aphasics with superior temporal lobe lesions often presented with disordered (i.e. paraphasic) speech output. Recent functional imaging evidence has confirmed that left auditory cortex, including unimodal fields, is indeed activated during speech production. Using PET, Paus and colleagues for example, demonstrated a correlation between rate of speech production (whispering 'ba-lu' repeatedly, with white noise masking auditory feedback) and activation of left secondary auditory cortex posterior to Heschl's gyrus, suggesting a correlation between 'phonological load' and auditory cortex activity (Ref. b). Another study used magnetoencephalography to gauge the timecourse and spatial distribution of activations during an object naming task (Ref. c). The investigators noted a cluster of activation in the vicinity of the left supratemporal plane that emerged in a time window (275–400 ms after stimulus onset) corresponding to when 'phonological encoding' during object naming is thought to occur (Ref. c). A recent fMRI study of silent object naming also reported activation in the left pSTP in a majority of subjects (G. Hickok *et al.*, unpublished; see Fig. I). These findings are consistent with the hypothesis that the left pSTP /pSTG participates, not only in speech perception, but also in some form of phonemic processing during speech production

(Ref. d). The functional significance of these activations is demonstrated by the finding that lesions to the left superior temporal lobe are associated with conduction aphasia (Ref. e), a syndrome which, because of the prevalence of phonemic errors in speech output, has been characterized by some as an impairment of phonemic encoding. Direct cortical stimulation of pSTG has also been shown to elicit a conduction aphasia-like deficit including frequent phonemic paraphasias (Ref. f).

References

a Wernicke, C. (1874/1977) Der aphasische Symptomencomplex: eine psychologische Studie auf anatomischer Basis. In *Wernicke's Works on Aphasia: A Sourcebook and Review* (Eggert, G.H., ed.), pp. 91–145, Mouton

b Paus, T. *et al.* (1996) Modulation of cerebral blood flow in the human auditory cortex during speech: role of motor-to-sensory discharges. *Eur. J. Neurosci.* 8, 2236–2246

c Levelt, W.J.M. *et al.* (1998) An MEG study of picture naming. *J. Cogn. Neurosci.* 10, 553–567

d Hickok, G. (2000) Speech perception, conduction aphasia, and the functional neuroanatomy of language. In *Language and the Brain* (Grodzinsky, Y. *et al.*, eds), pp. 87–104, Academic Press

e Damasio, H. and Damasio, A.R. (1980) The anatomical basis of conduction aphasia. *Brain* 103, 337–350

f Anderson, J.M. *et al.* (1999) Conduction aphasia and the arcuate fasciculus: a re-examination of the Wernicke–Geschwind model. *Brain Lang.* 70, 1–12
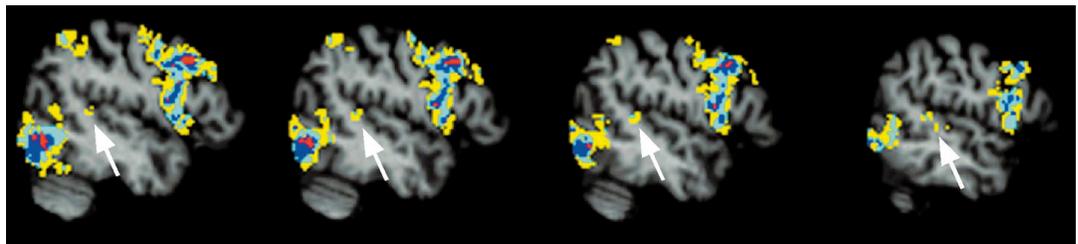
Fig. I. Functional anatomy of silent object naming. Contiguous sagittal slices in the left hemisphere showing activation of left auditory cortex (arrows) during a silent object naming task. Color scale indicates the proportion of subjects who showed activation in a given region (red = 0.1, dark blue = 0.86, light blue = 0.71, yellow = 0.57).

to be critical in this developmental process, and perhaps this auditory–motor interface network remains available to the adult and is useful in the performance of sub-lexical speech tasks. This framework explains why deficits on sub-lexical speech perception tasks have been observed in association with a variety of lesion patterns: damage either to lower-level sound-based representational systems in the posterior superior temporal lobe or to auditory–motor integration systems in the parietal and/or frontal lobe would be sufficient to interrupt performance on such tasks, although for different underlying reasons.

The idea that the inferior parietal lobe plays a role in auditory–motor transformations would bring this region in line functionally with other sensory–motor transformation systems in the parietal lobe. It may also provide a more principled account of the functional anatomy of phonological working memory[47]. Phonological working memory (comprising an 'articulatory subvocal rehearsal' component and a 'phonological store' component[50]) can be thought of as a system that uses articulatory mechanisms to keep phonemic representations active; it therefore requires a sensory–motor

interface. The subvocal rehearsal component appears to be supported by left frontal networks, whereas the phonological store mechanism relies on cortex in the left inferior parietal lobe[51,52]. We hypothesize that activations attributed to the functions of the phonological store reflect the operations of the proposed auditory–motor interface system in inferior parietal cortex. That is, inferior parietal cortex is not the site of storage of phonemic representations *per se* (or more precisely, it does not store the type of sound-based representations that are used during lexical access), but rather serves to interface sound-based representations of speech in auditory cortex with articulatory-based representations of speech in frontal cortex. Recently a connection between parietal lobe sensory–motor integration systems and working memory has been proposed in the primate visual system[45], a claim that is in line with the present conceptualization of phonological working memory.

We are suggesting therefore, that there is a fronto-parietal network predominantly in the left hemisphere, which functions to interface auditory and articulatory representations of speech. This system plays a key role in the development

of speech (i.e. learning to articulate the phonetic structure of the target language) and comprises the network underlying phonological working memory in the adult. Because this system provides a means for explicit access to sub-lexical speech segments, it is also recruited during the performance of tasks requiring explicit operations on sub-lexical information. Although parietal cortex appears to play a significant role in auditory–motor interaction, there are probably other, non-parietal networks that also function as an interface between auditory and motor representations[53–55], perhaps in the service of other language functions (e.g. normal speech production).

Conclusions

In this review we have argued: (1) that the posterior–superior temporal lobe bilaterally constitutes the primary cortical substrate for the construction of sound-based representations of speech; (2) that while both hemispheres participate, they probably make different contributions to speech perception (Box 1); (3) that left hemisphere frontal and parietal regions, sometimes implicated in aspects of speech perception (and also in phonological working memory), may be understood in terms of a system that supports auditory–motor interaction, by analogy to similar systems supporting visual–motor interaction; and (4) that auditory cortex in the left pSTP/pSTG participates in aspects of speech production (Box 2). The picture that emerges from these considerations is one in which pSTP/pSTG auditory networks bilaterally are important for assembling sound-based representations of speech, which interface both with left hemisphere motor speech-articulation systems via the parietal lobe, and with lexical–semantic representations via cortical fields in the vicinity of the left temporal-parietal-occipital junction.

References

1 Wernicke, C. (1874/1977) Der aphasische Symptomencomplex: eine psychologische Studie auf anatomischer Basis. In *Wernicke's Works on Aphasia: A Sourcebook and Review* (Eggert, G.H., ed.), pp. 91–145, Mouton

2 Luria, A.R. (1970) *Traumatic Aphasia*, Mouton

3 Caplan, D. *et al.* (1995) Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. *Neurology* 45, 293–298

4 Fiez, J.A. *et al.* (1995) PET studies of auditory and phonological processing: Effects of stimulus characteristics and task demands. *J. Cogn. Neurosci.* 7, 357–375

5 Boatman, D. *et al.* (1998) Right hemisphere speech perception revealed by amobarbital injection and electrical interference. *Neurology* 51, 458–464

6 Damasio, A.R. (1989) The brain binds entities and events by multiregional activation from convergence zones. *Neural Comput.* 1, 123–132

7 Rauschecker, J.P. (1998) Cortical processing of complex sounds. *Curr. Opin. Neurobiol.* 8, 516–521

8 Andersen, R. (1997) Multimodal integration for the representation of space in the posterior parietal cortex. *Philos. Trans. R. Soc. London Ser. B* 352, 1421–1428

9 Goodale, M.A. (1998) Vision for perception and vision for action in the primate brain. *Novartis Foundation Symp.* 218, 21–39

10 Blumstein, S.E. *et al.* (1977) Phonological factors in auditory comprehension in aphasia. *Neuropsychologia* 15, 19–30

11 Blumstein, S.E. *et al.* (1977) The perception and production of voice-onset time in aphasia. *Neuropsychologia* 15, 371–383

12 Goodglass, H. and Kaplan, E. (1976) *The Assessment of Aphasia and Related Disorders*, Lea & Febiger

13 Damasio, A.R. (1992) Aphasia. *New Engl. J. Med.* 326, 531–539

14 Kertesz, A. *et al.* (1993) The structural determinants of recovery in Wernicke's aphasia. *Brain Lang.* 44, 153–164

15 Naeser, M.A. *et al.* (1987) Relationship between lesion extent in 'Wernicke's area' on computed tomographic scan and predicting recovery of comprehension in Wernicke's aphasia. *Arch. Neurol.* 44, 73–82

16 Damasio, H. (1991) Neuroanatomical correlates of the aphasias. In *Acquired Aphasia*, pp. 45–71, Academic Press

17 Baker, E. *et al.* (1981) Interaction between phonological and semantic factors in auditory comprehension. *Neuropsychologia* 19, 1–15

18 Kleist, K. (1962) *Sensory Aphasia and Amusia: The Myeloarchitectonic Basis*, Pergamon Press

19 Benson, D.F. *et al.* (1973) Conduction aphasia: a clinicopathological study. *Arch. Neurol.* 28, 339–346

20 Albert, M.L. and Bear, D. (1974) Time to understand: a case study of word deafness with reference to the role of time in auditory comprehension. *Brain* 97, 373–384

21 Yaqub, B.A. *et al.* (1988) Pure word deafness (acquired verbal auditory agnosia) in an Arabic speaking patient. *Brain* 111, 457–466

22 Bachman, D.L. and Albert, M.L. (1988) Auditory comprehension in aphasia. In *Handbook of Neuropsychology* (Vol. 1) (Boller, F. and Grafman, J., eds), pp. 281–306, Elsevier

23 Buchman, A.S. *et al.* (1986) Word deafness: one hundred years later. *J. Neurol. Neurosurg. Psychiatry* 49, 489–499

24 Nielsen, J.M. (1946) *Agnosia, Apraxia, Aphasia: Their Value in Cerebral Localization.*, Paul B. Hoeber

25 McGlone, J. (1984) Speech comprehension after unilateral injection of sodium amytal. *Brain Lang.* 22, 150–157

26 Wada, J. and Rasmussen, T. (1960) Intracarotid injection of sodium amytal for the lateralization of cerbral speech dominance. *J. Neurosurg.* 17, 266–282

27 Zaidel, E. (1985) Language in the right hemisphere. In *The Dual Brain: Hemispheric Specialization in Humans* (Benson, D.F. and Zaidel, E., eds), pp. 205–231, Guilford Press

28 Mazoyer, B.M. *et al.* (1993) The cortical representation of speech. *J. Cogn. Neurosci.* 5, 467–479

29 Petersen, S.E. *et al.* (1988) Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature* 331, 585–589

30 Price, C.J. *et al.* (1996) Hearing and saying: the functional neuro-anatomy of auditory word processing. *Brain* 119, 919–931

Outstanding questions

- We have proposed that superior temporal lobe structures play an important role in constructing 'sound-based representations of speech.' This process is complex, probably involving multiple levels of representation. How does this general notion of sound-based representations map onto the different linguistic levels of representation (e.g. phonetic features, syllabic structure, etc.)? Are there neuroanatomical subdivisions within auditory cortex that correspond to these levels of representation?

- It is common for researchers to draw a distinction between 'acoustic' and 'phonological' processes in auditory perception, often with the neuroanatomical correlates of this distinction falling into the right and left hemispheres respectively. How does this distinction fit with the present hypothesis of bilateral speech perception systems? To what extent is this distinction driven by evidence from sub-lexical speech perception tasks, which we have suggested involve left lateralized frontoparietal systems rather than bilateral auditory-related systems?

- The vast majority of psychological research on speech perception has employed sub-lexical tasks such as syllable discrimination and identification. Assuming the present arguments to be correct, to what extent do the findings from this literature reflect the computations of speech perception systems in auditory cortex versus the operations of the proposed auditory-motor interface system?

31 Zatorre, R.J. *et al*. (1996) PET studies of phonetic processing of speech: review, replication, and re-analysis. *Cereb. Cortex* 6, 21–30

32 Binder, J.R. *et al*. (1994) Functional magnetic resonance imaging of human auditory cortex. *Ann. Neurol.* 35, 662–672

33 Dhankhar, A. *et al*. (1997) Functional magnetic resonance imaging assessment of the human brain auditory cortex response to increasing word presentation rates. *J. Neurophysiol.* 77, 476–483

34 Schlosser, M.J. *et al*. (1998) Functional MRI studies of auditory comprehension. *Hum. Brain Mapp.* 6, 1–13

35 Gage, N. *et al*. (1998) Auditory evoked M100 reflects onset acoustics of speech sounds. *Brain Res.* 814, 236–239

36 Kuriki, S. *et al*. (1995) Source analysis of magnetic field responses from the human auditory cortex elicited by short speech sounds. *Exp. Brain Res.* 104, 144–152

37 Poeppel, D. *et al*. (1996) Task-induced asymmetry of the auditory evoked M100 neuromagnetic field elicited by speech sounds. *Cognit. Brain Res.* 4, 231–242

38 Creutzfeldt, O. *et al*. (1989) Neuronal activity in the human lateral temporal lobe: I. Responses to speech. *Exp. Brain Res.* 77, 451–475

39 Boatman, D. *et al*. (1995) Auditory speech processing in the left temporal lobe: an electrical interference study. *Brain Lang.* 51, 269–290

40 Mesulam, M-M. (1998) From sensation to cognition. *Brain* 121, 1013–1052

41 Levelt, W.J.M. (1999) Models of word production. *Trends Cognit. Sci.* 3, 223–232

42 Demonet, J-F. *et al*. (1992) The anatomy of phonological and semantic processing in normal subjects. *Brain* 115, 1753–1768

43 Burton, M.W. *et al*. (1998) Neural basis of phonological processing: the role of segmentation. *Brain Lang.* 65, 249–252

44 Zatorre, R.J. *et al*. (1992) Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256, 846–849

45 Rizzolatti, G. *et al*. (1997) Parietal cortex: from sight to action. *Curr. Opin. in Neurobiol.* 7, 562–567

46 Hickok, G. (2000) Speech perception, conduction aphasia, and the functional neuroanatomy of language. In *Language and the Brain* (Grodzinsky, Y. *et al.*, eds), pp. 87–104, Academic Press

47 Aboitiz, F. and García V.R. (1997) The evolutionary origin of language areas in the human brain: a neuroanatomical perspective. *Brain Res. Rev.* 25, 381–396

48 Doupe, A.J. and Kuhl, P.K. (1999) Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* 22, 567–631

49 Werker, J.F. and Tees, R.C. (1999) Influences on infant speech processing: toward a new synthesis. *Annu. Rev. Psychol.* 50, 509–535

50 Baddeley, A.D. (1992) Working memory. *Science* 255, 556–559

51 Jonides, J. *et al*. (1998) The role of parietal cortex in verbal working memory. *J. Neurosci.* 18, 5026–5034

52 Awh, E. *et al*. (1996) Dissociation of storage and rehearsal in working memory: PET evidence. *Psychol. Sci.* 7, 25–31

53 Romanski, L.M. *et al*. (1999) Auditory belt and parabelt projections to the prefrontal cortex in the Rhesus monkey. *J. Comp. Neurol.* 403, 141–157

54 Dronkers, N.F. (1996) A new brain region for coordinating speech articulation. *Nature* 384, 159–161

55 Damasio, H. and Damasio, A.R. (1980) The anatomical basis of conduction aphasia. *Brain* 103, 337–350

56 Galaburda, A. and Sanides, F. (1980) Cytoarchitectonic organization of the human auditory cortex. *J. Comp. Neurol.* 190, 597–610

Inhibition of return

Raymond M. Klein

R.M. Klein is at the Department of Psychology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4J1.

tel: +1 902 494 6551
fax: +1 902 494 6585
e-mail: ray.klein@
dal.ca
http://www.dal.ca/
~klein/

Immediately following an event at a peripheral location there is facilitation for the processing of other stimuli near that location. This is said to reflect a reflexive shift of attention towards the source of stimulation. After attention is removed from such a peripheral location, there is then delayed responding to stimuli subsequently displayed there. This inhibitory aftereffect, first described in 1984 and later labeled 'inhibition of return (IOR)', encourages orienting towards novel locations and hence might facilitate foraging and other search behaviors. Since its relatively recent discovery, IOR has been the subject of intensive investigation, from many angles and with a wide variety of approaches. After describing the seminal contribution of Posner and Cohen ('Who'), this review will discuss what causes IOR and, once initiated, what effects IOR has on subsequent processing ('What'). The time course ('When') and spatial distribution ('Where') of IOR, and what is known about IOR's neural implementation ('How') and functional significance ('Why') are also discussed.

Orienting is one of the most primitive functions of living things. Even the leaves of a plant orient towards sunshine; the roots towards moisture. In more complex organisms more complex systems have evolved to orient the various receptors reflexively either towards or away from signal sources in the environment and to prepare the organism to select from a repertoire of behavioral actions[1]. Whereas the efficiency and rapidity provided by reflexive control of orienting would be critical in predation and defense, voluntary control over the orienting reflexes has undoubtedly been an important evolu-

tionary development. Efficient foraging for food (or other desirable objects, places, playmates) involves not only voluntary control over orienting but also the use of information stored in memory about one's previous orienting behavior. Once discovered, the locus of a food source that is not exhausted should be remembered as a place to return to. By contrast, places one has searched and not found what one is looking for, or places that have been stripped bare, should be remembered to be avoided. A mechanism that encourages orienting towards novel locations would be useful in these