



Università
Ca'Foscari
Venezia

DOTTORATO DI RICERCA
IN SCIENZE DEL LINGUAGGIO
Scuola di dottorato in Scienze del Linguaggio
Ciclo 25
(A.A. 2011 - 2012)

The Pronunciation of Mandarin Chinese

According to the ^{can}IPA Natural Phonetics & Tonetics Method

SETTORE SCIENTIFICO DISCIPLINARE DI AFFERENZA: L-LIN/01

Tesi di dottorato di Marco Cerini, matricola 955751

COORDINATORE DEL DOTTORATO
Prof. Alessandra Giorgi

TUTORE DEL DOTTORANDO
Prof. Luciano Canepari

Contents

The Pronunciation of Mandarin Chinese

According to the ^{can}IPA Natural Phonetics & Tonetics Method

| | |
|----|--|
| | 0. Introduction |
| 1 | 0.1. The official romanization – Hànyǔ pīnyīn fāng'àn |
| 2 | 0.2. The finals |
| 2 | 0.3. Phonotactics |
| | 1. Vowels and vocoids |
| 5 | 1.1. Vowels |
| 8 | 1.2. Phonemic conciseness vs realism: the interphonemic choice |
| 11 | 1.3. On the phonemic status of -i |
| 16 | 1.4. Diphthongs |
| 17 | 1.5. Vowel reductions |
| | 2. The r grammeme and the so-called érhua phenomenon |
| 21 | 2.1. Origin and use of érhua |
| 22 | 2.2.4. Articulatory characteristics of érhua |
| 25 | 2.5. Taxophonic variation in /Vɿ/ and /VVɿ/ |
| 27 | 2.6. Taxophonic variation in //Vnɿ// and /Vnɿ/ |
| 28 | 2.7. Various examples of érhua, according to our 'suggested' pronunciation scheme |
| 30 | 2.8. Frequency and acceptance of érhua in cultivated speech ^{or} non-neutral pronunciations |
| 31 | 2.8.3. An 'international' pronunciation scheme of érhua (for 'practical' teaching purposes only) |
| | 3. Consonants and contoids |
| 34 | 3.1. Main features of Mandarin consonantal system (with a table of neutral taxophones) |
| 35 | 3.2. Nasals |
| 36 | 3.3. Stops and stop-strictives |
| 39 | 3.4. Constrictives |
| 42 | 3.5. Approximants |
| 44 | 3.6. Laterals |
| 45 | 3.7. Casual speech |

| | | |
|-----|----------|--|
| | 4. | <i>Chinese tonemes, tones, taxotones & stress</i> |
| 47 | 4.1. | Tonality as a phonologically distinctive feature |
| 47 | 4.1.2-3 | An introduction to the four marked tonemes and the so-called ‘zero’ (unmarked) toneme |
| 49 | 4.2. | Taxotones and stress |
| 50 | 4.3. | Primary (or ‘full’) taxotones of the four marked tonemes |
| 51 | 4.4. | Rhythmic groups and taxotones |
| 51 | 4.5. | Further taxotones, including /,\$(\$)/ [,\$, ..\$], /,\$(\$)/ [‘\$, ‘\$], and /,\$(\$)/ [‘\$, ‘\$] |
| 53 | 4.6. | Taxotones in three-syllable sequences, including /‘\$/ [‘\$] → [‘\$] |
| 56 | 4.7. | The so called ‘zero’ toneme, or <i>qīngshēng</i> |
| 57 | 4.8. | Tonal behaviour in /‘\$°\$°\$/ sequences |
| 58 | 4.9. | Tonal behaviour of <i>bù</i> ‘not’, <i>méi</i> ‘not’, <i>yī</i> ‘one’ (plus <i>qī</i> ‘seven’ and <i>bā</i> ‘eight’) |
| | 5. | <i>An introduction to Chinese intonation</i> |
| 61 | 5.1. | Tonality and intonation |
| 63 | 5.2. | The <i>canIPA</i> tonogram |
| 64 | 5.3. | Studying and learning intonation |
| 65 | 5.4. | The intonation of English |
| 65 | 5.5. | In-tonemes and preintonemes |
| 67 | 5.6-9. | English intonational patterns |
| 69 | 5.10. | The intonation of Mandarin Chinese |
| 70 | 5.11. | The <i>canIPA</i> symbols inventory to annotate Chinese intonation |
| 71 | 5.12. | Mandarin Chinese preintonemes |
| 73 | 5.13. | Mandarin Chinese intonemes |
| 76 | 5.15-18. | Various kinds of questions and their intonational patterns |
| 80 | 5.19. | The continuative intoneme |
| 81 | 5.20. | Examples illustrating Mandarin Chinese intonation behaviour |
| 84 | 5.21. | Emphasis |
| 85 | 5.22. | Parentheses and quotations |
| 87 | 5.23. | Natural perception vs acoustic measurements |
| | 6. | <i>Phono-tonetic analysis of extended texts</i> |
| 89 | 6.1. | Phono-tonetic transcription of Aesop’s <i>The North Wind and the Sun</i> |
| 94 | 6.2. | Phono-tonetic transcription of <i>Liúyǒu yúdi</i> |
| | 7. | <i>Phonetic transcriptions in current handbooks</i> |
| 98 | 7.1. | A proposal for a simpler, teaching-oriented transcription to be used in Chinese textbooks |
| 102 | 7.2. | Phono-tonetic transcription of <i>Liúyǒu yúdi</i> according to the ‘simplified’ scheme |
| 103 | 7.3. | Scarce attention to phonetics, and pronunciation in general, by most teachers of Chinese |

| | | |
|-----|------------|--|
| 104 | 7.3.1. | An ‘enlightened’ use of <i>offIPA</i> in the old Linguaphone Institute textbook <i>Chinese</i> |
| 106 | 7.3.3. | The officially-sanctioned description of Mandarin Chinese as presented in the <i>JIPA</i> |
| 108 | 7.3.5-6. | Keeping <i>offIPA</i> ‘true to itself’ |
| | 8. | <i>Peking</i> |
| 112 | 8.1. | The Peking accent <i>vs</i> the Peking dialect |
| 113 | 8.2. | Classical Pekingese and contemporary Pekingese |
| 114 | 8.3. | Readily available examples of genuine Pekingese pronunciation |
| 114 | 8.4. | Fundamental traits |
| | 9. | ‘Praateries’ |
| 118 | 9.1. | Usefulness of computer-aided analysis applied to phonetics and tonetics |
| 119 | 9.2. | Detecting consonantal traits |
| 120 | 9.3. | Detecting vocalic traits |
| 121 | 9.4. | Analysing tonality |
| 122 | 9.5. | A few sentences transcribed in <i>canIPA</i> , with their relevant <i>Praat</i> diagrams |
| 125 | 10. | <i>Conclusions</i> |
| 127 | | <i>Bibliography</i> |

o. Introduction

We provide the modern neutral pronunciation of modern Mandarin Chinese, or *pǔtōnghuà* /₁phuːthoŋ˥˥wa/ [ˌph̥uː˥˥thoŋ˥˥wa], ie the ‘common language’, based on Peking dialect. We obviously use the official romanization system *pīnyīn* /ˈphin̄jin/ [ˈph̥in̄˥˥jin], accompanied by a careful phonetic transcription, to show and acquire a good pronunciation. ¹

At the beginning, a well meditated phonemic transcription will be provided too, in accordance with precise descriptive and teaching strategies. To improve readability, the phonemic transcription will be progressively omitted in the following chapters, wherever *pīnyīn* suffices to highlight any relevant phonemic facts.

The official romanization – Hànyǔ pīnyīn fāng’àn.

o.1. *Pīnyīn* stands at midway between the Chinese characters and our phonotonic transcriptions. This romanization scheme was developed, and officially adopted, in 1958, by the People’s Republic of China, and now is the worldwide *de facto* standard. Its full Chinese name is *Hànyǔ pīnyīn fāng’àn*, where *pīnyīn* roughly means ‘to put sounds together’. ² We shall present it here, even though, to be entirely loyal to the natural approach, any graphemic representation should come last, *after* the spoken language has been duly described and learnt.

The twenty-three *initials* are listed below, and arranged not alphabetically, but by logical phonic groups. Each is accompanied by its phonemic transcription and only one phonetic variant (the one used in fully-stressed syllables):

| | | | | |
|----------------------|------------------------|-------------------|----------------------------|------------------|
| <i>b</i> /p/ [p̚] | <i>p</i> /ph/ [ph̥] | <i>f</i> /f/ [f] | <i>m</i> /m/ [m] | |
| <i>d</i> /t/ [t̚] | <i>t</i> /th/ [th̥] | | <i>n</i> /n/ [n] | <i>l</i> /l/ [l] |
| <i>z</i> /ts/ [ts̚] | <i>c</i> /tsh/ [tsh̥] | <i>s</i> /s/ [s] | | |
| <i>j</i> /tɕ/ [tɕ̚] | <i>q</i> /tsh/ [tɕh̥] | <i>x</i> /s/ [ç] | <i>y</i> /j/ [j] & /ɥ/ [ɥ] | |
| <i>zh</i> /tʂ/ [tʂ̚] | <i>ch</i> /tʂh/ [tʂh̥] | <i>sh</i> /ʂ/ [ʂ] | <i>r</i> /ʐ/ [ʐ] | |
| <i>g</i> /k/ [k̚] | <i>k</i> /kh/ [kh̥] | <i>h</i> /x/ [x] | <i>w</i> /w/ [w] | |

o.2. The *finals* are 39 (with three occurrences of *i* and three *different* values), and we present them according to their *pīnyīn* graphic syllabic nucleus. This time, they are given in alphabetical order, even where their romanization has not been the happiest one (as we show in { }). The three finals which do not occur without an initial are marked with an asterisk (*): ³

A

a /a#/ [a], *ia*, #*ya* /ja#/ [ja], *ua*, #*wa* /wa#/ [wa];
ai /aɛ/ [aɛ], *uai*, #*wai* /waɛ/ [waɛ]; *an* /an/ [an], *ian*, #*yan* /jɛn/ [jɛn], *uan*, #*wan*
 /wan/ [wan], *üan*, #*yuan* (j/q/x + *uan*) /ɥɛn/ [ɥɛn];
ang /an/ [an], *iang*, #*yang* /jan/ [jan], *uang*, #*wang* /wan/ [wan];
ao /aɔ/ [aɔ], *iao*, #*yao* /jaɔ/ [jaɔ].

E

e /ɛ#/ [ɛ, ɛɤ, ɔɤ], *ie*, #*ye* /jɛ#/ [jɛ], *üe*, #*yue* (j/q/x + *ue*) /ɥɛ#/ [ɥɛ];
*ei** /ɛi/ [ɛi] ([ɛi]), *en* /ɛn/ [ɛn], *er* /ɛʐ/ [ɛʐ];
*eng** /ɛŋ/ [ɛŋ], *ueng*, #*weng* /wɛŋ, wɔŋ/ [wɛŋ, wɔŋ].

I

i /i#, #ji#/ [i, ji]; (s/z/c) /ɿ#/ [ɿ], (sh/zh/ch) /ʉ#/ [ʉ] {we adopt *i*, instead};
in, #*yin* /jin/ [in, #jin]; *ing*, #*ying* /jin/ [in, #jin];
ui, #*wei* /wei/ [wɛi] ([wɿ]) {*uei* would have been better than *ui*}.

O

o /ɔ#/ [ɔ, ɔɤ], *ou* /ɔu/ [ɔu] ([ɔu]), *uo*, #*wo* /wɔ#/ [wɔ, wɔɤ] ([wɔ, wɔɤ]);
*ong** /ɔŋ/ [ɔŋ], (/un/ [un]), *iong*, #*yong* /jɔŋ/ [jɔŋ]

U

u, #*wu* /u#, #wu#/ [u, wu]; *ü*, #*yu* (j/q/x + *u*) /y#, #ɥy#/ [y, #ɥy];
un, #*wen* /wɛn/ [wɛn]; *ün*, #*yun* (j/q/x + *un*) /ɥyn/ [#ɥyn];
iu, #*you* /jɔu/ [jɔu] ([jɔu]) {we employ *ü* whenever it stands for /y, ɥ/; besides,
uen and *iou* would have been better than *un* and *iu*}.

Phonotactics

o.3. Chinese phonotactics obeys to very strict limitations: in fact, just 400-odd distinct syllables exist in the modern language, tonemes not considered. Not all acceptable permutations are used in fact, like *shong*, though dictionaries list various

words with *zhong*, *chong* and *rong*. Other syllables do not occur with certain tones, for historical reasons: for instance, there is no *ēr* (*er* in the first tone).⁴

Mandarin phonology has been analyzed according to very different criteria, but we maintain that any Chinese syllable consists of an *initial* element and a *final* one, traditionally called *shēngmǔ* /ʃɛ̃ŋN_Jmu/ [ʃɛ̃ŋN_Jm̥u̥] ‘mother of the voice’, and *yùnmǔ* /\uʔyn_Jmu/ [\uʔy̥m̥_Jm̥u̥] ‘rhyme of the voice’.⁵

At the phonemic level, all *initials* are consonantal by their nature, even though, phonetically, a contoid is not necessarily present. In the case of close vowels //[#]i, [#]y, [#]u//, then, we actually have /ji, ɥy, wu/ [(j)i, (ɥ)y, (w)u].⁶ Before other vowels, our choice is /ʔ/: /ʔɛ, ʔɤ, ʔσ, ʔa/ (/ʔu/ does not occur) — which can be safely realized as such, [ʔ], or –less advisably– as [ɣ, ɦ, ʌ, ŋ] (: velar semi-approximant, lenis laryngeal approximant, uvular tap, and provelar semi-nasal). Nevertheless, when there is no risk of hiatus with a preceding vowel, /ʔ/ spontaneously reduces to ‘zero’, [∅].

In its full form, a *final* is a sequence of three parts: a *glide* /j, ɥ, w/; a *vocalic nucleus* /i, y, u, ɛ, ɤ, σ, a/; and a *coda*, which can be vocalic /i, ɛ, u, σ/, or consonantal /n, ŋ, ʐ/. For example, combining /ʃ, w, n; a, σ/, we obtain /ʃa/ *sha*, /ʃaσ/ *shao*, /ʃwa/ *shua*, /ʃan/ *shang*, and /ʃwan/ *shuang*. Also, glides can act as *initials*: /wa/ *wa*, /wan/ *wang*.

Many textbooks and dictionaries include a list of permitted syllables; one of the most exhaustive and reliable ones, in our view, is that provided by *The Contemporary Chinese Dictionary*, which is also the reference book with respect to the orthography of all Chinese words quoted in this work, including those featuring a ‘zero’ toneme ^{or} the so-called *érhuà* phenomenon.⁷

NOTES:

¹ We assume that our readers have a working knowledge of the *Canepari International Phonetic Alphabet* — ‘*canIPA*’, and of the Natural Method, as thoroughly described in CANEPARI, L. (1983) *Phonetic Notation · La notazione fonetica*, and in CANEPARI, L. (2007) *Natural Phonetics & Tonetics – articulatory, auditory, functional*.

The *canIPA* inventory is a groundbreaking development of the ‘official’ *IPA* (cursorily: ‘*offIPA*’), not just a mere extension, and it is purposely conceived to provide a very high accuracy rate, both at the phonetic and the tonetic level. Particularly in transcribing tonal languages, the unique *canIPA* tonemic and tonetic notation finally makes it possible for the transcriber to annotate the exact tonality of single syllable *and* of entire utterances, in a readable and unambiguous form.

² See CAO, X-ZH. (2000) *Hanyu pinyin yongchu duo*, p. 1-8; and LU, J-M. (2004) *Yuwen xian-daihua he Hanyu pinyin fang'an*, particularly p. 113-133.

³ In this, we follow CANEPARI, L. (2007) *A Handbook of Pronunciation*, § 11.0.6 *et seq*, with minimal amendments.

⁴ However, the Peking dialect also features a few words with peculiar phonology, like *dèn*, which not only are absent in the official inventory, but also not *permitted*, even in theory. See also ZHOU, Y-M. (2002) *Xiandai Beijinghua yanjiu*, particularly ㉓ 3 and 6.

⁵ Contrary to Western tradition, where dictionaries commonly list words according to their initial letter, classical Chinese dictionaries generally considered single characters –instead of words– arranged according to their ‘radical’ components and the number of calligraphic traits. Pronunciation, at first, was indicated by means of homophony (eg English *sea*, [now] pronounced like *see*).

Then, a mechanism called *fǎnqiè*, or ‘cross-cutting’, was developed: it consisted in a pair of characters, which indicated the initial and the ‘rhyme’, respectively (eg a combination like *chē-huáng*, to express *chuáng*; or, in English, *she-lock*, to express *shock*). Hence the importance of keeping initials well distinct from the ‘rhymes’. The latter, moreover, were particularly important for classical poetry; not surprisingly, then, ancient poets used to refer to specifically rhyme-arranged dictionaries, which nowadays provide precious –but often vague and ambiguous– hints at the pronunciation of older forms of Chinese (among others, see PULLEYBLANK, E. G. [1991] *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early mandarin*).

⁶ cf LIN, H-Y. (2007) *The Sounds of Chinese*, ㉓ 3, especially § 3.2 (and § 5.2.4). The author seems reluctant to grant full phonemicity to [j, ɥ, w], yet she methodically adopts [ji, ɥy, wu] –akin to what we ourselves do– instead of the more conservative approach, whose supporters still propose things like [i, y, u], as the sole phonetic counterpart of their theoretical ‘/i, y, u/’, even when an approximant is definitely there (both articulatorily and auditorily – and acoustically, of course). Coherently, she also employs [j, ɥ, w] before other vowels, and after other contoids (where present), eg *huan* [xwan], for ^{can}[ɣwan] (but always better than [huan], as found in too many books and articles).

⁷ The *Contemporary Chinese Dictionary* is also published in a Chinese-English bilingual edition, very convenient for those who are interested in Mandarin, but do not actually master the language up to the point of using a monolingual dictionary.

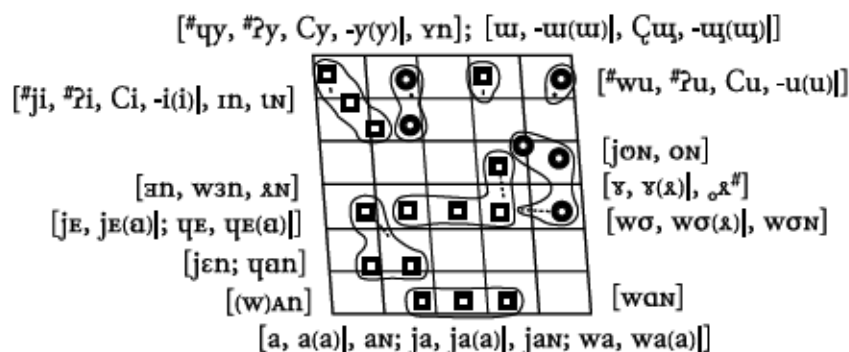
Again, *The Sounds of Chinese* also devotes quite a large amount of space to analyze the syllabic restraints determined by Chinese phonotactics (particularly, see § 5.3). We regard such an approach as slightly too theoretical, when compared to the highly pragmatic nature of our method; nonetheless, we would agree with many of the principles stated by the author, provided it is made clear that the phonetic reality –well understandably– is much more complex.

1. Vowels and vocoids

Vowels

1.1. After thorough consideration, we have chosen to include *eight* vowels in our analysis of Mandarin phonology: /i, y, u, u; e, ɤ, σ; a/. In fig 1.1-A, the various vocoids and their taxophones –ie contextual allophones– are grouped by phoneme; while, in fig 1.1-B, we present the same taxophones arranged in ‘functional’ groups: /V#, Vn, VN/. Let us start, then, by systematically considering each phoneme and its most typical realizations in neutral pronunciation (further –less representative– variants will be shown below, as well, more for recognition than for actual use).¹

fig 1.1-A. Mandarin fundamental vocoids and their taxophones, grouped by phoneme.



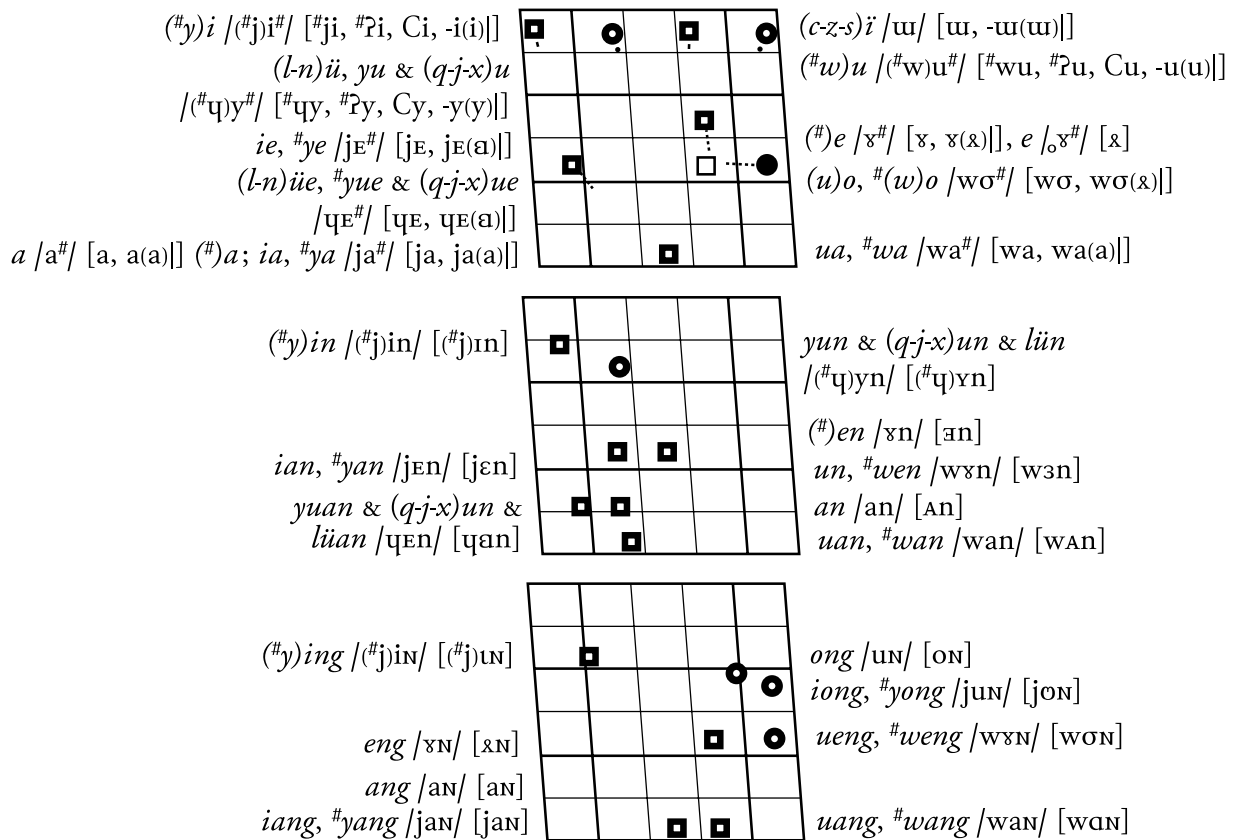
1.1.1. First, we find the close front phoneme /i/: /(#j)i, (#j)in, (#j)iN/ [(#j)i, (#j)in, (#j)iN] (C/y)i, (C/y)in, (C/y)ing: [ˈji] /ˈji/ yī, [ˌpɿ] /ˌni/ nǐ, [ˈjɿn] /ˈjin/ yín, [ˈt͡ɕɿn] /ˈtsin/ jīn, [ˌjɿN] /ˌjin/ yìng, [ˈt͡ɕɿN] /ˈtsin/ jìng. Oscillations are possible and not infrequent: [(#j)u, (#j)iu, (#j)iuN] and [(#j)ɛN, (#j)ieN, (#j)ieN]. The approximant /j/ is seldom kept even when a proper initial is present: in that case, and where deemed necessary, the actual pronunciation of such syllables as *lin(g)*, *pin(g)*, *din(g)* could be better phonemized as /ljin, -N/, /phjin, -N/, /tjin/ (or even /ljɤN, -N/, /phjɤN, -N/, /tjɤN/, which nevertheless should be reserved for more marked variants, as those found in heavily Peking-accented Mandarin).²

1.1.2. Secondly, there is the close front-central rounded phoneme /y/, with a fairly limited distribution: /(#ɥ)y, (#ɥ)yn/ [(#ɥ)y, (#ɥ)yn] (frequent variants with n: [(#ɥ)in,

([#]j)ɤn, ([#]j)ɤən]). The official orthography adopts *ü* only in *nü* and *lǚ*, while leaving *u* unmarked in *ju(n)*, *qu(n)*, *xu(n)*, *yu(n)*. We see no advantage in doing so, and we prefer to keep that friendly *ü* anywhere: [ːt̚ɕy] /ːt̚sy/ *jü*, [ːt̚ɕ̥y̥] /ːt̚shy/ *qü*, [ːt̚ɕ̥y̥] /ːsy/ *xü*, [ːt̚ɕ̥y̥] /ːɥy/ *yü*, [ːt̚ɕɤn, ːt̚ɕɤn, ːt̚ɕɤən] /ːtsyn/ *jün*, [ːt̚ɕ̥ɤn, ːt̚ɕ̥ɤn, ːt̚ɕ̥ɤən] /ːt̚shyn/ *qün*, [ːt̚ɕ̥ɤn, ːt̚ɕ̥ɤn, ːt̚ɕ̥ɤən] /ːsyn/ *xün*, [ːt̚ɕ̥ɤn, ːt̚ɕ̥ɤn, ːt̚ɕ̥ɤən] /ːɥyn/ *yün*.³

More conveniently, *pīnyīn* could have used *zū*, *cū*, *sū*, *yū*, by phonemically interpreting these syllables as we do: [t̚ɕy, t̚ɕ̥y̥, ɕy] /t̚sy, t̚shy, sy/. Indeed, /ɥy/ could even be /jy/; but since the phoneme /ɥ/ is there, it is more adequate to use it. In this way, as it seems convenient, it could also be used to show –even phonemically– the possible oscillations between [ɥyn, jyn, jɤən, ɥn]: /ɥyn/, and ‘/jyn, ɥn/’. (Therefore, it could have been possible to write *ziong*, *ciong*, *siong*, too, more internationally, instead of the official *jiong*, *qiong*, *xiong*, for [t̚ɕjɔn, t̚ɕ̥jɔn, ɕjɔn] /t̚sjɔn, t̚shjɔn, sjɔn/).⁴

fig 1.1-B. Mandarin fundamental vocoids and their taxophones: /V[#], Vn, Vn/.



1.1.3. Then we find a back rounded vowel, /u/ [u] (not to be confused with British English ‘/u:/’ –better: ‘/uu/’– whose various taxophones are rather mid-back, and

often diphthongized): [ˈɡuˈtu] /ˈkuˈtu/ *gūdú*, [ˈwu.ɹu.ɹu] /ɹwu.ɹu/ *wǔrǔ*. Given that we encounter [ʊN] as a secondary taxophone of [oN] /ɔN/, the alternative phonemization ‘/uN/’ must be considered, at least for scientific purposes: [ˈɡʊN\kʊN] /ˈkʊN\kʊN/ & ‘/ˈkʊN\kʊN/’ *gōnggòng*.⁵

1.1.4. Following, we have the open vowel /a/, with its distributions and various taxophones: / (C)a[#], (w/C)an, (j)an, wan/ [(C)a, (w/C)an, (j)an, wan]. A few examples: [ˈpaːɰa, ˌɰa] /ˈatsa/ *āza*, [l̥tʰa] /l̥tha/ *tà*, [ɹa] /ɹan/ *ǎn*, [ˈpʰan] /ˈphan/ *pán*, [ˈpʰan] /ˈphan/ *páng*, [wan] /wan/ *wǎn*, [ˈjan] /ˈjan/ *yáng*, [wan] /wan/ *wǎng*.

As far as //jan, ɹan// are concerned, we think they are better represented by /jɛN, ɹɛN/ (in spite of the official rendering with *an*): [ˈjɛN] /ˈjɛN/ *yán*, [tʃɛɰ] /tʃɛN/ *diǎn*, [ɹɛɰ] /ɹɛN/ *yüàn*, [ˈtʃɰɹan] /ˈtʃɰɹan/ *qián*.

1.1.5. The vocalic inventory of neutral Chinese continues with three mid vowels: /ɛ, ɤ, σ/. Since the interjection *ê* is rather theoretical, the only environment where we find /ɛ/ [ɛ] alone (ie not combined in diphthongs, /ɛi, aɛ/) is in the two sequences /jɛ[#], ɹɛ[#]/ [jɛ, ɹɛ].

For the former, *pīnyīn* writes *ye*, when lacking a graphic initial; otherwise, *ie* /ɹtsjɛ, ˌtʃhɛ, ˌsjɛ, \nɛ, \ljɛ/ [tʃjɛɛ, ˌtʃhɛ, ˌsjɛ, \nɛ, \ljɛ] *jiě, qiě, xié, niè, liè*, and so on. Any combination is permitted, except *fi* and *zhie, chie, shie, rie* (since the post-alveolars cannot match with /j/).

As far as *zie, cie, sie* and *gie, kie, hie* are concerned, they have historically merged into the single series *jie, qie, xie*, seen above. Interestingly, some characters in the Peking Opera, *Jīngjù* /ˌtsɪn\tsy/ [ˌdʒɪn\tsɰ], often chant their lines using a form of stage pronunciation which still retain archaic syllables, like *zie* and *siang*.

For the latter sequence, we have *jue, que, xue* and *yue*, but *lüe* and *nüe* (no other combinations occur, even those –like *müe*– that might be articulated without problems). As with *ü(n)*, ‘our’ *pīnyīn* will always show *üe*, instead: /ɹtsɹɛ, ˌtʃhɹɛ, ˌsɹɛ, \nɹɛ, \lɹɛ/ [tʃɹɛɛ, ˌtʃhɹɛ, ˌsɹɛ, \nɹɛ, \lɹɛ] *jüě, qüē, xüé, nüè, lüè*.

1.1.6. The mid-back vowel /ɤ/ plays an important role in giving Mandarin its distinctive sound, and shows the following distribution: / (#ɹ)ɤ[#], (#ɹ)ɤɹ; (#ɹ)ɤN, wɤN; (#ɹ)ɤN, wɤN/ [(#ɹ)ɤ(ɤ)], (#ɹ)ɤɹ; (#ɹ)ɤN, wɤN; (#ɹ)ɤN, wɤN] *e, er, en, wen|Cun, eng, weng*. The last one has the variant [wɔN] (which, in case, could be indicated as /wɔN/). A

few examples: [ˈkʰɿ\kʰɿ(ɿ)] /ˈkʰɿ\kʰɿ/ *kēkē*, [ˈt͡ʂou˥˥ʔən˥˥lɑɛ, ˈʔɛl˥˥lɑɛ] /ˈt͡ʂou˥˥ʔən˥˥lɑɛ/ *Zhōu Ēnlái*, [ˈʔɿɿt͡ʂjɛɛ] /ˈʔɿɿt͡ʂjɛ/ *érqiě*, [ˈg̊əm˥p̊ɿɿ] /ˈkʰɿn˥pʰɿn/ *gēnběn*, [ˈswɿn˥wɿn] /ˈswɿn˥wɿn/ *Sūn Wén*, [ˈf̊ɿn˥f̊ɿn] /ˈf̊ɿn˥f̊ɿn/ *fēngshēng*, [ˈfu˥wɿn; ˈwɿn] /ˈfu˥wɿn/ *fùwēng*.

1.1.7. Finally, /(#ʔ)σ/ [(#ʔ)σ(ɿ)] *o*, which can occur in the interjections /ʔσ, jσ/ [σ(ɿ), jσ(ɿ)] *o*, *yo*, or be preceded by /w/ in the sequences /mwσ, pwσ, phwσ, (C)wσ/ [m̊σ(ɿ), p̊σ(ɿ), p̊h̊σ(ɿ), f̊σ(ɿ), (C)wσ(ɿ)], *mo*, *bo*, *po*, *fo*, *wo/Cuo*. A few examples: [ˈm̊σ\m̊σ(ɿ)] /ˈmwσ\mwσ/ *mómò*, [ˈb̊σ˥p̊h̊σ(ɿ)] /ˈpwσ˥phwσ/ *bópó*, [ˈlwσ˥twσ(ɿ)] /ˈlwσ˥thwσ/ *luòtuò*, [ˈd͡z̊wσ˥t͡sh̊wσ(ɿ)] /ˈtswσ˥tshwσ/ *zuò cuò*, [ˈɿn˥jσ] /ˈɿn˥jσ/ *hángyō*, [ʔσ(ɿ)] /ʔσ/ *ò!*

As far as (*i*)ong is concerned, a traditional approach favours ‘/(j)un/’, and we consider it as perfectly legitimate (along with ‘/ɥun/’, for certain secondary taxophones). Nonetheless, we have observed that [on] and [jon] are by far the most typical realizations, and the phonemic status bestowed on /σ/ suggests us to switch to the more realistic /σ(ɿ)/. Two examples will suffice: [ˈg̊on\k̊on] /ˈkon\kon/ *gōnggòng*, [ˈt͡ʂj̊on˥t͡ʂj̊on] /ˈtshj̊on˥tshj̊on/ *qióngjiǒng*.⁶

Phonemic conciseness vs realism: the interphonemic choice

1.2. We are perfectly aware that from an intraphonemic point of view, ie just within the Chinese language, [ɿ, ɛ, ɜ, ɶ, ɿ] and [ɛ] could –or should– be considered as taxophones of the same phoneme. Besides, in contact with labials (*mo*, *bo*, *po*, *fo*, *ou*, *wo/Cuo*), [σ] could be considered as a variant of /ɿ/ too, as many Chinese phonologists do (*off* ‘/wɿ/’ or ‘/wɶ/’). In that case, a more ‘Solomonic’ symbol would be the ubiquitous ‘/ə/’, or –even better– the same ‘/ɜ/’ that we have employed in our simplified *Pronuncia cinese per italiani* (‘Mandarin Pronunciation for Italians’) as a practical *passe-partout* for the various occurrences of /ɿ/: [ɜ] (while keeping [ɛ, σ] apart, for the reasons given above).

We are confident in our interphonemic choice to list /ɛ, σ/ among the vowel phonemes of Mandarin Chinese. Anyway, more restricted (certainly, more abstract, because intraphonemic) inventories are possible as well, and –in fact– frequently adopted by many authors. Our approach is different and has been duly motivated, yet our readers might find it interesting to try out and compare various solutions, to evaluate the pros and cons of each one.

1.2.1. Thus, for the sake of completeness and *par condicio*, we present here one of such vowel inventories (coherent, nonetheless, with the same rigorous principles applied to our analysis throughout this book): /i, y; ə, a; u, u/.

Accordingly, our /jɛ, ɥɛ, ɛi; wσ, (j)σu, (j)σN/ should appear as ‘/jə, ɥə, əi; wə, (j)əu, (j)wəN/’ (or ‘(j)uN’, for the last one, as previously noted). In other words, our /ɛ, σ/ would be treated as taxophon(em)es of the ubiquitous ‘/ə/’, when in contact with /i, j/ or /u, w/; inevitably, /aɛ, aσ/ should revert to their traditional renderings ‘/ai, au/’ (though detached from phonetic reality as they can be).

Up to this point, one may feel content with this, and put up with the double or triple mental passages needed to switch from such a highly sketchy *phonematography* to something loosely resembling actual pronunciation, not to mention a passably accurate *phonography*. But some problems arise when trying to render *ê!* and *o!*, two very common interjections. In our analysis, we do not have to think twice: /ɛ, σ/ (usually: è, ò /ɛ, |σ/ [|\ɛ, |\xi; |\sigma, |\varrho, |\zeta], but Chinese dictionaries permit any toneme for both). The point is that interjections and onomatopoeia are written down in Chinese characters too, as any other lexeme, which may lead to think that they can be lexicalized and phonemicized accordingly. This is debatable, since ‘unorthodox’ phones appear sporadically in any language, precisely in interjections and onomatopoeia, but are not generally included in the phonemic inventory.

1.2.2. A glaring example: in Italian, laughter is represented by *ah ah ah!* (or something like that); the basic pronunciation should be something like [ʰ(h)a:haʰa], but we can actually hear anything ranging between [h, h̃, h̄, h̅, x, x̄; a, A, a, e, ɜ, a, ʌ, ʌ̄], just to list a few. None of these belong to neutral pronunciation, of course, except [a], ie the phoneme /a/. Now, how should we treat this? In fact, /h/ is a *xenophoneme* in Italian; as such, it is frequently omitted, and often it has to be: an English loan like *hi-fi* is currently pronounced [aiʰfai], then, while [haiʰfai] would be regarded as pedantic or affected. But *ah ah ah!* is purely Italian; thus, the only solution to represent it is accepting [h] as an *incidental phoneme*, /h/ – to say, a phone that deserves to be exceptionally phonemicized, without actually being a fully functional component of Italian phonology. ⁷

1.2.3. If we want to stick to the ‘/i, y; ə, a; u, u/’ scheme rigidly, there is little choice but retaining / \langle E, σ \rangle / as *incidental phonemes* in Chinese, like / \langle v \rangle / (which, more exactly, is rather a *xenophoneme*). To make it even more captious, one may go as far as treating [E, σ] as *incidental monophthongizations* of ‘/əi, ai/’ and ‘/əu/’ respectively, still failing to transcribe them unequivocally. We presume that our readers have done the maths and drawn their own conclusions: a ‘skeletal’ phonology may appear logical and easy to remember, but it is not necessarily functional and easy to use.

We prefer, instead, to regard ‘/ə/’ as an *archi-phoneme*, and to split it into three different *semi-phonemes*: / \langle E, γ , σ \rangle /. It has to be underlined that these are not the same as *incidental phonemes*: they are true phonemes, instead, but simply not ‘crucial’ ones. In fact, native speakers may perceive no difference between [σ U] and [γ U], for instance, or [Λ E, Aə] and [AI], as well as between [uN, uN] and [oN, oN]; but we insist —once more— that phonology has to be internally coherent, but also as useful and explicit as possible.

That is why we have departed from the *offIPA* tradition of employing only ‘primary’ symbols at the phonemic level; and, according to the most adamant traditionalists, not merely ‘primary’, but as near to their graphemic counterparts as possible, too! Consequently, we should have even used ‘/e/’ instead of the nearly-passable ‘/ə/’, since there is no opposition between [e, E, ε, ɜ, γ] &c!

Fortunately, the latter practice has been dismissed, but the former still rules the land: as a consequence, most English dictionaries continue to employ ‘/e/’ to represent modern British [E], for which ‘/ε/’ would be far better (but it is finally —and happily— gaining popularity among scholars).⁸

1.2.4. In some analyses of Chinese, we find sober compromises like ‘/e, γ , o/’ or, better, ‘/ε, γ , γ /’, but also masterpieces of schizophrenia, such as ‘/e, γ , γ /’ or ‘/ε, γ , o/’ — and even ‘/ε, ə, γ , o/’, mixing phones and phonemes! This is highly confusing for foreign learners, who often rely on these poor examples of *offIPA* misuse to build their own Mandarin phon(em)ic inventory, when first attempting to learn *pīnyīn* and try to pronounce the language.

For example, native speakers of French, Italian, Portuguese and German consider /e/ as actually [e], not certainly ‘anything in between [e] and [æ]’ (and /ε/ is [ε] as well, in those languages; neutral German even distinguishes between ‘/e:, 'ε:,

'ɛ/ [ɛ, ɛ̃, ɛ̂]'.⁹ What foreign learners really need is a phonemic scheme devised according to *interphonemic principles*: realism and multilingual applicability. In practical terms, this means choosing, for each phoneme, the symbol that encompasses all the taxophones (or may be able, at least, to suggest most of them), without discriminating 'primary' symbols against 'secondary' ones.

1.2.5. In *canIPA*, there is no such thing: *a symbol is a symbol*, and if it is best for its task, we employ it. So, even at first glance, Mandarin (but also English and Spanish) /ɛ, σ/ suggest that the learner is dealing with something different from [e, o] and [ɛ, ɔ], while at the same time implying a certain auditory and articulatory resemblance with both pairs. Besides, while our /ɛ/ is able to cover successfully [ɛ, ɛ̃, ɛ̂], '/e/' makes it necessary to revive the unrealistic '/jan, ɣan/', to explain [jɛn, ɣan] *yan, yuan*.

Multilingual applicability is the other advantage of the *interphonemic* approach, in that it helps polyglots (or those who aim at being one) and phoneticians build a sort of personal, constantly growing 'phon(em)ic mindset'. They can resort safely to it, when attempting to learn –or analyze– another language, or improve their own pronunciation of one they already know.

Just mastering [e, ɛ, ɛ̃] and [o, σ, ɔ] helps us avoid gross *faux pas* in most European languages and Japanese (and Chinese, as we have seen!), but also improves our understanding of different accents. For instance, English *awe* /'σ:/ is [ʰσ:] in modern Standard British, but [ʰσ̄] in Mediatic British, and [ʰɔ:] in most American accents; to an untrained ear accustomed to hear and speak only American English, [ʰσ̄] may easily suggest *owe* /'σ̄/ [ʰσ̄] at first, but once the listeners are able to detect [o] in isolation, they would promptly recognize a mono-timbric diphthong like [ʰō], and distinguish it from [ʰσ̄].¹⁰

On the phonemic status of -i.

1.3. This last vowel phoneme requires a longer, but necessary, explanation, almost a chapter on its own. This is due to the fact that many Chinese scholars keep a conservative approach to Mandarin phonetics, being largely influenced by the assumptions of some pioneers in the field, the Swedish sinologist Karlgren among them. It is nearly given for granted, thus, that Mandarin has two *retroflex* –or *apical*– vowels. In our view, both belong to the same phoneme.

We strongly support ‘/ʍ/’ as the phonemicization of what *pīnyīn* transcribes as *i* after *z(h)*, *c(h)*, *s(h)* and *r*. As explained below –cf *Consonants*– we adopt *ī* instead, in order to mark, even at the graphemic level, and in parallel to /ʍ/, that we are dealing with another fundamental component of Mandarin phonology, not with taxophones of /i/. Our choice falls upon /ʍ/, as this appears to better represent all the taxophones emerged from long and accurate analyses of numerous recordings of neutral Mandarin voices (especially those of professionally trained speakers).

1.3.1. Many –if not all– Sinologists favour two symbols not enclosed in the *offIPA* inventory: ‘ɿ, ʅ’; the former represents *ī* in *zī, cī, sī* — the latter, *ī* in *zhī, chī, shī, rī*. Though called ‘vowels’, both are actually presented as ‘syllabic’ approximants (we prefer ‘intense’), totally assimilated to the point of articulation of the contoid that comes before them. They are often –but not systematically– considered as two distinct phonemes, though they clearly are in complementary distribution.

A very similar view is shared by the International Phonetic Association, which, in the 2003 issue of its *Journal*, proposed ‘ɹ̥’ as a unified expedient to represent both ‘ɿ, ʅ’. The official descriptions of these two phon(em)es, then, are ‘(apico)laminal denti-alveolar’ approximant and ‘apical postalveolar’ approximant, respectively. None can be entirely accepted.

1.3.2. Correctly, no groove is assigned to the first contoid, which nevertheless cannot be described as a full approximant (in our system, the proper symbol would be [ʂ]), nor is it totally assimilated to /ts(h), s/: in that case, we should perceive a continuous ‘buzz’, though lighter than that of [z].

More realistically, we have a vocoid, [ʍ], *with negligible assimilation* to the preceding segment: in fact, just a very short portion of the vocoid –its contiguous part to /ts(h), s/– exhibits a certain degree of *co-articulation*, but this is absolutely automatic, and need not be indicated in phonetic transcriptions, let alone in a *phonemic* one! We admit that an *intense semi-approximant*, [ʂ], may result *auditorily* quite similar to [ʍ], but the *articulatory* nature of this phone is primarily vocalic.

To sceptical ^{6%} traditionalist readers, we suggest a simple experiment, provided they can pronounce Chinese at least decently: articulate a long, slow and emphatic *sī*, paying attention to any transition from *s* to *ī*; if this is done correctly, the result should resemble something like [[s:zzzʂʂʍ]. ¹¹

This being said and recognised, yet we have to determine what really matters –ie what is important and ‘typical’– and what is simply the outcome of inevitable (but secondary) adjustments. From this point of view, it appears reasonable to skip even intense (or ‘syllabic’) [ʒ, z, ʒ̥, z̥, δ, ɔ̃] and to keep only [ʉ] as the real distinctive element of *sī* [su], which we therefore phonemicize as /su/ (with /ts(h)u/, too), discarding ‘ɿ’ and similar expedients.

1.3.3. A slightly different criterion must be applied to /çu/, ie when /u/ follows the post-alveolar initials: *zhǐ, chǐ, shǐ, rǐ*. In this case, the automatic assimilation affects the entire articulation of /u/, and cannot be regarded as a merely transitional event: even an untrained ear can detect a continuous ‘buzz’, incomplete –or even absent– in /u/; auditorily, the latter has a very different timbre too, even if the mid-dorsum generally remains in the same ‘box’ as [u].

Much debated is the true nature of this /çu/. As we explain hereby, neutral pronunciation actually admits more than just one realization: it is surprising that, instead of simply recognizing this fact and providing all the relevant transcriptions, *every single* analysis that we have examined insists in presenting only one. Inevitably, very different phonic renderings have been proposed so far, depending on which sources the authors took into account, and how restrictive their concept of ‘neutrality’ is.

For example, in his *Hànyǔ yǐyīn jiàochéng*, Cáo Wén often refers to how ‘Peking speakers’ would pronounce something, as if the *pǔtōnghuà* spoken in Peking were to be assumed as a model for neutral pronunciation. Undoubtedly, Pekingese speakers have the best chances to conquer a purely neutral pronunciation, considering that their local dialect served as the basis for Modern Chinese; yet, without proper training, anyone with an even slight Peking accent would be spotted immediately, sometimes by foreigners as well.

Nonetheless, the Peking accent is considered a ‘prestigious’ one, due to its connection with the capital city and its presence in the media, thus qualifying as what we call a ‘mediatic’ accent: not neutral, for sure, but accepted by most as nearly-neutral, and no doubt more widespread than a real neutral accent, which has to be learnt expressly.

With few exceptions, those handbooks where ‘ɿ’ is still employed describe it as something of the same nature as ‘ə’, apparently, but ‘narrower’... whatever this

means. Nevertheless, when one examines the relevant illustrations proposed by the same handbooks, they usually show a curled-shaped tongue, with very little space between its tip and the post-alveolar ridge. It is clear that, by ‘retroflex vowel’, the authors actually mean an *intense post-alveolar contoid*.

As far as the exact nature of the contoid is concerned, opinions vary, but the dominant position is that summarized by Cáo and others as ‘a less fricative [ɰ]’.¹² Provided ‘less fricative’ does not stand for a *semi-constrictive* phone (^{can}[ʒ]), we can safely infer that an *intense post-alveolar approximant* is actually meant (^{can}[ʝ]), in accordance with the officially sanctioned description.

1.3.4. In fact, such a realization is very common in neutral Mandarin (as well as in mediatic Pekingese) and, in general, in all northern China. It can be adopted safely by foreigners learning the language, provided they refrain from that most tiresome habit of curling their tongues so backwards, that they do not produce a true *post-alveolar* contoid anymore, but rather an *apico-palatal* one: [ʝ].

It is a speech defect –or, at best, a form of *hypercorrection*– that even a certain number of Chinese exhibit, especially when their native dialect contemplates no [ʑ], and they are willing to ‘speak proper’: in doing so, they literally *retroreflect* their apex, instead of raising it gently behind the alveolar ridge, ‘just to be on the safer side’. The resulting phones [ʈ, ʂ, ʐ] &c are normal in many Dravidian languages, for example, but not in Mandarin, where they arouse an auditory impression of ‘jabbering’ instead, if used as initials.

Another peculiarity to be avoided is splitting [ʝ] into a sequence like [wʐ] (or worse, [wʑ], even elongated in [wʑ:], which might be misunderstood as /ʂʐ/). This trait is pretty common among native speakers of ‘rhotic’ English accents, typically those from the United States and Canada, in particular when they have just started learning Mandarin. Not rarely, instead of [ʐ], these speakers would actually use any one of their most familiar *r* phones, namely [ɹ, ɹ̥, ɹ̄, ɹ̥̄] &c (with their ‘intense’ counterparts, [ɹ̄, ɹ̥̄, ɹ̄̄, ɹ̥̄̄] &c, for /ʐ/). These contoids act quite well as surrogates, but produce a fairly recognizable accent: those who aim at a more convincing pronunciation, then, should choose [ʝ].¹³

1.3.5. The intense post-alveolar approximant, though, is not the preferred choice of professionally trained speakers. Instead, we have noticed a consistent use of a

somewhat ‘milder’ phone, precisely [u̥], which –for the sake of conciseness– we shall describe here as a *laterally contracted* [u]. Since we have detected an entire set of such *latero-vocoids* as natural realizations of the so-called *érhuà* phenomenon, we suggest our readers to refer to the relevant § 2, where we describe all the features involved in this –quite complex– co-articulation process.

Here, suffice to say that the presumable reason why professional voices favour [u̥] is both of an *articulatory* and *euphonic* nature. By comparing the orograms of [ʒ] and [u̥], it is quite clear that the latter phone requires little –if any– movement of the apex, yet the auditory result is not very different from that of [ʒ]. Therefore, in connected speech, the apex may pass from one articulation to another with relative ease, especially when the front half of the tongue is crucial (as in /i, y; j, ɥ; t(h), ts(h), s, l, n/ &c).

Subjective and questionable as it can be, euphony certainly does play a role, even if [u̥] was not selected on purpose, nor was it ‘genetically engineered’ to improve the Chinese phonetic inventory. It is a natural taxophone, instead, and it has always been there: but those for whom a beautiful and mellow voice is the essence of a job well done –ie dubbers, actors, broadcasters, singers, and the like– know perfectly what sounds better... and instinctively adopt it.

It is worth noticing that using [u̥] suggests no far-fetched ‘stage pronunciation’ at all, like –for example– [r, ʀ] in old-fashioned French and German theatrical delivery (even though an alveolar *r* did not belong to the neutral pronunciation of those languages anymore). For this reason, we argue that foreign speakers should at least attempt to learn and use [u̥] instead of [ʒ] (or alternating between them).

There is a certain timbric difference between [u] and [u̥], though mostly as a side effect of the co-articulation phenomena occurring in the latter (lateral contraction, in the first place, primarily responsible for the characteristic ‘buzz’ typical of any [ʏ]; not so much the post-alveolarity, which plays just a marginal role). Apart from this, [u] and [u̥] share almost the same mid-dorsum position.

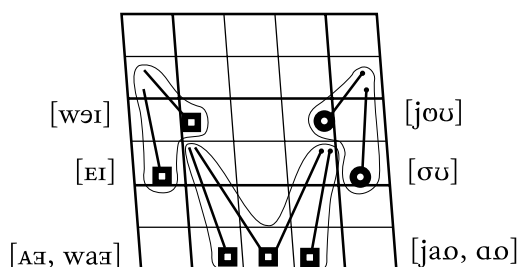
1.3.6. Nonetheless, we have noted a more advanced taxophone, not very common, but adopted by professional voices, as well: some examples of it can be heard at the beginning of the Chinese dubbing of *Hiroshima mon amour* (Alain Resnais, 1959), which indeed sounds somewhat awkward to native listeners. Scattered occurrences of the same taxophone are found in various sound files attached to the

late 70's *Practical Chinese Reader*; a few present-day tv anchormen use it occasionally too. A separate symbol is not needed, but, for truly fronted realizations, something like ‘[i̟]’ might be useful. In fact, kinesthesia reveals that it is possible to articulate a fully functional [ʏ] as front as in the area of [i], but no fronter than that, otherwise it would be very hard to contract both sides of the tongue and, at the same time, raise the apex.

Diphthongs

1.4. Modern Mandarin phonology permits only four diphthongs (fig 1.2): /ɛi, aɛ, aɔ, ɔu/ [ɛi, aɛ, aɔ, ɔu]. The former two may be preceded by /w/, the latter two by /j/, with slightly different vocalic traits: /wɛi, waɛ, jaɔ, jɔu/ [wɛi, waɛ, jaɔ, jɔu]. We insist that /i, u/ (and /y/) are always *vowels*, while /j, w/ (and /ɥ/) are always *approximants*, thus *consonants*, not miraculous ‘semi-something’; consequently, we cannot help disapproving of such renditions as ‘/ej, aj, aw, ow/’ or ‘/uei, uai, iau, iou/’ either (see our reference to /j, ɥ, w/, § 3.5.3).

fig 1.2. Mandarin diphthongs.



Pīnyīn duly represents simple diphthongs as *ei, ai, ao, ou* (though *ae* would have been more realistic, and more coherent with *ao* too); *wei, wai, yao, you* are also very reasonable. When *wai* and *yao* follow another initial, their spelling change into *uai* and *iao*, making grapho-syllabification easier in some cases: *chēnguài* ‘to blame, to rebuke’ is unmistakably /tʂhɤn\kwɛɛ/, while *chéngwài* ‘outside of a town or a city’ leads automatically to /tʂhɤn\wɛɛ/; otherwise, *pīnyīn* would abound in apostrophes or hyphens, like older romanization schemes: *chēn’guài* vs *chéng’uài*, or *chēn’gwài* vs *chéng’wài*.

1.4.1. But, then, we find absurd inconsistencies: instead of changing *wei, you* into *uei, iou*, the official orthography prescribes *ui, iu*, with no real advantages for

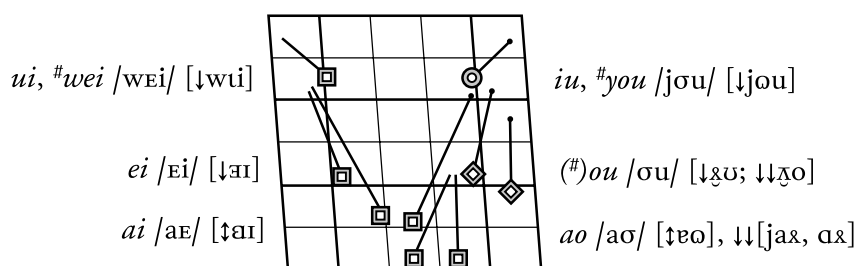
the learners. In all honesty, it is odd even at the typographical level: lacking the real *nuclei*, we are forced to put the relevant diacritics on the syllabic *codae* instead, ie $u\bar{i}$, $i\acute{u}$, $u\check{i}$, $i\grave{u}$, rather than the more friendly and correct ‘ $u\bar{e}i$, $i\acute{o}u$, $u\check{e}i$, $i\grave{o}u$ ’!

Besides, *ui* and *iu* easily lead foreign students to think that / $w\bar{e}i$, $j\acute{o}u$ / are actually to be reduced to ‘/wi, ju/’ in composite syllables. We can personally testify that even some Chinese teachers at the Beijing Language and Culture University are convinced of such a ‘reduction’, so strong does *pīnyīn* influence their understanding of Mandarin phonology. To prove their point, they would refer to certain secondary variants, where –nonetheless– we do not hear anything like ‘[wi, ju]’, but rather [w $\bar{u}i$, $j\acute{o}u$] (and [ɛ \bar{u} , $\xi\bar{u}$] for / $\bar{e}i$, σu /, see fig 1.3).

Such variants belong to (nearly) neutral pronunciation as well: foreign listeners should be able to recognize them readily, but they are no way necessary for a good pronunciation, nor do they represent a true exception to the rule. Only if we really needed a different phonemicization (merely of academic interest, by the way), a compromise like ‘/wii, juu/’ would be somewhat worth consideration.

A few examples (with variants, where present): [ː $\bar{f}e\bar{i}$; ː $\bar{f}e\bar{i}$] / $\bar{f}e\bar{i}$ / $f\bar{e}i$, [ː $\bar{t}w\bar{e}\bar{i}$; ː $\bar{t}w\bar{u}i$] / $\bar{t}w\bar{e}i$ / $du\bar{i}$, [ː $\bar{w}\bar{a}\bar{i}$; ː $\bar{w}\bar{a}\bar{i}$] / $\bar{a}e$ / $\acute{a}i$, [ː $\bar{w}\bar{a}\bar{i}$] / $w\bar{a}e$ / $w\grave{a}i$, [ː $\bar{a}\bar{o}$; ː $\bar{a}\bar{o}$] / $\bar{a}o$ / $\acute{a}o$, [ː $\bar{a}\bar{o}$; ː $\bar{a}\bar{o}$] / $\bar{a}\bar{o}\sigma$ / $h\grave{a}o$, [ː $\bar{j}\bar{a}\bar{o}$] / $\bar{j}\bar{a}\sigma$ / $y\acute{a}o$, [ː $\bar{\sigma}\bar{u}$; ː $\bar{\sigma}\bar{u}$] / $\bar{\sigma}u$ / $\bar{O}u$, [ː $\bar{m}\bar{o}\bar{u}$; ː $\bar{m}\bar{o}\bar{u}$] / $\bar{m}\sigma u$ / $m\grave{o}u$, [ː $\bar{t}\bar{j}\bar{o}\bar{u}$; ː $\bar{t}\bar{j}\bar{o}\bar{u}$] / $\bar{t}\bar{j}\sigma u$ / $di\bar{u}$. Less advisably: / σu / [ː $\bar{\sigma}\bar{o}$], / $(j)\bar{a}\sigma$ / [ː $\bar{j}\bar{a}\bar{x}$, $\bar{a}\bar{x}$].

fig 1.3. Mandarin diphthongs (variants).



Vowel reductions

1.5. Mandarin vowels may undergo some timbric reductions $\%$ or shortenings, mainly when the relevant weakly-stressed or unstressed syllables occur in positions different from word-final ones, or in intonemes. As a matter of fact, in actual language, apart from in slow and accurate speech, some neutralizations occur elsewhere too. It is fundamental to know such reductions, at least to be able to understand spoken Chinese well. Foreigners need not use such reductions actively, al-

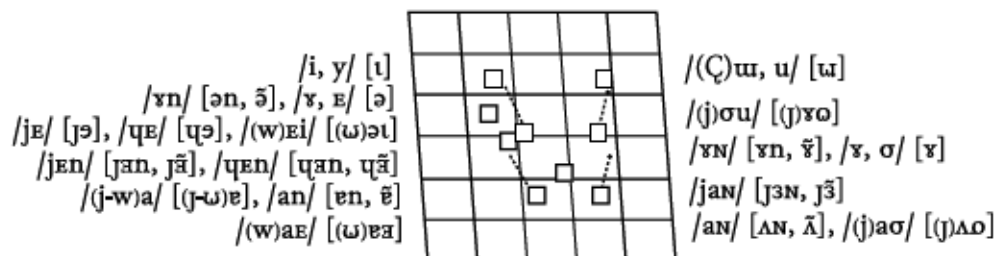
though an appropriate use of them would decidedly improve their pronunciation. In order to adequately show the behavior of different unstressed syllables (for vowel, consonant, and pitch reductions), sometimes we also had to use some rare words, rather than omitting something significant.

1.5.1. fig 1.4 shows the most important reductions which *may* occur in fast % casual speech, and in connection with the so-called *qīngshēng*, /o\$/, or ‘zero toneme’. This does not imply ‘absence of tone’, as many Chinese themselves think (and teach foreigners, too!), but simply a feeble tonal contour as that found in syllables which, being unstressed, lose their original (or theoretical) ‘full’ toneme.

The examples given are to be interpreted as in a sort of decreasing scale, starting from a pre-pausal position, in an intoneme (for the first variants), up to modifications produced within a sentence, in a pre-intoneme (for the other variants). For /Vn, Vn/, the variants with [Ṽ] have been included in the vocograms, but not in the examples; it would have been a sterile repetition, after all, as the vocoid remains in its ‘box’, though nasalized. But it is important to note it, because the vocalic timbre remains the only discriminant, when [Vn] and [Vn] merge into [Ṽ].

Anyway, in order not to make our transcriptions more complex than necessary, we will not employ such reductions ordinarily. On one side, they do not occur automatically, nor universally; on the other, foreign learners –to whom our ‘normalized’ transcription is primarily aimed– can safely use normal vocoids anywhere (but they should be able to recognize reduced forms, when native speakers use them).

fig 1.4. Vowel reduction in /o\$/ (in quick speech, not in an intoneme; whereas, in an intoneme, the distinct timbers of fig 1.1-2 are used).

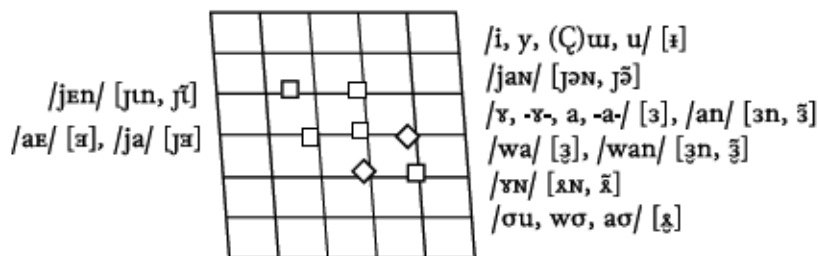


1.5.2. As can be seen, [i] applies to /i(n/N), y(n)/: [t̚d̚d̚ -d̚] *dìdì*, [k̚wɑŋ·ʒi; -ʒi] *guānxi*, [t̚jɑ̄.t̚h̚ȳ; -t̚t̚] *xiàqiū*; and [u], to /u, u/ (with [u] for /C(u)/): [j̚z̚z̚; -z̚] *yìsì*, [l̚ʒ̚ŋ̚.ʒ̚z̚; -z̚z̚] *rènsì*, [s̚u·y̚; -v̚u] *shūfu*.

In addition, we have [ə], for /jɛ, ɥɛ, (w)ɤN/, and [əɪ] for /(w)ɛi/: [ʔɛjɛ̂jɛ̂ʔɛ̂jɛ̂; ʔɛ̂jɛ̂, ʔɛ̂jɛ̂] *jiějìe*, [ʔɛ̂jɛ̂jɛ̂jɛ̂; ʔɛ̂jɛ̂jɛ̂, ʔɛ̂jɛ̂jɛ̂] *àiren*, [ʔmɛ̂jɛ̂mɛ̂jɛ̂; ʔmɛ̂jɛ̂] *mèimeì*; then [ɤ], for /ɤ, (w)ɤN, (w)σ/, and [(j)ɤO], for /(j)σu/: [ʔkɤ̂ĝɤ̂; ʔĝɤ̂] *gēge*, [ʔtɤ̂jɛ̂jɛ̂; ʔɛ̂jɛ̂] *dǐngzhe*, [ʔɛ̂jɛ̂jɛ̂jɛ̂; ʔɛ̂jɛ̂] *xiānsheng*, [ʔɛ̂jɛ̂jɛ̂m̂jɛ̂; ʔm̂jɛ̂] *jièmò*, [ʔnwɤ̂jɛ̂jɛ̂; ʔjɤ̂] *nuǎnhuo*, [ʔjɤ̂jɛ̂jɛ̂jɛ̂; ʔjɤ̂jɛ̂] *yòngtóu*, [ʔphɤ̂jɛ̂jɛ̂jɛ̂; ʔjɤ̂jɛ̂] *péngyou*.

Lastly, we have: [ɐ; ɐɪ] in /(w)a(n); (w)aɛ/: [ʔpɐ̂jɛ̂jɛ̂; ʔbɐ̂] *bàba*, [ʔmâmâ; ʔmɐ̂] *māma*, [ʔɛ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔjɤ̂jɛ̂] *xìhuan*, [ʔthɐ̂jɛ̂jɛ̂jɛ̂; ʔtɐ̂jɛ̂] *tàitai*; and [ʌ; ʌO] in /(j)aN; (j)aσ/: [ʔwɐ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔjɤ̂jɛ̂] *wǎnshang*, [ʔtɐ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔdɐ̂jɛ̂] *zhīdao*.

fig 1.5. Further reductions in syllables with ‘zero’ tones (in even quicker speech, not in an intoneme; then, in an intoneme, the realizations of fig 1.4 are also possible).



1.5.3. In even quicker speech, and not in an intoneme, further reduction are possible (fig 1.5): [ɪ] for /i, y, u, u; ʔ/ *ī, ì, u*: [ʔɛ̂jɛ̂jɛ̂jɛ̂jɛ̂ -dɛ̂, -dɛ̂] *dìdì*, [ʔkwɐ̂jɛ̂jɛ̂jɛ̂; ʔzɛ̂, -zɛ̂] *guānxi*, [ʔɛ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔtɛ̂jɛ̂, -tɛ̂jɛ̂] *xiàqu*, [ʔjɤ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔzɛ̂jɛ̂, -zɛ̂jɛ̂] *yìsì*, [ʔzɛ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔzɛ̂jɛ̂, -zɛ̂jɛ̂] *rènshǐ*, [ʔɤ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔvɛ̂jɛ̂, -vɛ̂jɛ̂] *shūfu*; and [ɤ] for /σu, wσ, aσ/: [ʔnwɤ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔjɤ̂] *nuǎnhuo*, [ʔjɤ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔtɤ̂jɛ̂, -tɤ̂jɛ̂] *yòngtóu*, [ʔphɤ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔjɤ̂jɛ̂, -jɤ̂jɛ̂] *péngyou*, [ʔtɤ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔdɤ̂jɛ̂] *zhīdao*.

Besides, [ɛ̂] in /aɛ, ja/ [ɛ̂, jɛ̂]: [ʔthɐ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔtɐ̂jɛ̂] *tàitai*, [ʔɛ̂jɛ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔɛ̂jɛ̂jɛ̂jɛ̂jɛ̂] *rénjia*; and [ɜ] for /ɤ/ an for any other /a/ a: [ʔkɤ̂ĝɤ̂; ʔĝɤ̂] *gēge*, [ʔtɤ̂jɛ̂jɛ̂jɛ̂; ʔɛ̂jɛ̂] *dǐngzhe*, [ʔɛ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔɛ̂jɛ̂] *àiren*, [ʔɛ̂jɛ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔɛ̂jɛ̂] *xiānsheng*, [ʔpɐ̂jɛ̂jɛ̂jɛ̂; ʔbɐ̂, -bɐ̂] *bàba*, [ʔmâmâ; ʔmɐ̂, -mɐ̂] *māma*, [ʔɛ̂jɛ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔjɤ̂jɛ̂jɛ̂jɛ̂jɛ̂, ʔjɤ̂jɛ̂jɛ̂jɛ̂jɛ̂] *xìhuan*, [ʔwɐ̂jɛ̂jɛ̂jɛ̂jɛ̂jɛ̂; ʔzɛ̂jɛ̂jɛ̂jɛ̂jɛ̂] *wǎnshang*.

NOTES:

¹ It is indispensable to read through and carefully consider CANEPARI, L. (2007) *Natural Phonetics & Tonetics – articulatory, auditory, functional*: ① 1-7, to understand the basics of ^{can}IPA and of the ‘natural’ method; and ① 8, entirely devoted to *vowels* (vocalic phonemes) and *vocoids* (vocalic phones).

² See LIN, H-Y. (2007) *The Sounds of Chinese*, § 3.4.2 & 8.1.5.

³ As in CAO, W. (2008) *Chinese Pronunciation Practice*, p. 12. In this booklet, the author seems to have resorted to such a kind of ‘improved’ *pīnyīn*, precisely to avoid IPA as much as possible, in teaching beginners the fundamentals of Chinese pronunciation. During a conversation with him, in 2008, the author frankly admitted that many foreign students were somewhat distracted by the ‘cohabitation’ of *pīnyīn* and phonetic symbols on the page. Therefore, given the tiny amount of time devoted to the methodic learning of pronunciation, and the huge amount of time and energy required to assimilate Chinese characters, he had to make a choice, and favoured *pīnyīn* over IPA.

In our view, this does not prove at all that phonetic transcriptions are superfluous, or even detrimental, but that they must be the logical outcome of a serious *preparatory stage* supported by a scientific approach, not a purely decorative embellishment, ‘just to make it look better’. In this sense, Cao’s choice was right, but not the idea that pronunciation can be really learnt through, without the aid of a well meditated and realistic transcription *with phonetic symbols*.

⁴ In CANEPARI, L. & CERINI, M. (2008) *Pronuncia cinese per italiani*, § 7.21, we proposed a similar romanization scheme (more for scientific interest than for practical use).

⁵ Native speakers of many English dialects, and German (just to mention two major languages) may find it easier to approach [ʊŋ] than [oŋ, ɔŋ]. What really matters is that the speaker keeps a certain consistency.

⁶ cf CAO, W. (2004) *Hanyu yuyin jiaocheng*, p. 84: the author adopts ‘[iuŋ]’ as the basic rendition, adding nonetheless that ‘in Peking, some speakers would also use different pronunciation, such as [yŋ] and [iuŋ]’, ie our ^{can}/qun, jun/ respectively.

⁷ See CANEPARI, L. (2004) *Manuale di Pronuncia Italiana*, particularly ¶ 2-6; and, more concisely, CANEPARI, L. (2007) *A Handbook of Pronunciation*, ¶ 3 ‘Italian’.

⁸ See CANEPARI, L. (2009) *English PronunciationS*, ¶ 1-12.

⁹ See CANEPARI, L. (2007) *A Handbook of Pronunciation*, ¶ 5 ‘German’.

¹⁰ See also CANEPARI, L. (2009) *English PronunciationS*, ¶ 53 & 55.

¹¹ In this time of majestic technological achievements, many instrumental (‘acoustic’) phoneticians underestimate the importance of kinesthesia and proprioceptive abilities. To those interested in ‘redeeming’ themselves, we may suggest CATFORD, J. C. (1988) *A Practical Introduction to Phonetics*.

¹² See CAO, W. (2008) *Chinese Pronunciation Practice*, p. 25 & 32.

¹³ We strongly suggest to read through CANEPARI, L. (2009) *English PronunciationS*, ¶ 26 ‘The <whole truth> on English r’.

2. The *r* grammeme and the so-called *érhuà* phenomenon

2.1. As a result of an evolutive process, which lasted for centuries, the north-western dialects (Pekingese and Standard Chinese itself among them), developed a series of monosyllabic unstressed suffixes, still in use today, like *zǐ* and *li*...¹ Being very frequently used, these suffixes were progressively eroded, until they reduced to a mere *-r* coda; at first, it retained a certain syllabicity (probably as an intense con-toid, [Ç]), but its precise articulation is still under debate. Later, that coda merged with the ‘rhyme’ of the preceding syllable, modifying some of its components too, eg its vowel quality. Historically, such a process had been deeply rooted in the language well before the seventeenth century, when European missionaries described Chinese pronunciation in detail, for the first time.²

This phenomenon is called *érhuà* in Chinese, literally ‘*r*-ization’; we could render it as ‘rhotacism’, which is partially adequate, or ‘syllabic retroflexion’, which is partially inadequate; it might be better to keep the original term (as with *stod*, when describing Danish pronunciation), or to resort to something like ‘*r*-colouring’ if really needed. *Érhuà* not only influences the phonic outline of a lexeme; moreover, any *r* has its own semantic and even grammatical relevance: it is a true grammeme, then, with a phonic signifier and various realizations, as we will see right away.

2.1.1. A small number of verbs (including ‘attributive’ verbs, corresponding to our adjectives) become nouns by adding *-r*:

bāng /p̄paN/ [p̄paN] ‘to help’ vs *bāngr* /p̄paNɹ/ [p̄pã̃] ‘gang’;

gān /k̄kaN/ [k̄kaN] ‘to be dry’ vs *gānr* /k̄kaɹ/ [k̄kɛ̃] ‘dried food’;

huà /x̄kwa/ [x̄kwa] ‘to depict’ vs *huàr* /x̄kwaɹ/ [x̄kwɛ̃] ‘a painting’;

huó /x̄kwɔ/ [x̄kwɔ(x)] ‘to live’ vs *huór* /x̄kwɔɹ/ [x̄kwɔ̃] ‘activity’;

shù /ʃu/ [ʃu] ‘to stand vertically’ vs *shùr* /ʃwɔɹ/ [ʃwɔ̃] ‘vertical brush stroke in Chinese calligraphy’.

2.1.2. Besides, a few noun minimal pairs are distinguished according to the same criterion:

tóu /^hthou/ [ʰtʰou] ‘head’ vs *tóur* /^hthouʒ/ [ʰtʰouʒ] ‘chief’ (interestingly, *chief* derives from Latin *caput* ‘head’, through French *chef*);
xìn /sin/ [ɕi̯n̩] ‘letter’ vs *xìnr* /sjɛʒ/ [ɕjɛʒ̥] ‘information, news’;
bàntiān /pan̄-thjɛn/ [!b̩ʌ̃n̄-thjɛn] ‘a long time’ vs *bàntiānr* /pan̄-thjaʒ/ [!b̩ʌ̃n̄-thjɛʒ] ‘half a day’;
māoyǎnr /maoʒjaʒ/ [ˈmaoʒjɛʒ̥] ‘peephole’ vs *māoryǎn* /maoʒjɛn/ [ˈmaoʒjɛʒ̥n̩] ‘opal’.

2.1.3. Finally, an *r*-suffixed word may easily suggest a diminutive, reductive %_r affectionate acceptance:

bīnggùnr /pin̄|kwɛʒ/ [ˈb̩ɪn̄|k̥wɛʒ̥] ‘ice lolly’;
nánhàir /nan̄ʰlaʒ/ [ˈn̩ʌn̄ʰlɛʒ] ‘baby boy’;
ménfèngr /mɛn̄|fɛn̄ʒ/ [ˈm̩ɛn̄|f̥ɛ̃ʒ̥] ‘a tight crack between a door and its frame, or between two doors’;
xiǎo Zhāngr /sjaō-tʂan̄ʒ/ [..ɕjaō-t̥ʂã̃] ‘young Zhāng’;
lǎo māor /laō-maʒ/ [..l̩aʒ̄-maʒ] ‘old dear kitten’;
àiqíng /ʔaēʰtʂhɛn̄ʒ/ [!:(ʔ)ʌ̃t̥ʂh̥j̃ã̃] ‘unimportant love’;
méi shìr /m̩eī|ʃɛʒ/ [ˈm̩eī|ʃ̥ɛʒ̥] ‘no problem’;
xiǎoshuōr /sjaō-ʃwoʒ/ [..ɕjaō-ʃwõ] ‘novelette’.

2.2. With all its taxophones, *érhuà* provides one of the most prolific fields of research on neutral Chinese pronunciation: if a simple addition of [ʒ] to any syllabic rhyme $-V(V)$, Vn and $Vng-$ might be sufficient for an ‘International’ kind of pronunciation to sound understandable and convincing, phonetic reality is –as usual– quite more complex.

Only certain phones are able to coexist with [ʒ], which is primarily a *postalveolar approximant*, [ʒ]; any foreign speaker of Chinese may feel content with it, and employ it everywhere safely, including for *érhuà*. This is a wise and practical choice, totally loyal to the natural method. Nevertheless, audio files have revealed such a wide range of taxophonetic oscillations, that it is no longer possible to passively accept what the ‘phono-sinological’ tradition has nearly passed off as a sacred dogma.

2. The *r* grammeme and the so-called *érhuà* phenomenon

2.2.1. For decades, nonsensical and frankly embarrassing transcriptions like ‘[Vɾ]’ have been proposed in handbooks and essays, nothing being said about that curious mixture of *pīnyīn* and *IPA*... and sometimes the former prevails over the latter! Luckily, some ‘enlightened’ scholars at least resort to less shameful solutions, eg ‘/Vɾ/’ or ‘[Vɾ]’, but still inadequate both phonemically and phonetically. Such inadequacy derives both from the conceptual and typographical limits of the *offIPA*, and from the fact that *érhuà* is still, too often, treated vaguely.

What we hear from professional voices (actors, dubbers and broadcasters) shows just that problematic sequences are systematically avoided, and adjusted to the few ones allowed. Furthermore, an actual [ʒ] rarely occurs in neutral *érhuà*, since it would require a considerable effort if methodically applied to continuous speech. Instead of a vocoid plus a truly postalveolar approximant (with potential nasalization, as well), [Vʒ, $\tilde{V}\tilde{z}$], in truly neutral and spontaneous pronunciation –neither emphatic nor mannered– /Vʒ/ sequences are more often realized as [V $\underset{\sim}{V}$, $\tilde{V}\tilde{\underset{\sim}{V}}$], ie a simple vocoid followed by a *provelar* (: midback) *vocoid with lateral contraction*.

2.3. As shown in fig 2.1, the apex is *slightly* raised behind the alveoli, certainly not ‘curled’, and that is why we disagree with the so-called ‘retroflex vowel [ɤ]’ frequently encountered in traditional descriptions of *érhuà*. A certain degree of postalveolarity in any [Y] is indisputable, but the essential feature to be underlined is the lateral contraction, which inevitably causes the apex to move back and raise.

Concisely, any [Y] is a coarticulation of *vocalic traits*, ie a specific quality or ‘timbre’ (determined by the mid-dorsum position and the degree of lip rounding), inherent voiceness and potential nasalization; and of *consonantal traits* fairly similar to those of a *slightly postalveolarized prevelar rounded semi-approximant with lateral contraction*, [ɹ]; the corresponding *approximant* is nothing but the famous ‘American /ɹ/’, [ɹ] (while Southern British /ɹ/ is, rather, a *postalveolar rounded approximant*, [ɹ̥], *pace* what most phoneticians still seem to think).³

The very symbol ‘[ɹ]’ suggested us the presser-foot-like diacritic in ‘[Y]’. Yet, there are some differences worth noticing: while *pūtōnghuà* [Y] are generally *pro-velar*, [ɹ, ɹ̥] are *pre-velar*; the latter two are also *intrinsically rounded*, instead any [Y] owns its potential roundedness to that of its basic [V]. In case a separate symbol for the ‘presser foot’ were needed, almost ‘distilling’ it from the actual coarticulation, that

would logically be a *provelar unrounded semi-approximant with lateral contraction*, [ɥ]. Therefore: [V] + [ɥ] → [Y].

fig 2.1. The orograms of the laterally contracted vocoids.

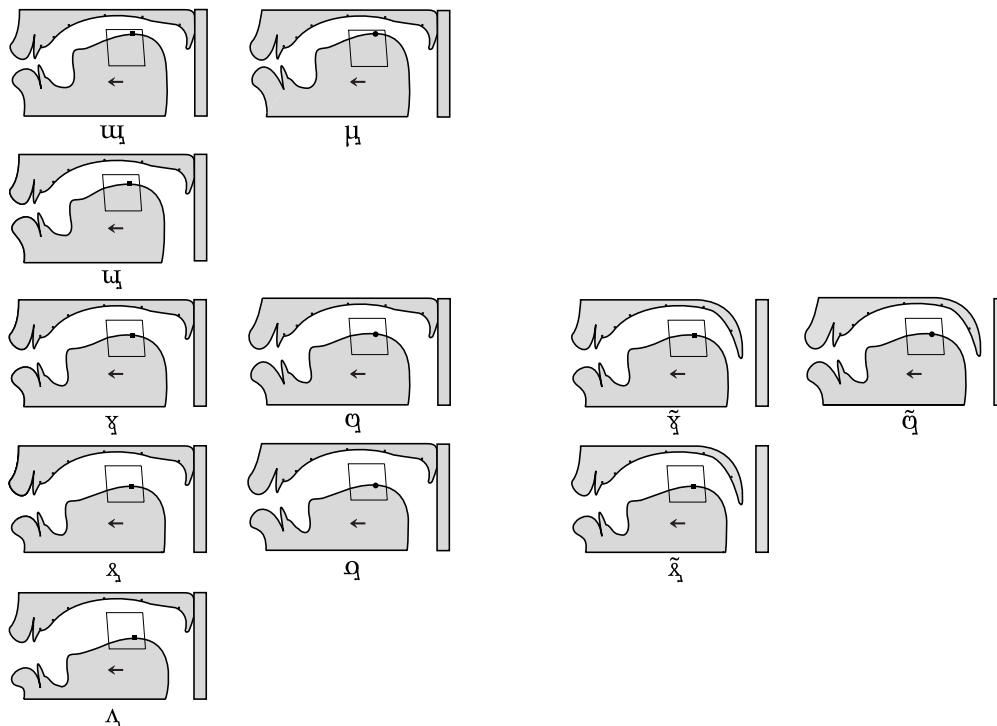
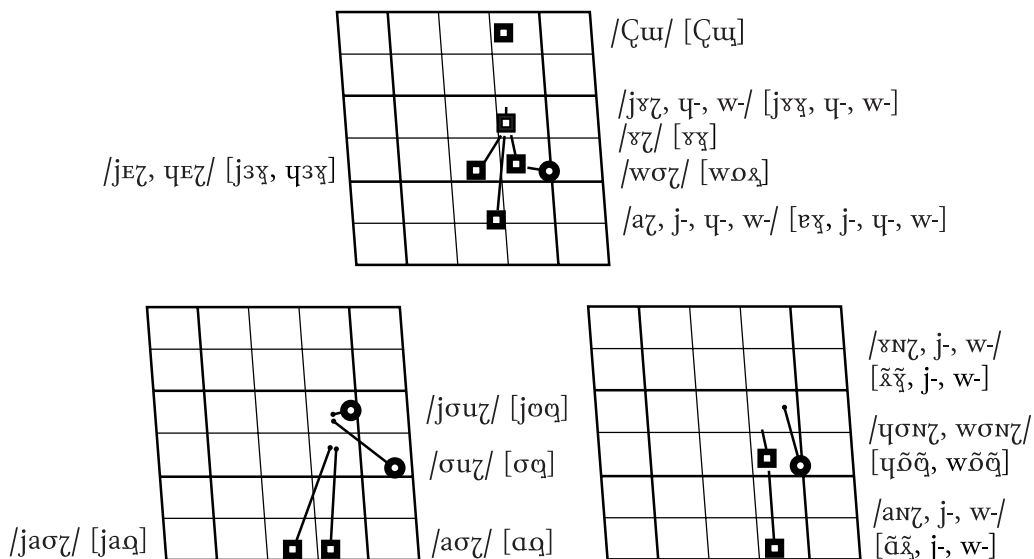


fig 2.2. The vocograms of the laterally contracted vocoids for the *suggested pronunciation*.

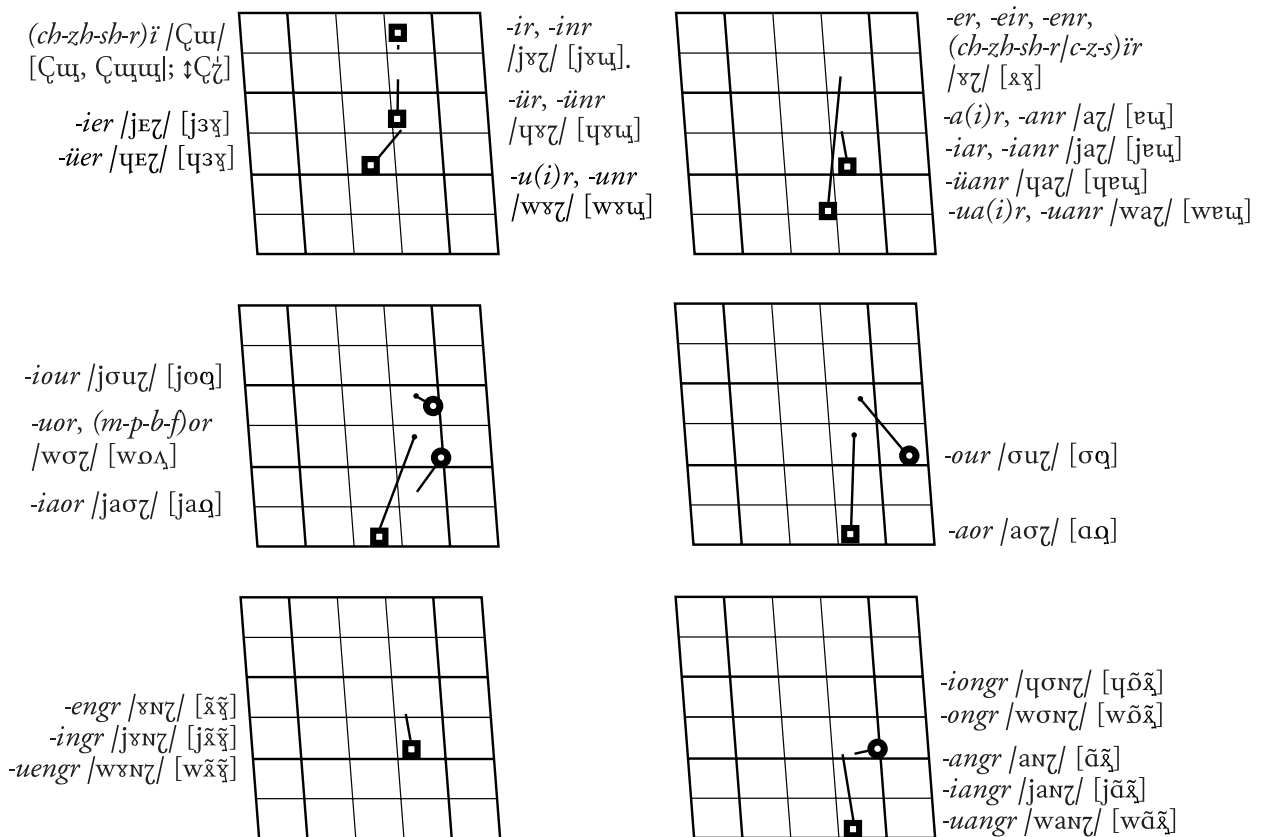


2.4. Before analyzing how these peculiar vocoids behave in *érhuà*, it is important to note that, even in /ʈɥw/, we normally find a lateral articulation, [ɥ], rather than an intense approximant like [ɥ̥], as various *offIPA* notations –eg ‘[ɥ̥, ɥ̥]’– would suggest. Many sinologists use ‘[ɥ]’, a nonstandard symbol whose *offIPA* equivalent

2. The *r* grammeme and the so-called *érhuà* phenomenon

should be ‘[ʅ]’, described as a ‘syllabic retroflex fricative’ or even ‘retroflex vowel’ (though written with a contoid glyph!). Undoubtedly, [ʅ] belongs to neutral pronunciation as [ʉ] does, but it is associated with a tenser articulation and a certain emphatic component (if not with a marked Peking accent).

fig 2.3. The vocograms of the laterally contracted vocoids with no simplification (but without secondary variants).



2.5. Starting with /()Vʅ/, /()a, ʃ, wɔ/ and /ʅ/ interact without phonemic adjustments: /aʅ, jaʅ, waʅ; ʃʅ; wɔʅ/. Chinese taxophonics does not allow //iʅ, yʅ// as such, and adjusts them to /jʃʅ; ɥʃʅ/. Exceptionally, //iʅ, yʅ// present oscillations that might suggest quite peculiar phonemicizations: [iʃ, yʃ] /iʃʅ; yʃʅ/; but all the sound files that we have examined clearly show that these oscillations are always related to intentionally slow *er* careful speech. In addition, //Cwʅ, Cʉʅ// change into /Cʃʅ, Cʉʅ/, in order to be better distinguished from [Cw, Cʉ] and especially from [Cʉʉ, Cʅ]. Herein we list the most frequent taxophones, including those with †[ʅ] (where ‘†’ suggests a less spontaneous pronunciation, often used by native speakers with a regional background and an intentional effort to approach a neutral pronunciation).

2.5.1. Let us notice that ‘^s[...]’ means *Suggested Pronunciation*, that is a sober and sufficient simplification meant to employ as few taxophones as possible: in our opinion, those who are willing to acquire a native-like –yet manageable– kind of pronunciation are recommended to choose it. Our readers will not fail to notice that the *suggested* variant does not always coincide with the first variant shown on the list, and it needs not to: for example, [ɐʷ] is the most typical for /aʷ/, yet we ‘suggest’ [ɐ̃], because it sounds as much as natural to natives, and it employs what we consider to be a very versatile taxophone, [ɿ].

ar; yar, iar; war, uar /aʷ, j-, w-/ [ɐʷ, j-, w-] & [-ʷ] & †[-ʷ] –^s[ɐ̃, j-, w-]
er & ir /ɛʷ/ (//ʷɛʷ//) [ɛ̃ʷ] & [ɛ̃ʷ] & †[ɛ̃ʷ] –^s[ɛ̃ʷ]
yir, ir; yür, ür /jɛʷ, ɸɛʷ/ [jɛ̃ʷ, ɸ-] & [jɛ̃ʷ, -] & †[jɛ̃ʷ, ɸ-] –^s[jɛ̃ʷ, ɸ-]
wor, uor; (m-b-p-f)or /wɔʷ/ [wɔ̃ʷ, wɔ̃ʷ; ɕɔ̃ʷ, ɕɔ̃ʷ] & †[wɛ̃ʷ; ɕɛ̃ʷ] –^s[wɔ̃ʷ]

2.5.2. Some sequences present such a wide range of taxophonic oscillations, that more than one phonemicization would be possible (and even advisable, for scientific coherence), though beginners need not use all the taxophones. For *yer, ier; yüer, üer* //jɛʷ, ɸɛʷ//, the sound files suggest three groups of realizations: less frequently, †[jɛ̃ʷ, ɸɛ̃ʷ], which plausibly lead to ‘/jɛʷ, ɸ-/’; more often, [jɛ̃ʷ; ɸɛ̃ʷ] /jɛʷ, ɸ-/ , totally assimilated to //iʷ, yʷ//; yet the most typical realizations are [jɛ̃ʷ; ɸɛ̃ʷ], which we still prefer to assign to /-ɛʷ/ (in a more useful interphonemic analysis, according to the principle of Natural Phonetics, rather than to /-ɛʷ/, in a supposed ‘more economical’ –but more abstract, and less useful– intraphonemic analysis). So, we recommend the last taxophonic sequences shown, [jɛ̃ʷ; ɸɛ̃ʷ] (or, at most, †[jɛ̃ʷ, ɸɛ̃ʷ], which risk to sound too precise, because of the use of the contoid [-ʷ], instead of the more typical laterally contracted vocoids, cfr fig 2.1): ⁴

yer, ier & yüer, üer /jɛʷ, ɸɛʷ/ [jɛ̃ʷ, ɸ] (& †[jɛ̃ʷ, ɸ]).

2.5.3. Similarly, //uʷ// provides two choices: /wɛʷ/, which we consider primary; and /uʷ/, certainly not indispensable, but useful to improve one’s understanding of spoken Chinese:

$(w)ur /wʁʒ/ [wʁʒ] \& \uparrow[wʁʒ] -^s[wʁʒ]$ (& $//uʒ// [uʒ]$).

2.5.4. The second elements of the diphthongs $/(w)ei$, $(w)ae/$ are incompatible with the place of articulation of $/ʒ/$; they are therefore reduced to a ‘zero’ element, leaving $/(w)ʁʒ$, $(w)aʒ/$; in theory, this causes homophony with *er* and *ar*, though the very few cases where this happens are substantially negligible. For $//(w)eiʒ//$, $[(w)ɛʒ]$ is not totally uncommon, but it is better to adopt $[(w)ʁʒ]$ as the ‘*s*’ norm, in order to keep an advantageous consistency with other $/ʁʒ/$ sequences.

2.5.5. Regarding $/aσʒ$, $jaσʒ/$, both vowels are preserved, though a slight difference in $/a/$ can be observed: with $/j/$, the $/a/$ phoneme is prevented from being realized as far back as $[ɑ]$, which otherwise would be very natural, considering the inherent midback quality of $[ɑ]$, $oʌ$. A similar behaviour distinguishes $/σuʒ/$ from $/jσuʒ/$, as clearly shown (cfr fig 2.3), in spite of the high number of variants currently occurring for this diphthong: we would drag the reader’s attention to the high number of taxophonic oscillations registered for this diphthong: some of them would not justify a fully $/(j)σuʒ/$ phonemization; $/(j)σʒ/$ surely suffices for an ‘International’ pronunciation, nonetheless it cannot properly represent secondary realizations like $[σuɑ]$, $ʒuɑ$ (if really needed, ‘ $/(j)σ(u)ʒ/$ ’ might work well as a practical shorthand for both $/(j)σuʒ$, $(j)σʒ/$).

aor $/aσʒ/ [ɑɑ] \& \uparrow[ɑσʒ] -^s[ɑɑ]$
yaor, iaor $/jaσʒ/ [jaɑ] \& \uparrow[jaσʒ] -^s[jaɑ]$
our $/σ(u)ʒ/ [σɑ] \& \uparrow[σuʒ]$
your, iur $/jσ(u)ʒ/ [jσɑ] \& \uparrow[jσuʒ]$

2.6. In *-nr*, $/n/$ disappears, too. Consequently, *gēr* ‘song’ and *gēnr* ‘root’ are pronounced alike, $/kʁʒ/ [kʁʒ]$; as well as *bàr* ‘handle’ and *bànr* ‘obstacle, petal’, $/pɑʒ/ [pɑʒ]$. It is important to note that *pīnyīn* is certainly not the actual phonic structure of Chinese (as many are lead to think), but just an official romanization, which does not exactly correspond to the phonemes of the language. As a matter of fact, *pīnyīn* might be more similar to a hypothetical, more abstract, phonological level, more like our possible $// //$ symbolization. For instance, we might have the sequences $//inʒ$, $ynʒ$, $wʁnʒ//$, for *inr*, *ünr*, *enr*; but $/jʁʒ$, $ʁʒ$, $wʁʒ/$ are certainly better.

As for //jɛnʒ, ɤɛnʒ//, it must be remembered that /jɛn, ɤɛn/ are simply more realistic phonemicizations for [jɛn, ɤan] than ‘/jan, ɤan/’; when /n/ disappears, the syllabic nucleus resumes its original theoretical vowel quality: /jaʒ, ɤaʒ/ (//janʒ, ɤanʒ//), with the same values as described above for /aʒ/, eg [ˈɤan̩˥˥tʂʰɤɛŋ̩] /‘ɤɛn˥˥tʂʰɤaʒ/ yuánqūānr.

2.6.1. Also -ng /ŋ/ is dropped before /ʒ/, but it does not vanish totally; instead, it nasalizes the preceding vowel and the following /ʒ/, with a certain number of taxophones. It must be considered that [ŋ̃] are quite complex and unstable co-articulations: beside a slight raising of the apex and a lateral contraction, they involve the lowering of the velum, too, while having to keep the timbre consistent with that of the original [V]. In short, the actual realizations may vary: we have chosen the most frequent ones, including those with [ŋ̃ʒ].

The neutralizations and readjustments explained for ordinary /V(n)ʒ/ largely apply to /Vŋʒ/, as well: *ingr* /jɤŋʒ/ (with a rare [ĩɤ̃] realization, too, //Cinʒ//); *ongr* can be explained either as a variant of *wengr*, thus /wɤŋʒ/ (with a fully audible [w] or, better, a milder [ɰ]), or as an autonomous ‘rhyme’, /ɔŋʒ/. This requires a concise explanation: by and large, realizations referable to /wɤŋʒ/ instead of /ɔŋʒ/ are fairly common and certainly a good choice for foreign speakers; but with the labial phonemes, /m, p, f, v/, the glides [w, ɰ] merge with the preceding /C/, producing a [C̥] articulated with rounded lips:

yingr, ingr /jɤŋʒ/ [jĩɤ̃, jɤ̃ɤ̃] (& †[ĩɤ̃, jĩɤ̃ʒ], //inʒ//) – ^s[jĩɤ̃ɤ̃];
angr; yang, iangr; wangr, uangr /aŋʒ, j-, w-/ [ãɤ̃, j-, w-] & †[ãʒ, j-, w-]
engr /ɤŋʒ/ [ɤ̃ɤ̃] (& †[ɤ̃ʒ])
wengr, ongr /wɤŋʒ/ [wɤ̃ɤ̃, wĩɤ̃, wɤ̃ɤ̃] (& †[wĩɤ̃ʒ, wɤ̃ɤ̃ʒ]) – ^s[wĩɤ̃ɤ̃]
ongr /ɔŋʒ/ [õɤ̃, õĩɤ̃] (& †[õʒ]) – ^s[õĩɤ̃]
yongr, iongr /ɤŋʒ/ [ɤ̃ĩɤ̃, ɤ̃õĩɤ̃] (& †[ɤ̃ĩɤ̃ʒ]) – ^s[ɤ̃ĩɤ̃].

2.7. We now present a few examples, transcribed according to our *Suggested Pronunciation*, though our readers will not find it difficult to prepare more transcriptions on their own, with all the –often capricious– taxophones of a purely ‘native’ pronunciation. Besides, once the few ‘suggested’ taxophones are automatized and the speaker’s fluency improves, thanks to intensive listening and speaking exercises (and to the influence of co-articulation, as well), the other taxophones very often come by themselves.

2. The *r* grammeme and the so-called *érhuà* phenomenon

- A. /aʒ/: [ɐʒ] *ar, air, anr*; [jɐʒ] *iar, ianr*; [ɥɐʒ] *üanr*; [wɐʒ] *uar, uair, uanr*:
 /aʒ/ (//a(i)ʒ, anʒ//): *huà-huà* [ˈxwa|kʰwɐʒ];
 /jaʒ/ (//jaʒ, janʒ//): *chuāngliánr* [ˈtʃhwɑŋˈliɐʒ];
 /ɥaʒ/ (//ɥanʒ//): *yüánqüānr* [ˈɥɑŋˈtɕhɥɐʒ];
 /waʒ/ (//wa(i)ʒ, wanʒ//): *hǎowánr* [ˈxɑŋˈwɐʒ].

- B. /ɛʒ/: [ɛʒ] *er, eir, enr, ír*; [jɛʒ] *ir, inr*; [ɥɛʒ] *ür, ünrr*; [wɛʒ] *ur, uir, unrr*:
 /ɛʒ/ (//ɛiʒ, ɛ(n)ʒ, ɯʒ//): *bèikér* [ˈb̥ɛiˈkʰɛʒ];
 /jɛʒ/ (//i(n)ʒ//): *shǐjìnr* [ˈʃɿˈj̥iːnʒ];
 /ɥɛʒ/ (//y(n)ʒ//): *qūqūr* [ˈtɕhɥˈtɕɥɛʒ];
 /wɛʒ/ (//uʒ, wɛiʒ, wɛnʒ//): *kǒuwèir* [ˈkʰɔˈw̥ɛiˈwɛʒ].

Again, *yer, ier & yüer, üer* might be given as /jɛʒ, ɥɛʒ/ too, yet we prefer to interpret them precisely as /jɛʒ, ɥɛʒ/, thus:

- /jɛʒ/: *shǐjiēr* [ˈʃɿˈj̥iːɛʒ];
 /ɥɛʒ/: *mùjüér* [ˈm̥ɥˈj̥üːɛʒ].

- C. /σuʒ, jσuʒ/: [σuʒ] *our*, [jσuʒ] *iur*; /wσʒ/: [wσʒ, ʒσʒ] (*u*)*or*; /(j)ασʒ/: [αʒ] *aor*, [jaʒ] *iaor*:
 /σuʒ/: *tǔdòur* [ˈtʰuˈtɕuʒ];
 /jσuʒ/: *zhǐliúr* [ˈʒ̥liˈuʒ];
 /wσʒ/: *shānpōr* [ˈʃɑŋˈp̥oʒ];
 /ασʒ/: *nǎimáor* [ˈn̥ɑˈm̥ɑʒ];
 /jaσʒ/: *màimiáor* [ˈm̥ɑˈmiˈɑʒ].

- D. /aŋʒ/: [ɑ̃ʒ] *angr*; [jɑ̃ʒ] *iangr*; [wɑ̃ʒ] *uangr*:
 /aŋʒ/: *piānchǎngr* [ˈpʰjɛŋˈtɕhɑ̃ʒ];
 /jaŋʒ/: *dòujiāngr* [ˈd̥oˈj̥iːɑ̃ʒ];
 /waŋʒ/: *tiānchuāngr* [ˈtʰjɛŋˈtɕhwɑ̃ʒ].

- E. /ɛŋʒ/: [ɛ̃ʒ] *engr*; [jɛ̃ʒ] *ingr*; [wɛ̃ʒ] *uengr*:
 /ɛŋʒ/: *bǎndèngr* [ˈb̥ɑ̃ˈn̥d̥ɛŋʒ];
 /jɛŋʒ/: *qǐ míngr* [ˈtɕh̥iˈm̥j̥ɛ̃ʒ];
 /wɛŋʒ/: *xiǎo wèngr* [ˈɕj̥ɑ̃ˈw̥ɛ̃ʒ].

- F. /ɔŋʅ/: [õǔ] *ongr*; [ʏõǔ] *iongr*:
 /ɔŋʅ/: *tōukòngr* [ˈthɔʊ\kɥõǔ];
 /jɔŋʅ, ʏ/: *kūqiónggr* [ˈkhuˈtɥʏõǔ].

2.8. Compared to the huge number of the Chinese-speaking population worldwide, it is just a tiny minority that can use *érhuà* natively and properly, being it a peculiarity of Peking and of the Peking area: many geolects, including some of the ‘north-eastern’ group, do not contemplate the *érhuà* phenomenon at all, or lack those phones necessary to realize it correctly. In other dialects, something comparable to *érhuà* has crystalized in a more conservative phase than that of Peking *érhuà*, or it has taken a different course.

Beside dialectal interference, the Chinese script, too, influences speech habits of those for whom *érhuà* is an (often erratically) acquired feature: the *r* grammeme, when explicitly registered, is represented by the character *ér* ‘son’ *added to* the primary character. Because of that, many non-Pekingese speakers, especially when poorly educated, may indeed pronounce two distinct syllables, each one with its own segments; other speakers would instead articulate only one, but possibly with minor ^{or} non-standard neutralizations... not to be imitated, such as:

- *[!ʂʏ.ǔ] /ʂwɣʅ/ *shùr* ‘number; vertical stroke in Chinese calligraphy’;
- *[!ɕī fu.ǔ] /ˈsi_ofwɣʅ/ *xífur* ‘wife’;
- *[.ɕjaɣˈtɥjɔʊ.ǔ] /,sjaɔˈtshjɔʊʅ/ *xiǎo qiúr* ‘small ball’;
- *[!tshwɔ.ǔ] /tshwɔʅ/ *cuòr* ‘mistake’.

2.8.1. Proper *érhuà* enjoys prestige in speaking *pǔtōnghuà*, but its use is hardly coherent outside Peking; even there, ‘Pekingese’ speakers (treated as such for being born in the city, or for having acquired the local accent after a long stay there) are uncertain whether to use *érhuà* in formal contexts, considering it too colloquial. Not surprisingly, when bound to a *formal* register, most broadcasters and reporters, as well as government spokespersons, read aloud official texts as they are written: without the character *ér* and –consequently– with no *érhuà*. Instead, a markedly *colloquial* register shows the opposite trend: *érhuà* abounds, sometimes overwhelmingly –if not anarchically– in vernacular Pekingese.

2.8.2. From a general perspective, the *formal vs colloquial* opposition is not the best criterion to explain *érhuà*, as many scholars still do. In fact, we are dealing with the result of a historical process *inherent to the language*, from which the new *r* derives: a true grammeme, not a mere embellishment. Regardless of any artificiality, *érhuà* is as much a part of the spoken language as the aforementioned suffixes *-zi* and *-li* are. To put it bluntly, certain words *must* have *-r*, others *may* have it, while others are not entitled to. The most authoritative reference work in the field is, by and large, the *Contemporary Chinese Dictionary*, where any lexeme requiring or allowing *érhuà* is clearly marked. Both native speakers and foreign learners could simply resort to it and look up the relevant word(s) they doubt about, but unfortunately few people actually do so. Therefore, it is not easy to determine a universally shared rule in cultivated speech, not even among ‘Pekingese’ speakers.

2.8.3. In an ‘International’ kind of pronunciation, slightly –but reasonably–simplified, when compared to the ‘normalized’ neutral one, just three sequences might suffice, namely [aʒ, ɜʒ, oʒ], to which any other /V(V)ʒ, Vnʒ/ shall adhere and conform. For /Vnʒ/, [ãʒ, ẽʒ, õʒ] are fair enough to make oneself understood; but attention should be paid not to omit this nasalization, given that [ãʒ, ẽʒ, õʒ] are *not* to be merged with [aʒ, ɜʒ, oʒ]. Let us examine a few examples: *bānr* [ˈpaʒ] ‘job; scar’ vs *bāngr* [ˈpãʒ] ‘band, gang’; *bèir* [ˈpɜʒ] ‘totally; generation’ vs *bèngr* [ˈpɛʒ] ‘separation; little coin’. After all, such nasalization is not difficult at all, provided the speakers are able to lower their velum, as illustrated in fig 2.1. For *yīngr*, we recommend [jɜʒ]; while, for *yōngr* and *iongr*, the learner could usefully evaluate the alternatives, [jõʒ, ɥẽʒ, ɥõʒ], though we recommend *iongr* [jõʒ] and *ongr* [õʒ], for being the most practical ones: the important thing is to be coherent in the course of time. It goes without saying that an approximation to native vocoids produces a more authentic result and is, of course, highly recommendable (and appreciated).

A. Related to ⁱ[aʒ], we have: ⁱ[aʒ] *ar, air, anr*; ⁱ[jaʒ] *iar, ianr*; ⁱ[ɥaʒ] *üanr*; ⁱ[waʒ] *uar, uair, uanr*.

ⁱ[jaʒ]: *chuāngliánr* ⁱ[ˈtʂhwan˥˥jaʒ], ^s[ˈtʂhwan˥˥jɛʒ];

ⁱ[ɥaʒ]: *yuánquānr* ⁱ[ˈɥɛŋ˥˥tɕhɥaʒ], ^s[ˈɥaŋ˥˥tɕhɥɛʒ];

ⁱ[waʒ]: *hǎowánr* ⁱ[.haɔ˥˥waʒ], ^s[.ɤɔ˥˥wɛʒ].

B. Related to $i[ɜ̄]$, we have: $i[ɜ̄]$ *er, eir, enr, ǐr*; $i[jɜ̄]$ *ir, ier, inr*; $i[ʏɜ̄]$ *ür, üer, ün̄r*; $i[wɜ̄]$ *ur, uir, unr*.

$i[jɜ̄]$: *shǐjǐnr* $i[.ʂʅ'|\tɕjɜ̄]$, $s[.ʂɥ|\tɕjɜ̄]$;
 $i[ʏɜ̄]$: *qūqūr* $i[.tɕhy·tɕhyɜ̄]$, $s[.tɕhy·tɕhy]$;
 $i[wɜ̄]$: *kōuwèir* $i[.khɔu|wɜ̄]$, $s[.khɔy|wɜ̄]$.

C. Related to $i[ɔ̄]$, we have: $i[ɔ̄]$ *our*; $i[jɔ̄]$ *iur*; $i[wɔ̄]$ *uor*; $i[aɔ̄]$ *aor*; $i[jaɔ̄]$ *iaor*.

$i[ɔ̄]$: *tǔdòur* $i[.thu|tɔ̄]$, $s[.thɥ|tɔ̄]$;
 $i[jɔ̄]$: *zhīliúr* $i[.tʂ'|\jɔ̄]$, $s[.dʒɥ'|\jɔ̄]$;
 $i[aɔ̄]$: *nǎimáor* $i[.naɛ'maɔ̄]$, $s[.naʂ'maɔ̄]$;
 $i[jaɔ̄]$: *màimiáor* $i[.maɛ'mjaɔ̄]$, $s[.maʂ'mjaɔ̄]$;
 $i[wɔ̄]$: *shānpōr* $i[.ʂam̄|phwɔ̄]$, $s[.ʂam̄|p̄hɔ̄]$.

D. Related to $i[ā̄]$, we have: $i[ā̄]$ *angr*; $i[jā̄]$ *iangr*; $i[wā̄]$ *uangr*.

$i[ā̄]$: *piānchǎngr* $i[.phjɛŋ|tʂhāā̄]$, $s[.phjɛŋ|tʂhāā̄]$;
 $i[jā̄]$: *dòujiāngr* $i[.tɔu·tɕjā̄]$, $s[.dɔy·tɕjā̄]$;
 $i[wā̄]$: *tiānchuāngr* $i[.thjɛŋ|tʂhwā̄]$, $s[.thjɛŋ|tʂhwā̄]$.

E. Related to $i[ǎ̄]$, we have: $i[ǎ̄]$ *engr*; $i[jǎ̄]$ *ingr*; *uengr*.

$i[ǎ̄]$: *bǎndèngr* $i[.pan|tǎ̄]$, $s[.b̄aŋ|tǎ̄]$;
 $i[jǎ̄]$: *qǐmíng* $i[.tɕhi'mjǎ̄]$, $s[.tɕh̄|mjǎ̄]$;
 $i[wǎ̄]$: *xiǎowèngr* $i[.ɕjaɔ|wǎ̄]$, $s[.ɕjaɔ|wǎ̄]$.

F. Related to $i[ō̄]$, we have: $i[ō̄]$ *ongr*; $i[jō̄]$ & $[yō̄]$ *iongr*.

$i[ō̄]$: *tōukòngr* $i[.thɔu|khō̄]$, $s[.thɔu|khō̄]$;
 $i[jō̄]$: *kūqióng* $i[.khu'tɕhɥō̄]$, $s[.khu'tɕhɥō̄]$.

NOTES:

¹ A detailed analysis of the historical origin and development of the *érhuà* phenomenon, and of its dialectal variation, is provided by LI, S-J. (1984) *Hanyu <ér> [ɛ] yin shi yanjiu*.

² cf also LU, Y-ZH. (2001) *Qingsheng he erhua*, which, besides *érhuà*, provides abundant examples of words with a 'zero' toneme, too.

2. The *r* grammeme and the so-called *érhuà* phenomenon

³ Again, see CANEPARI, L. (2009) *English Pronunciation*, Ch 26 ‘The ‹whole truth› on English *r*’, as it helps avoid any further misunderstanding (and misuse) of the relevant phonetic symbols, especially in consideration of the noticeable difference –in this regard– between ^{off}IPA and ^{can}IPA.

Besides, cf LIN, H-Y. (2007) *The Sounds of Chinese*, § 2.1.3.2, on /ʅV/; and § 3.4.5, where ‘rhotacized/retroflexed vowels’ are concisely described. Though the symbols adopted by the author are inevitably quite generic, yet it is quite clear that a certain degree of taxophonic variation has been duly noticed (which is remarkable, when compared with the bewildering superficiality found elsewhere).

⁴ For a typically traditional treatment of *érhuà*, see LI, M. & SHI, P-W. (1986, rist. 2004) *Hanyu putonghua yuyin bianzheng*, Ch 12.

3. Consonants and contoids

3.1. The consonantal system of Mandarin Chinese shows some interesting peculiarities, including some alternative possibilities of phonemic analysis. For example, the phonemic status of [tʂʰ, tʂ, ʂ] is debatable; in fact, *pīnyīn* has chosen their ‘official’ acknowledgement –as *j, q, x*– to the detriment of vowels, especially /y/: this is rendered as *u*, even if its phonemic values are indisputably /y, ʏ/. On the contrary, it would have been sufficient to systematically use *ü*, instead of fluctuating between *yu, ju, qu, xu* and *nü, lü* – all the more so because, with *nü, lü*, both an Umlaut and a tonal mark are to be used anyway. If no *ü* had to be used at all, the most reasonable –and typographically coherent– solution would have been keeping *iu* and *iou* as the post-consonantal counterparts of *yu* and *you*, without switching schizophrenically from one to another. Then, simply *yu, jiu, qiu, xiu* for [ʏy, tʂy, tʂʰy, ʂy], and *you, jiou, qiou, xiou* for [jʊʊ, tʂjʊʊ, tʂʰjʊʊ, ʂjʊʊ].

fig 3.1. Table of Mandarin contoids.

| | bilabial | bilabial rounded | palatalized bilabial | labiodental | labiodental round. | dental | alveolar | postalveolar | bilabialized pre-palatal (* prepalat.) | palatal | postpalatal rounded | velar | velar rounded | uvular | laryngeal | laryngeal rounded | palatalized laryngeal |
|----|--------------------|---------------------|----------------------|-------------|--------------------|----------------------|----------|----------------------|--|---------|---------------------|--------------------|---------------|--------|-----------|-------------------|-----------------------|
| N | m | [m̠] | [m̠] | [m̠] | [m̠] | [n] | n | [ŋ] | [ɲ]* | | | [ŋ] | | ɴ | | | |
| K | p ^h [b] | [p ^h b̠] | [p ^h b̠] | | | t ^h [d] | | | | | | k ^h [g] | | | [ʀ] | | |
| KS | | | | | | tʂ ^h [dʒ] | | tʂ ^h [dʒ] | [tʂ ^h dʒ] | | | | | | | | |
| X | | | | f[v] | [f̠ v̠] | | | | | | | | | x | [χ] | [χ̠] | [χ̠] |
| S | | | | | | s[z] | | ʃ[z] | [ʃ z] | | | | | | | | |
| J | | | | | | [ʒ] | | ʒ | | j[j] | ɥ[ɥ] | [h] | w[ʋ] | [x] | [h] | [h̠] | [h̠] |
| L | | | | | | | l | | | | | | | | | | |

Considering the vowel phonemic inventory that we have adopted, it appears more reasonable not to consider [tʂʰ, tʂ, ʂ] as phonemes, but simply as taxophones of /ts, tsh, s/ when followed by /i, y; j, ʏ/.¹ All this, of course, is irrespective of both morphological and lexically-derived considerations, since they are not at all absolute. In fact, [tʂʰ, tʂ, ʂ] historically derive from two distinct series: /ts(h), s/ + /i, y;

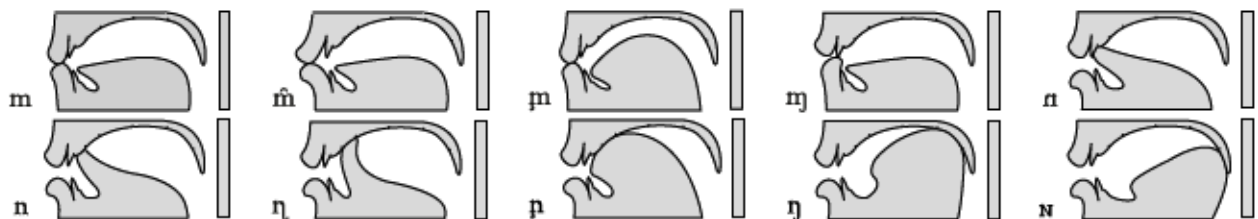
j, ɥ/ and /k(h), ʃ/ + /i, y; j, ɥ/. The ancient *Phags-pa* inscriptions, as well as comparisons with modern dialects, testify that both series still existed in Early Mandarin (around the XIII century AD), but then gradually merged.² Nevertheless, we insist that phonology must be explicit, in order to be useful. The simple reason why we have chosen to assign [tʃ̥, tʃ̥, ʃ̥] to /tʃh, ts, s/ only, then, is that this series is nearer to the phonetic reality than /kh, k, ʃ/ is, both at the auditory and articulatory level.

Nasals

3.2. There are three nasal phonemes (fig 3.2): /m, n, ŋ/ [m, m̂; ŋ, n, ŋ, ɲ, ɳ; ɴ], which *pīnyīn* correctly renders as *m, n, ng*. In syllable-final position, only /n, ɴ/ occur, and the latter always maintains its uvular articulation, even before any consonants. Instead, /n/ regularly assimilates, except in very slow and accurate pronunciation; assimilation is only partial for /ɲj, ɲɥ/, though: [ɲj, ɲɥ]; for /n[#]w/, [ŋw] is a possibility. Besides, let us notice the reduced forms of *wǒmen* and *tāmen*, which become monosyllabic in very fast speech, with /m[#]/: [w̥m, ˉtam]. The tendency of /n[#]/ to assimilation is so strong that we can find cases such as: [ʒ̥ːd̥jɛˉˌw̥zɛnˌl̥xˌmaˌ] (& [ˉˌw̥zɛlˌl̥x]) /ʒ̥ːˈtsjɛˉˌw̥zɛnˌl̥xˌmaˌ?/ *Jiéhūn le ma?* Mandarin has /n/ [ɲ] + /i, y; j, ɥ/, as said, although [n] is possible as well, and can be safely used by foreign learners.

On the other hand, before the dentals, /t, th, ts, tʃh, s/, the more specific symbol would be [ɲ], but this is not at all necessary. On the contrary, it is useful to clearly mark [m̂] in /mwσ/ [m̂σ(ʃ)]. Syllable-final and word-final nasals always have an inaudible release, even before a pause: [ɲ̥, ɴ̥]. (This is quite normal for other oriental languages such as Cantonese, Thai, Cambodian, Vietnamese, Tagalog, Malay, and Indonesian – even for final voiceless stops, now lost in Mandarin.)

fig 3.2. Nasal contoids.



3.2.1. Here are some examples: [ˉˌmao] *māo*, [ˉˌm̂σ] *mō*, [ɲɑɑɑ] *nǎo*, [ˈp̂hɛɲ] *pén*, [ˈp̂hɛɲ] *péng*; [ˌx̂ɛɲ|ʔ̂x̂(ʃ)] *hén è*, [ˈtʃ̥ɥɑɲ.th̥] *quántǐ*, [ˈd̥ĵɳ.tʃ̥ɑɑɑ] *jǐnzǎo*, [ˈd̥wɑɲ|l̂ĵɳ, -Al] *duànliàn*, [ˈʒ̂ɑmˈmɪn] *rénmín*, [ˈl̂ŵzɛmˈp̂ɑɲ] *lúnbān*, [l̂k̂ĥɑɲ.f̂ɑɑ]

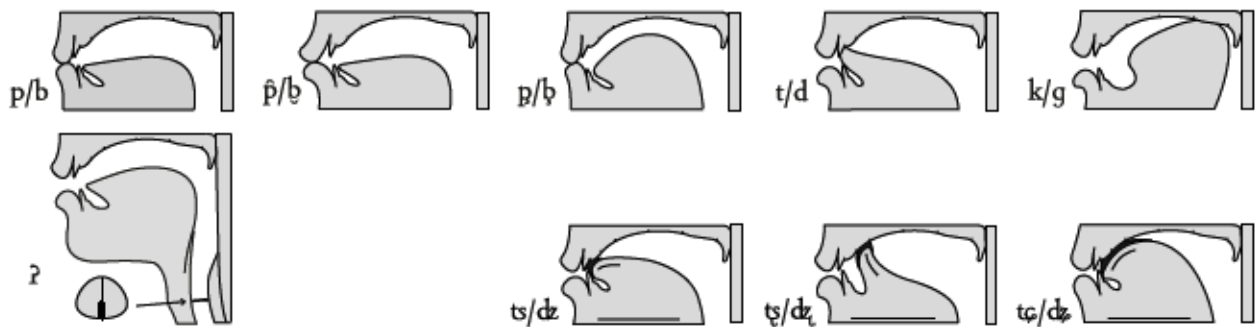
kànfǎ, [ˈb̥jɛŋ˥˥t͡ɕi] *biānji*, [ˈt͡ɕjɛŋ˥˥t͡ɕjɛŋ˥˥] *jìnzhǐ*, [ˈpʰɑŋ˥˥t͡ɕjɛŋ˥˥] *pànjüé*, [ˈb̥jɛŋ˥˥t͡ɕjɛŋ˥˥] *biǎnyú*, [ˈg̊ɑŋ˥˥t͡ɕjɛŋ˥˥] *gǎnrǎn*, [ˈb̥iŋ˥˥k̚wɑŋ˥˥] *bīnguǎn*, [ˈt͡ɕjɛŋ˥˥wɑŋ˥˥, -ŋ|w-] *jiànwàng*, [ˈt͡ɕjɛŋ˥˥wɑŋ˥˥] *liánhuān*.

With /N/: [ˈji|lɥ ·pʰiŋ˥˥ʔɑŋ] *yìlù píng'ān*, [ˈf̥ɛŋ˥˥m̥jɛŋ˥˥] *fēngmiàn*, [ˈkʰoŋ˥˥p̥jɛŋ˥˥] *kǒngbù*, [ˈlɑŋ˥˥f̥ɛi] *làngfèi*, [ˈf̥ɛŋ˥˥t̥ɕi] *fēngdǐ*, [ˈkʰoŋ˥˥t̥ɕi] *kǒngxì*, [ˈpʰiŋ˥˥t͡ɕjɛŋ˥˥] *píngyú*, [ˈf̥ɛŋ˥˥t͡ɕjɛŋ˥˥] *shēngrì*, [ˈf̥ɛŋ˥˥t͡ɕjɛŋ˥˥] *fēngshàn*, [ˈpʰiŋ˥˥k̚wɑŋ˥˥, k̚wɑŋ˥˥] *píngguǒ*.

Stops and stop-strictives

3.3. Our transcriptions aim at being as clear and realistic as possible; at the same time, they are to be meant as a convenient *normalization* of numerous taxophones, some of which tend to be more typical and recurrent than others. It is precisely such a statistical recurrence criterion that our readers should consider, when comparing normalized transcriptions with those of some specific sentences uttered by given speakers. In this last case, all the peculiarities of a single speaker must be taken into account, and transcribed accordingly (especially at the *hyper-phonetic* level, ‘[]’, where necessary).

fig 3.3. Stop and stop-strictive contoids.



Traditionally, Chinese stops and stop-strictives (fig 3.3) are said to distinguish minimal pairs through ‘aspiration’: *off*/*C^h*, *C/*. In fact, ‘aspirated’ consonants are, rather, */Ch/* sequences, where */h/* is therefore a true phone and phoneme, indeed. Since voicing is not phonemic, many Chinese phoneticians –simplistically– explain ‘*/C^h*, *C/*’ exactly as ‘*[C^h*, *C]*’ anywhere, as if there were no variation. Others, at least, introduce ‘*[C̥]*’ for */C/*; again, without further distinctions, regardless of their phonetic reality, which –as usual– is a little more complex than that. ³

3.3.1. In genuine pronunciation, we have observed –at least– three degrees of variation, depending on whether a syllable has a *primary* stress, [ʰ], a *secondary* one, [ᵀ], or if it is *unstressed*, [₀]. For stops and stop-strictives, we have /Kh, KSh/ [ʰ, ᵀ], and /K, KS/ [₀, ᵀ, ₀]. This is absolutely automatic for native speakers, who instinctively give priority to fully stressed syllables, while producing much laxer articulations in unstressed ones.

Yet, for /Kh, KSh/, native speakers might also employ [ʰ] instead, with a *semi-constrictive* (obviously weaker than its constrictive counterpart [h]), or even [ʰ], with the *lenis laryngeal approximant* more commonly found in /ᵀKh, ᵀKSh/. That depends on many factors, including enunciation speed and paraphonic superstructures, like *emphatic* or (on the contrary) *muddy speech*, not to mention those vocal traits that good actors often adopt for theatrical purposes. For example, it is clear that an ancient military hero giving orders to his soldiers will prefer a distinct, cut-glass, somewhat stentorian delivery, while a Táng lady-in-waiting is supposed to talk in a most refined, mellifluous voice, coherent with her social status.

Regarding [₀], we shall employ [₀] in general transcriptions, but it is important to underline that a residual ‘aspiration’ may be present, here and there, especially in relatively slow speech: if not [h], possibly something even lighter and shorter, eg a *semi-approximant*, [₀h].

3.3.2. Similar flexibility must be applied to [ᵀ, ᵀ, ₀], respectively *voiceless lenis*, *partially voiced* (actually, a mixed-phonation contoid) and *voiced* contoids. First of all, it is hard to present Mandarin /K, KS/ as simply voiceless [ᵀ]: in that case, native speakers of languages with a true /ᵀ, ᵀ/ opposition would notice no difference between –say– their /t/ and what *d* is supposed to represent in *pinyin*. Instead, to Italian speakers, for instance, a sequence like *dàodá mùdìdì* [ʰtao˥˥ ta ˩mu˥˩ti˥ti˩] would almost certainly sound like a hypothetical Italian sequence /ˈtaodamu di˥ti˩/, or even /ˈdaodamu di˥di˩/, depending –among other things– on the hearers’ own speech habits, which always influence their perception of unfamiliar phones.

Since /t/ is a truly voiceless [t] in Italian, such a misinterpretation suggests that the Chinese /t/ produces an auditory impression slightly ‘milder’ (or ‘laxer’) than its Italian counterpart, even when fully stressed.

In our normalized scheme, in fact, we would have [ʰt̚˥˥˩˥˩˩˩ -di˩]: nevertheless, as good recordings easily demonstrate, [ᵀ] occurs too, as well as [ᵀ, ₀]. Os-

cillation, again, depends on many factors and on the characteristics of single speakers. Anyway, very few Chinese hearers perceive any difference, especially among *stop phones* (while voicing is somewhat easier to detect in *stop-strictive phones*, thanks to the /S/ component that makes voicing much more evident).

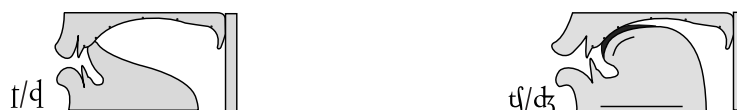
Even if /'K, 'KS/ were occasionally articulated as purely voiceless [C̥], and not as voiceless lenis [C̣], a degree of distinctiveness between ‘aspiration’ and ‘non-aspiration’ would be guaranteed anyway. In fact, only /_oKh, _oKSh; 'K, 'KS/ would actually overlap, coinciding phonetically (but not functionally, from a phonemic point of view).

3.3.3. Native speakers of languages like English and German will find that acquiring Chinese taxophones is a relatively easy task for them: their mother tongues primarily oppose /C̥, C̣/, but ‘aspiration’ and *mixed phonation* are very important components, at the phonetic level, and can be employed intentionally, once perceived adequately; for those who prefer an easier approach, we regard /Kh, KSh/ [C̥h] and /K, KS/ [C̣] as a versatile compromise. This is certainly ‘safer’ for foreign learners: first, it ensures a larger auditory distance between /Ch/ and /C/ than simply [C̥h; C̣] (which is not wrong anyway), and it avoids unintentional ‘aspirations’ by those speakers whose native language –or accent– features this phenomenon: besides English and German, many Southern Italian regional accents, for instance (while neutral Italian has no ‘aspiration’, neither phonemic nor phonetic).

To acquire [C̣], one may learn to isolate and consciously articulate the latter contoid in *Bob* [ḅ], *dad* [-ḍ] (alveolar, though, not dental as in Chinese, [ḍ]), *gag* [-g̣], before a pause or an unvoiced contoid. In fact, if we listen carefully, we shall notice a subtle difference between *Bob comes* [-ḅ 'kh-; -ẓ] and *Bob goes on* [-b 'g-; -z].

English has no unitary /ts/, yet the sequence *ds* as in *dads* [-ḍẓ] offers a rough approximation to [ḍẓ]; for [ḍẓ, ḍẓ], one should start from English *judge*, first comparing *j-* [-ḍẓ-] and *-dge* [-ḍẓ], then gradually approximating to the correct points of articulation (fig 3.4 shows the orograms of [ṭ, ḍ; ṭʃ, ḍʒ], for a useful comparison between English [ḍ, ḍẓ] and Mandarin [ḍ, ḍẓ, ḍẓ]).

fig 3.4. The orograms of [ṭ, ḍ; ṭʃ, ḍʒ], for a useful comparison between English and Chinese.



3.3.4. Let us now see the following examples: [ˈpɪn] *bīn*, [ˈp̥ɔ] *bó*, [l̥p̥h̥ɔ] *pò*, [ˈtʃjɛ] *dié*, [ˈkhwəɪ] *kui*, [ˈp̥hɪn̥p̥hən] *pīngpāng*, [l̥t̥h̥j̥ḁɔ̥w̥ɥ̥ɥ̥] *tiào-wú*, [ˈk̥h̥ɛn] *kēng*; [l̥p̥ḁb̥ḁ], [-b̥ɤ] *bàba*, [l̥t̥d̥], [-d̥ɪ] *didi*, [ˈk̥ɣ̥·g̥ɣ̥], [-g̥ə] *gēge*, [l̥ɣ̥ɥ̥.p̥h̥ɛ̃ɣ̥], [-p̥ɣ̥ɣ̥] *shìpai*, [ˈp̥hɪu̥.t̥h̥ḁɔ̥], [-t̥ɔ̥] *pútao*, [ˈp̥hɪu̥.t̥ɔ̥.t̥ɔ̥j̥ɔ̥ɥ̥], [-t̥ɔ̥] *pútaojiǔ*, [l̥t̥sh̥ɥ̥.k̥h̥ɣ̥], [-k̥ɔ̥] *cìke*.

3.3.5. As far as stop-strictives are concerned, therefore, we have /ts, tsh/ [t̥s, t̥ɕ, ɔ̥ɕ; t̥sh, t̥sh, ɔ̥ts] *z, c*; before /i, y; j, ɥ/, they are realized as [t̥ɕ, t̥ɕ, ɔ̥ɕ; t̥ɕh, t̥ɕh, ɔ̥t̥ɕ] *j, q*. Finally, /tʃ, tʃh/ [t̥ʃ, t̥ʃ, ɔ̥t̥ʃ; t̥ʃh, t̥ʃh, ɔ̥t̥ʃ] *zh, ch*. As previously said, [ɔ̥ts, ɔ̥t̥ɕ, ɔ̥t̥ʃ] must be intended as [ɔ̥ts(h), ɔ̥t̥ɕ(h), ɔ̥t̥ʃ(h)], ie with a possible residual ‘aspiration’.

Here are some examples (expressly chosen with final syllables with a zero-toneme): [ˈt̥sɛ] *zāi*, [l̥t̥shw̥ɛ̃n] *cuàn*, [ˈt̥ɕy] *jū*, [ˈt̥ɕh̥j̥ɔ̃n] *qióng*, [l̥t̥ʃɛɪ] *zhèi*, [ˈt̥ʃhɥ] *chí*; [ˈp̥i-d̥zɥ], [-d̥zɥ] *bízǐ*, [ˈɥ̥ɪn̥-t̥shɛ̃ɣ̥], [-t̥shɛ̃ɣ̥; ˈj̥ɪn̥-] *yúncai*, [l̥t̥j̥ɛ̃n̥.d̥z̥], [-d̥z̥] *diànji*, [ˈt̥h̥j̥ɛ̃n̥.t̥ɕh̥ɪ], [-t̥ɕh̥ɪ] *tiānqi*, [ˈt̥shɛ̃ɣ̥.d̥z̥ɥ], [-d̥z̥ɥ] *cáizhǐ*, [ˈp̥h̥ɛɪ.t̥ʃh̥ɛ̃n], [-t̥ʃh̥ɛ̃n] *péichen*.

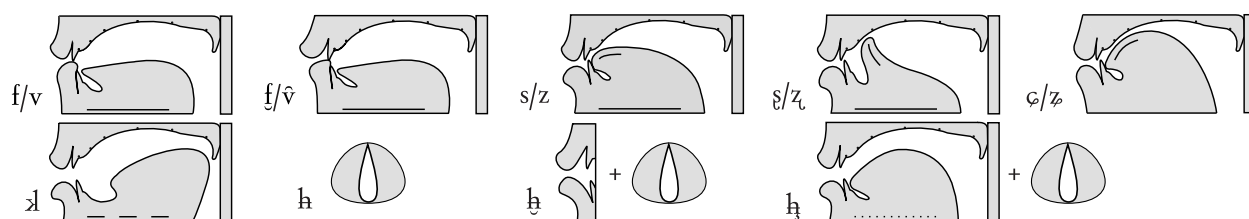
Constrictives

3.4. There are three constrictive phonemes, all phonemically voiceless: /f, s, ʃ/; they are also phonetically voiceless in syllables with a primary stress, [ˈ\$], as well as in those with a secondary stress *preceding* another [ˈ\$], ie /(')\$'\$\$/ [ˈ\$]\$ & [ˈ\$]\$\$ (see 4): *fǎnfù* [ˈf̥ɛ̃n̥|f̥ɥ], *sān suì* [ˈs̥ɛ̃n̥|s̥w̥ɛ̃ɪ], *shìshǐ* [l̥ʃ̥ɥ̥|ʃ̥ɥ̥]; *fánfēnshù* [ˈf̥ɛ̃n̥|f̥ɛ̃n̥|ʃ̥ɥ̥], *sān-sì suì* [ˈs̥ɛ̃n̥|s̥ɥ̥|s̥w̥ɛ̃ɪ], *shǒushùshǐ* [l̥ʃ̥ɥ̥|ʃ̥ɥ̥|ʃ̥ɥ̥]. It is to be remembered that, in a typical [ˈ\$]\$\$ trisyllable, the former [ˈ\$] is auditorily less prominent than the latter, due to the influence from the rhythmic structure.

3.4.1. When the most prominent syllable *comes first*, as in the /'\$(\$(\$))/ [ˈ\$(\$)] & [ˈ\$(\$,\$)] sequences, the *following* less stressed syllables are articulated with much less energy, too. They are of course very weakly stressed *per se*; but, regardless of phonological considerations, in fact two degrees of ‘weakness’ exist: in relatively slower speech, we actually have /'\$(\$)/ [ˈ\$,\$] instead of the [ˈ\$(\$)] found in normal conversation. Similarly, as enunciation speeds up considerably, it is easy to observe that all trisyllabic sequences tend to converge to one similar ratio, corresponding to a [ˈ\$(\$,\$)] pattern, in terms of phonatory energy. It is precisely such syllable-strength attenuation that produces a kind of ‘voicing gradation’ similar to that observed in stop and stop-strictive contoids: [ˈC̥, ɪ̥C̥, ɔ̥C̥] *f*, /s/ [ˈs̥, ɪ̥s̥, ɔ̥s̥] *s* and [ˈC̥, ɪ̥C̥, ɔ̥C̥] *x*, /ʃ/ [ˈʃ̥, ɪ̥ʃ̥, ɔ̥ʃ̥] *sh*.

A few examples: [ˈma·ɣan], [ˈma·vən] *máfan*, [k̚q̚.ɔ̚.z̚u], [-z̚u] *gàosu*, [ˈk̚wan·ʒ̚i], [-ʒ̚i] *guānxi*, [ˈl̚jan·ʒ̚u], [ˈl̚jan·ɰ] *liángshǐ*. This is obviously referred to fluid and moderately fast speech; while, in very slow *or* careful speech, the speaker may want to avoid any voicing at all, *producing voiceless contoids anywhere*. Many Chinese teachers adopt such a style of enunciation to teach beginners the fundamentals of Mandarin pronunciation.

fig 3.5. Constrictive and semi-constrictive contoids.



3.4.2. Nevertheless, we do not support the [ᶱᶱ, ᶱᶱ, ᶱᶱ] pattern, for it might sound artificial to native speakers. Secondly, it would require the same amount of phonatory energy on every syllable, *regardless of stress*. But stress is quite relevant instead, in order to produce tones and intonation properly.

As speed increases, a simple two-way gradation spontaneously emerges, with *fortis* [ᶱᶱ] and *lenis* [ᶱᶱ] as extremes; syllables with a secondary stress freely oscillate between the two (for, as usual, paraphonics and the position in a rhythmic group do play a role): [ᶱᶱ], [ᶱᶱ, ᶱᶱ], [ᶱᶱ]. Due to its relative ease and flexibility, such a pattern seems very suitable indeed for teaching purposes.

At an even higher speed, the taxophonic scenario becomes a bit more complicated, with the pronunciation of professional broadcasters and dubbers diverging from that of neutral (or nearly-neutral) native speakers in spontaneous conversation. The ‘professional’ choice favours [ᶱᶱ], [ᶱᶱ, ᶱᶱ], [ᶱᶱ, ᶱᶱ] (again, with smooth and flexible transitions between the edges), from which we have devised the normalized scheme ‘[ᶱᶱ, ᶱᶱ, ᶱᶱ]’ presented above.

Statistically, [v, z, ʒ, ɰ] are much more common in a preintoneme than in an intoneme. Besides, ‘laxer’ articulations abound in informal and colloquial environments: more [ᶱᶱ], then, together with ‘reduced’ vocoids. However, we will keep the three-level gradation, in transcribing isolated words or short sentences, in order to highlight the *possible* phonetic ‘extremes’, at all times. In transcribing audio files (see ¶ 6), the symbols relevant to actual realizations will be employed, of course.

3.4.3. Even lenitions –both for voiced and voiceless phones– can be heard sometimes: [f̥, ʋ̥; s̥, z̥; ʃ̥, ʒ̥; ʂ̥, ʐ̥]; however, it is hard to define an explicit criterion for this last point, as it appears to be a side effect of many combined factors. Such phenomena interact with appropriate paraphonic features, to convey a sense of straightforwardness ^{or} familiarity; in more formal delivery, however, the same style could be regarded as sloppy and, if excessive, even regional (eg too ‘Pekingese’). Foreign speakers should therefore pay attention not to make a bad imitation of native speakers, just for the sake of sounding ‘cool’, or simply more fluent.

3.4.4. Though not precisely a constrictive, [ɣ] *h* must be included here as well, since we have chosen to describe it –primarily– as a *uvular semi-constrictive*, [ɣ]. Our choice has a decisive advantage, hidden behind this unitary symbol, free from annoying diacritics: as a matter of fact, it is extremely difficult, if not impossible, to intentionally produce such a phone in current speech, with the absolute certainty of not articulating a full *constrictive*, [χ], or an *approximant* instead, [ɣ̠]. In syllables with a primary stress, native speakers themselves oscillate airily, probably more than with any other consonant. Thus, [ɣ] appears to be the ideal ‘normalized’ choice, capable of suggesting both the main taxophones, at the same time; secondly, by attempting to articulate [ɣ], foreign learners will happen to automatically reproduce the same –very natural– taxophonic variation observed in neutral pronunciation. ⁴

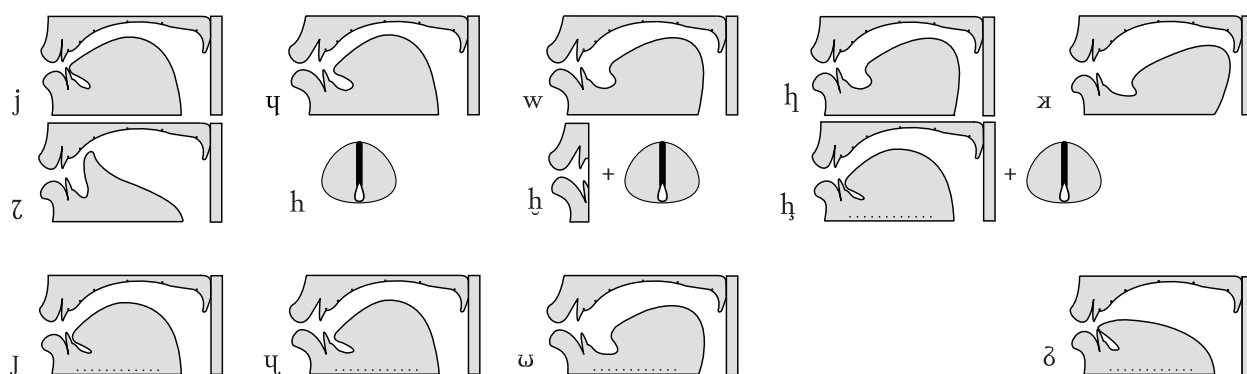
Statistically, [ɣ] is most common in syllables with a secondary stress, while we frequently find a *velar approximant*, [ŋ], in unstressed ones (be noted: a change in PLACE –not MANNER– OF ARTICULATION). Then, for *héhǎo* /ʰk̚x̚k̚lɑo/ and *hǎohē* /k̚lɑoʰk̚x̚/, we actually have [ʰk̚x̚k̚lɑo, ʰk̚x̚k̚lɑo]; while, for *huǒhòu* /k̚wɔk̚k̚ou/ and *còuhē*, /tshou.k̚x̚/, we have [k̚wɔk̚k̚ou, ʰtshou; tshou.k̚x̚, ʰtshou].

In fact, any of the taxophones of /h/, [ɦ, h], would work as well, at the phonemic level; but phonetically, only [ɦ] is auditorily ‘strong’ enough to replace [h], not so much [ɣ] (while, certainly, [ɣ] would not be fit). More examples: [ʰk̚x̚ʃ̚t̚] *héhì*, [ʃ̚t̚k̚x̚] *shìhé*, [ʰk̚x̚w̚k̚w̚ɛ̚ɛ̚] *hòuhuǐ*, [ʰp̚t̚.k̚x̚(ʃ̚), ʰt̚k̚x̚] *bòhē*, [ʰt̚k̚x̚.k̚w̚, ʰk̚w̚k̚x̚] *shìhòu(r)*, [ʰʃ̚an\k̚l̚ɛ̚] *shānghài*, [k̚w̚k̚n\k̚w̚ɛ̚] *huǎnghuà*, [ʰt̚i\k̚z̚an\k̚w̚ɛ̚] *Xīzàngghuà*, [ʰk̚x̚\tsh̚u] *hē cù*.

Approximants

3.5. In our analysis, Mandarin Chinese has five approximant phonemes (fig 3.6). The first one is postalveolar, /ʐ/ [ʐ] *r*: [ʎʐ̚] *rè*, [ʐən] *rén*, [ʎwʐ̚] *rùn*, [ʐoŋ] *róng*, [ʎʐ̚] *rì*, [ʐʌŋʐ̚ʂoʊʐwʌʌŋ, ʐʌŋʐ̚ʂoʊʐwʌʌŋ] *réngrán róuruǎn*. The various occurrences of /ʐ/ at the end of a syllable –ie the *érhuà* phenomenon– are treated in depth in Ch 2, since most speakers actually realize /V(V)ʐ[#], VNʐ[#]/ not as sequences ending in [ʐ], but rather as peculiar vocoids featuring both lateral contraction and a slight post-alveolarization.

fig 3.6. Approximant and semi-approximant contours (with [ʐ], for the theoretical /w/ [ʐ])



3.5.1. Besides [ʐ], the corresponding *post-alveolar constrictive* [ʎ] is an important taxophone of syllable-initial /ʐ/; in fact, there is a certain degree of free variation, with the possibility of an intermediate articulation too, ie the *semi-constrictive* [ʎ̥] (though there is no real need to employ a distinct symbol, in ordinary transcriptions). However, the constrictive is *not* the primary rendition of Chinese *r*, contrary to what nearly all phoneticians ^{or} sinologists insist on presenting in their analyses, where only ‘/ʎ/’ and ‘[ʎ]’ are generally employed. ⁵

From a certain point of view, that is an understandable choice: for one thing, ‘/ʎ/’ looks like the logical counterpart of /ʃ/, in terms of phonemic opposition (which theoretical linguists and acoustic phoneticians consider to be more than enough); secondly, the *offIPA* inventory lacks a unitary symbol for ^{can}[ʐ], and the only two ‘candidates’ worthy of consideration, *off*[ɹ, ɹ̥], are not entirely adequate... not to mention that there is little agreement, among *offIPA* enthusiasts, on what ‘[ɹ, ɹ̥]’ precisely stand for. As far as ^{can}[ɹ] is concerned, we employ it univocally for British English *r*, which is certainly post-alveolar, but also *laterally contracted*, and it has a

certain amount of *lip rounding*, thus it is different from [ʒ] — cf British English [ˈrɑːo] /ˈrɑːo/ *row* ‘dispute’ vs Chinese [ʒɑo] /ˈzɑo/ *rào* ‘coil’.

Nevertheless, another problem arises with ‘/ʒ/’: how to represent *érhuà* phonemically? No doubt, ‘/kʰʒ/’ would seem a bit exaggerated as a phonemicization for *gēr*. The sources consulted so far show a variety of solutions, from ‘/ɹ/’ or ‘/ɹ̥/’ (with the possible ambiguities mentioned above), to an embarrassing ‘/r/’. Sometimes, the question is ignored altogether, and the *érhuà* ends up being banally transcribed with a mixture of *pīnyīn* and a few scattered *offIPA* symbols, not always with a clear distinction between phonemes and phones (even [r]!). It is self-evident that /ʒ/ has the great advantage of being adequate anywhere (just one symbol), and of suggesting the most common *real* pronunciation (with no diacritics).

3.5.2. Following, we have /j, w/ [j, w], and also /ɥ/ [ɥ], for it cannot be considered to be a simple taxophone of /j/ (in that case, a syllable like /ɥɛ/ *yüe* should be re-interpreted as */jɥɛ/, with no real advantages). In weak syllables with a ‘zero toneme’ –but not before a pause– the following *semi-approximant* taxophones are possible: *semi-palatal* [j], *semi-post-palatal rounded* [ɥ], and *semi-velar rounded* [w]. A few examples: [ˌjan˥˩˥˩˥˩, ˨˩˥˩] & [˨˩˥˩˥˩] *yǎngyang*, [˨˩˥˩˥˩, ˨˩˥˩] & [˨˩˥˩˥˩] *diēdie*, [˨˩˥˩˥˩˥˩, ˨˩˥˩] & [˨˩˥˩˥˩] *jiějie*, [˨˩˥˩˥˩˥˩, ˨˩˥˩] & [˨˩˥˩˥˩] *libian*; [˨˩˥˩˥˩˥˩] *yǔyǔ*, [˨˩˥˩˥˩˥˩˥˩] *xüěqüè*, [˨˩˥˩˥˩˥˩˥˩, ˨˩˥˩] & [˨˩˥˩˥˩˥˩] (& [˨˩˥˩˥˩˥˩˥˩, ˨˩˥˩]) *dìxióng*; [˨˩˥˩˥˩˥˩] *wáng*, [˨˩˥˩˥˩˥˩] *huī*, [˨˩˥˩˥˩˥˩˥˩, ˨˩˥˩] & [˨˩˥˩˥˩˥˩] *tǎolun*.

3.5.3. Some might even argue that /j, ɥ, w/ are not truly necessary, and could be replaced by their ‘vocalic counterparts’ /i, y, u/, at least at the phonemic level. But that could hardly be accepted; however, instead of simply saying ‘no’, we prefer to offer our readers a chance to judge by themselves. According to such a minimalist approach, then, /ʰ(C)JV(V), ʰ(C)JVN/ should be rendered as ‘/ʰ(C)VV(V), ʰ(C)VVN/’.

A few concrete examples will instantly underline what the basic fault is, in this kind of analysis: ‘/tʃhiɛn/’ *qiàn*, ‘/sɥɛ/’ *xüě*, ‘/tʃɥwan/’ *zhuāng*. Apparently, there is no major difference with respect to our /tʃhɛn, sɥɛ, tʃɥwan/; but, let us examine the phonetic reality and listen carefully: in normal speech, we would normally encounter [tʃhɛ̃n, sɥ̃ɛ, tʃɥ̃wan]. Surely, the most prominent segments are the vocalic nuclei, namely [ɛ̃, ɥ̃, ɔ̃], where most of the phonatory energy insists and focuses on.

Nevertheless, should we rigorously accept something like ‘/tʃhiɛn, ʃsyɛ, ʔtʃuan/’, the phonetic outcome would inevitably be [tʃh̥iɛ̃n, ʃs̥ỹɛ̃, ʔtʃuañ], with major prominence on [ɲ̥ ʃ̥ỹ̃, ʔu]. It could not be otherwise, because /iɛ, yɛ, ua/ are certainly *diphthongs*, as much as /ɛi, aɛ, aɔ, ɔu/ are. It is self-evident that such an absurdity could be avoided only by means of a (no less problematic) phonemicization as ‘/ʔtʃhiɛn, ʃsyɛ, ʔtʃuan/’. But, even so, we would expect, as a primary rendition, something like [tʃh̥iɛ̃n, ʃs̥ỹ̃ɛ̃, ʔtʃuañ]: in other words, *three dissyllabic sequences containing hiatuses*, while all native speakers perceive and treat them as *purely monosyllabic* words, instead!

Of course, that similar sequences can occur in extremely slow *or* emphatic speech, often with an intentional *staccato* enunciation, does not prove at all that /j, ɥ, w/ are superfluous; rather, the contrary is true, provided everyone accepts that phonology must represent what is typical and frequent, not just incidental.

Laterals

3.6. For the last manner (of articulation) necessary for Mandarin, we find just one phoneme, /l/ [l] (which remains unchanged even before /i, y; j, ɥ/): [ʔlɑɛ] /ʔlɑɛ/ *lái*, [ʔlɔŋ] /ʔlɔŋ/ *lóng*, [l̥jɛ̃ɲ̥] /l̥jɛ̃ɲ̥/ *liǎn*, [l̥ỹ] /l̥ỹ/ *liù*. Native speakers of English should pay attention, moreover, not to pronounce Mandarin *l* as the so-called English ‘dark’ *l*, found in preconsonantal or prepausal position: *velarized lateral*, [ɭ], or –less conspicuously– *velarized semi-lateral*, [ɭ̥]. Though neutral English /l/ is [l] too, when preceding a vowel, there are some English speakers –especially Americans and Scottish– who happen to produce slightly velarized phones, if a Chinese /l/ occurs next to back vowels, or to the back approximant /w/. It is a speech defect with no phonemic implications, but it is particularly annoying in terms of phonetic authenticity and acceptability, and must be avoided at all costs.

fig 3.7. Lateral contoids (with a useful comparison with phones occurring in English).



Casual speech

3.7. We present hereby some supplementary information, regarding *casual speech*, ie the kind of pronunciation normally employed in informal situations, where native speakers not only talk faster than usual, but also most frequently tend to ‘indulge themselves’ in relaxed articulations.

What has been said so far, with respect to vocalic and consonantal reduced forms in ordinary pronunciation, largely applies to casual speech too, of course (as well as the typical paraphonic setting with *raised larynx*, ⟨.:⟩, not only for women; other Chinese languages, instead, use a normal –or modal– laryngeal setting, or even one with a lowered larynx). In addition to that, non-native speakers should be able to recognize two less common types of reduction, affecting vowels and consonants *in fast speech and in a pre-intoneme* (ie not in an intoneme), for they can make Chinese less easy to understand. They are not at all indispensable to acquire a good pronunciation; however, should foreign learners want to apply them in an appropriate manner, naturalness would certainly be improved.

First, the *partial devoicing of close vowels* after /p(h), t(h), k(h); ʈ(h), ʨ(h); f, s, ʃ, ʂ/, ie not after sonants, and with the fourth toneme, particularly its ‘long’ taxotone [˨˨˨] (see below): /i, y, u, u/ [i̥, y̥, u̥; ɨ̥, ʏ̥; ʉ̥; ɤ̥, ʊ̥]. Be noted that partially devoiced vowels do not feature any ‘creaky voice’, ie laryngealization, [ʲ], as their fully voiced counterpart would, when associated with low taxotones.

A few examples: [˨˨˨pʰi˥˥˥] /pʰi˥˥˥/ *pìrú*, [˨˨˨b̥u˥˥˥f̥a˥˥˥] /p̥u˥˥˥f̥a˥˥˥/ *bùfǎ*, [˨˨˨thi˥˥˥] /thi˥˥˥/ *tìshēn*, [˨˨˨d̥u˥˥˥k̥h̥ɔ˥˥˥] /tu˥˥˥k̥h̥ɔ˥˥˥/ *dùkǒu*, [˨˨˨d̥ʒ̥y̥˥˥˥thwan] /tsy˥˥˥thwan/ *jùtuán*, [˨˨˨k̥hu˥˥˥t̥ʂ̥a˥˥˥] /k̥hu˥˥˥t̥ʂ̥a˥˥˥/ *kùchǎ*, [˨˨˨g̥u˥˥˥k̥h̥ɔ˥˥˥] /ku˥˥˥k̥h̥ɔ˥˥˥/ *gùkè*, [˨˨˨t̥shu˥˥˥t̥ʂ̥i˥˥˥] /t̥shu˥˥˥t̥ʂ̥i˥˥˥/ *cìjǐ*, [˨˨˨d̥ʒ̥u˥˥˥m̥u˥˥˥] /tsu˥˥˥mu˥˥˥/ *zìmǔ*, [˨˨˨t̥ʂ̥u˥˥˥f̥an] /t̥ʂ̥u˥˥˥fan/ *chùfàn*, [˨˨˨d̥ʒ̥i˥˥˥n̥t̥swɔ˥˥˥] /tsin˥˥˥t̥swɔ˥˥˥/ *jìngzuò*, [˨˨˨f̥u˥˥˥p̥u˥˥˥] /fu˥˥˥pu˥˥˥/ *fùbù*, [˨˨˨ʂ̥i˥˥˥n̥j̥a˥˥˥] /sin˥˥˥jan/ *xìnyǎng*, [˨˨˨ʂ̥y̥˥˥˥t̥ʂ̥h̥y̥˥˥˥] /sy˥˥˥t̥ʂ̥hy̥˥˥˥/ *xùqǐ*, [˨˨˨ʂ̥u˥˥˥f̥u˥˥˥] /ʂ̥u˥˥˥fu˥˥˥/ *shùfú*, [˨˨˨x̥u˥˥˥t̥ʂ̥a˥˥˥] /x̥u˥˥˥t̥ʂ̥a˥˥˥/ *hùzhào*.

Besides, in unstressed syllables with ‘zero’ toneme, again *not in an intoneme*, *approximant realizations* are possible for /s/: [z] instead of [s], and [j] instead of [ʃ]; for /ʃ/, [ɰ] is seldom reduced to [z]. In very fast and uncontrolled pronunciation, stop-strictives may feature the same reductions: then, /ts(h)/ [z] also for [ts, dz], and [j] for [tʂ, dʂ]; /t̥ʂ(h)/ [z], also for [t̥ʂ, d̥ʂ].

A few examples: [˨˨˨k̥ɑo˥˥˥s̥u˥˥˥] /k̥ɑo˥˥˥s̥u˥˥˥/ *gàosu*, [˨˨˨p̥ei˥˥˥d̥z̥u˥˥˥] /p̥ei˥˥˥d̥z̥u˥˥˥/ *bēizǐ*, [˨˨˨t̥ʂ̥h̥ɔ˥˥˥t̥ʂ̥ɔ˥˥˥] /t̥ʂ̥h̥ɔ˥˥˥t̥ʂ̥ɔ˥˥˥/ *cóucou*, [˨˨˨t̥ɔn˥˥˥z̥i˥˥˥] /t̥ɔn˥˥˥si˥˥˥/ *dōngxi*, [˨˨˨j̥ɛ˥˥˥n̥d̥ʒ̥u˥˥˥] /j̥ɛ˥˥˥n̥d̥ʒ̥u˥˥˥/ *yǎnjing*, [˨˨˨t̥ʂ̥h̥u˥˥˥t̥ʂ̥y̥˥˥˥] /t̥ʂ̥h̥u˥˥˥t̥ʂ̥hy̥˥˥˥/ *chūqu*, [˨˨˨t̥ʂ̥u˥˥˥ɰ˥˥˥] /t̥ʂ̥u˥˥˥ɰ˥˥˥/ *zhǐshǐ*, [˨˨˨p̥an˥˥˥d̥z̥u˥˥˥] /p̥an˥˥˥t̥ʂ̥u˥˥˥/ *bāngzhu*, [˨˨˨t̥ʂ̥h̥i˥˥˥n̥t̥ʂ̥u˥˥˥] /t̥ʂ̥h̥i˥˥˥n̥t̥ʂ̥u˥˥˥/ *qīngchu*.

NOTES:

¹ In LIN, H-Y. (2007) *The Sounds of Chinese*, the author seems to share our view on *j, q, x*.

² cf PULLEYBLANK, E. G. [1991] *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early mandarin*, p. 6-8.

³ For instance, ‘[Ç]’ are used in CAO, W. (2000) *Hanyu fayin yu jiuyin* (bud, sadly, not in *Hanyu yuyin jiaocheng* and *Chinese Pronunciation Practice*), as well as in BRUCE, J. P. (1930) *Chinese*, a pioneering Mandarin course by the Linguaphone Institute, and entirely transcribed in IPA (included a tonetic notation which even clearly marks the various tones!).

Sadly, by using ‘[Ç]’, the authors normally mean ‘*unvoiced unaspirated* contoids’, not our *partially devoiced* contoids. In other words, ‘[b̥, d̥, ʝ̥]’ &c should *visually* prevent foreign learners –especially native speakers of English– from realizing ‘[p, t, k]’ as [ph, th, kh], due to the interference of common orthography. In fact, the diacritical circling simply means *devoicing*, in the official system, with no further indication whether devoicing is partial or total; thus, *off*[Ç] may correspond to *can*[Ç], but also *can*[Ç̥] (and *can*[Ç̥], too). Consequently, *off*[b̥, d̥, ʝ̥] &c can be legitimately –though very *questionably*– used, in place of [p, t, k].

However, we cannot help regarding this expedient as a ‘not-so-clever trick’. Surely, it avoids one mistake, but it easily leads to make another one, and for the same *visual* factor mentioned above: who can misinterpret ‘[p, t, k]’ as [ph, th, kh] could certainly misinterpret ‘[b̥, d̥, ʝ̥]’ as [b, d, ʝ], too, which is simply not correct in [ʃ, ʒ].

⁴ Most handbooks on Chinese pronunciation simplistically resort to ‘[x]’ (as they do with ‘[ŋ]’, instead of *can*[ŋ]), with no gradation or further tones. We made a different choice, but we admit that [x, ŋ] can be perfectly legitimate, if used by non-native speakers, for purely communicative purposes.

⁵ *off*[ɹ] corresponds to a number of different *can*IPA symbols, but could be a decent replacement for *can*[ʒ], in simplified transcriptions. However, it is the questionable use of *off*[ɹ] to render the American English *r*, whose articulation has little to do with a true [ʒ], that forces the *off*IPA aficionados to resort to [ɹ̥], nonetheless always implying something like ‘[ɹ̥]’, ie an *approximant*, or a *semiconstrictive* phone.

4. Chinese tonemes, tones, taxotones & stress.

4.1. One of the most remarkable features of all Chinese languages is that they employ different *tonemes*, besides vowels and consonants, as phonemic features: a variation in pitch, then, is capable of distinguishing syllables that otherwise would be perfectly homophonic. It is a fascinating challenge for foreign learners, but also a typical source of endless frustration. That is why it is paramount to practice all tonemes –and their phonetic counterparts: the *tones*– with as much care and perseverance as possible.

Whether ancient Chinese was a tonal language since the beginning is still under debate, but it appears plausible that tonemes developed at a relatively early stage of its long history. However, what can be targeted as the ‘ancestral’ dialects spoken by Hàn tribes some three-thousand years ago, very presumably, possessed quite a large syllabic inventory, and the mere combination of vowels and consonants would have been more than enough to avoid ambiguous homophonies.

4.1.1. Tonemes are very common in the East-Asian language area, to which Chinese belongs: from Tibetan to Thai, from Burmese to Vietnamese, we invariably find quite restrictive syllabic inventories enriched by the use of tonemes. Otherwise, the lexemic monosyllabism so typical of those languages would simply be impossible, as it is the case in Japanese: though the Japanese *akusento* has a tonal nature, that language has no tonemes, in the strict sense of the term. No surprise, then, that polysyllabic words abound in purely indigenous Japanese vocabulary: *atarashii* /ɛtɛɾɛʃiɪ/ [ɛ'tɛɾɛ.ʃiɪ] ‘new’, *kokoro* /kɔkɔɾɔ/ [kɔ'kɔɾɔ] ‘heart’ vs *shin*, the Sinic pronunciation of both the corresponding Chinese characters (as in *shinbun* /siŋbɯŋ/ [ɕiŋɸɯŋ] ‘news’, and *shinri* /siŋɾi/ [ɕiŋɾi] ‘mentality, psychology’).¹

4.1.2. As far as Chinese ‘dialects’ are concerned, some of them count as many as eight tonemes; others, like Shanghainese, are more concerned with the tonal contour of entire words, rather than that of single syllables, and the five traditional tonemes nowadays interact in a simple two-way opposition. *Pǔtōnghuà* and the Peking dialect lie somewhat in between, with only four marked tonemes, or *shēngdiào*

[ˈʂʌN\|tjʌŋ], collectively called *sì-shēng* [ˈsɨ̃ʍ̃ˉʂʌN] ‘four voices’. They are indicated by their ordinal number –T-1, T-2, T-3, T-4– or can be named after four of the eight tonemes of traditional Chinese phonology (since these names do not necessarily reflect modern pitch contours, it is better to avoid them):

– *yīnpíngshēng* [ˈjim̄ˈpɦɪNˉʂʌN, ˈpɦɪNˉ-, ˈpɦɪNˉ-], ‘lower level’, nowadays rather a steadily *high* and level toneme, in neutral Chinese;

– *yángpíngshēng* [ˈjan̄ˈpɦɪNˉʂʌN, ˈpɦɪNˉ-, ˈpɦɪNˉ-], ‘higher level’, nowadays a mid-to-high rising toneme (in its basic form);

– *shǎngshēng* [ˌʂʌNˉʂʌN], ‘ascending’, nowadays nearly always a low and slightly descending toneme, *only seldom with a rising coda*;

– *qùshēng* [ˈtɕɦ̥ỹˉʂʌN], ‘departing’, apparently coherent with the modern steep, abruptly high-to-low falling contour. ²

Numerous expedients have been devised to indicate tonemes in Latinized transcriptions; that adopted by *pīnyīn* marks with a diacritic the most prominent *graphic* vowel in each ‘final’ (which not always coincides with the *phonic* vowel):

– T-1, *qīn* /ˈtʂɦɪn/ [ˈtɕɦɪN], and *qiān* /ˈtʂɦjɛn/ [ˈtɕɦjɛN];

– T-2, *lóu* /ˈlɔu/ [ˈlɔU], but *liú* /ˈljɔu/ [ˈljɔU] (instead of ‘*lióu*’);

– T-3, *hěn* /ˈkɦɛn/ [ˈkɦɛ̃N], but *gǔn* /ˈkwɛn/ [ˈkʷɛ̃N] (instead of ‘*guěn*’);

– T-4, *hèi* /ˈkɦɛi/ [ˈkɦɛ̃ɪ], but *huì* /ˈkɦwɛi/ [ˈkʷɛ̃ɪ] (instead of ‘*huèi*’).

4.1.3. The marked tonemes only occur in syllables with a *primary stress*, [ˈ\$], or a *secondary* one, [ˌ\$]. However, since syllables are rarely uttered in isolation, some of them may easily undergo a further accentual reduction and become *destressed syllables*, [°\$]; their phonatory energy is not strong enough to articulate a full tonal contour, and what is left is a shorter tone –with reduced intensity and no peculiar movements– whose actual pitch depends on that of surrounding tones, especially the immediately preceding ones.

The same criterion can be applied to *unstressed syllables*, *ie* those which never take a primary stress, not even when considered in isolation. Actual pitches –again– are determined by the tonal environment; but, unlike in destressed syllables, they cannot be linked to any of the marked tonemes. Therefore, another tonal group must be considered: a true *unmarked toneme*. The Chinese call it *qīngshēng* [ˈtɕɦɪNˉʂʌN] ‘light voice’; another popular term is ‘neutral tone’, as if it had no tonal identity at

all. We, on the other hand, agree with the simple definition ‘zero tone’, which –of course– has to be interpreted as ‘zero *toneme*’: T-0, underlining that it is none of the other four (and it is the least prominent, too).³

Dealing with Chinese tonemes –and all the relevant taxotones– is decidedly hard, not just for foreigners whose mother tongue does not use pitch for lexical purposes, but also for many Chinese whose native dialect has a different tonal behaviour. Very often, the more similar this is to the neutral scheme, the harder it is for the speaker to avoid cross-dialectal interference. For example, in the small borough of Lónghúzhèn, near Zhèngzhōu, Hénán Province, the locals pronounce T-1 and T-2 in a very similar way to T-3 and T-4 of *pǔtōnghuà* respectively (and the other way around): thus, they have the right tones, but these are not assigned to the right tonemes, and a sentence like *tā xiúé Hànyǔ* ‘he/she is learning Mandarin’ becomes a puzzling **tǎ xiúé Hànyǔ*.

However, controlling the pitch to use tones correctly is not an option: practicing tones ‘the hard way’ from the start is the only way –if not to master them– at least to attain a decent level of accuracy. It will not be easy, and a certain skill on the side of the learner is certainly required, but any effort is worth the price.

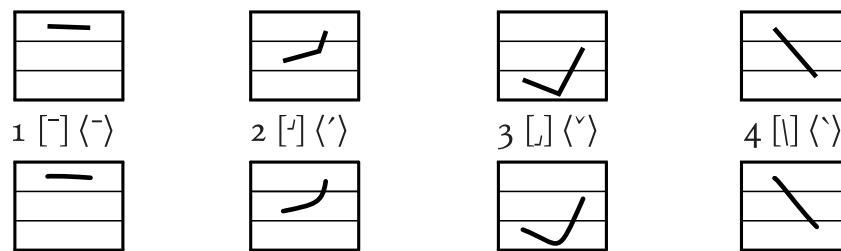
Taxotones and stress

4.2. Many students learning Chinese often fail to realize that the four tonemes are not fixed and unchangeable entities; in an attempt to ‘sound proper’, then, they tend to produce ‘full’ pitch contours under any circumstances, just like in uttering isolated syllables. Too many teachers do not sufficiently insist on the importance of linking the proper (taxo)tones in a natural speech *continuum*, provided their pupils are able to put ‘the right *shēng* (toneme) on the right *zì* (character)’. Surely, the speakers will manage to communicate unambiguously, but such a kind of Chinese will sound emphatic and unnatural.

Instead, each toneme features at least two contextual variants, or *taxotones* (better than simply *allotones*), which have to be mastered, and employed in connection with all the relevant vocalic & consonantal reductions, where possible (not to mention, of course, that grammar and vocabulary must be kept under control, as well): quite a challenge!

4.3. For each one of the marked tonemes, there is a primary ‘full’ taxotone (fig 4.1), to be used in [‘\$] (and –of course– in emphatic enunciation): [ˉ], [ˊ], [ˋ], [ˋˊ]; the last two are further limited to pre-pausal position (or, to be more precise, when –at the end of a rhythmic unit not semantically and structurally linked to the next one– an interruption in the tonal chain intervenes).⁴

fig 4.1. The four marked tonemes (schematic and realistic).



Consequently, in normal speech, the various tonemes are expressed by different taxotones capable of interlinking more smoothly, thanks to a number of *continuous and reciprocal small adjustments* (or *sandhi*, should we want to show-off by using that technical term borrowed from ancient Sanskrit linguists), which produce a rather complex and composite pitch contour. Taxotones (see fig 4.2) are definitely not an unnecessary nuisance: without them, it would be uneasy and tiresome for both the speaker and the listener.

In syllables with a secondary stress, the primary taxotones are somewhat reduced, both in duration and in intensity; we have, then, [‘\$]: [ˊ], [ˊˊ], [ˋ], [ˋˊ] (not considering other taxotones modified by intonation). Since the ‘zero’ toneme is regularly associated with [◦\$], the actual pitch contour is so feeble that it can be represented by a single dot, and only three symbols are necessary: [ˊ], [ˊˊ], [ˋ] (see fig 4.8-9); but, in slow speech and in an intoneme, T-O *may* receive more phonatory energy and substantially merge with [‘\$]: [ˊˊ], [ˊˊˊ], [ˋˋ].

To be entirely rigorous, a notation like ‘[ˊˊ]’ would suffice to express a secondary stress with medium pitch, but we prefer ‘[ˊˊˊ]’ to keep consistency with [ˊˊ], [ˋˋ]; besides, in non-tonal languages, [ˊ] may be completely unrelated to pitch. For example, in English, *Mexico* /ˈmɛksɪkəʊ/ is actually [ˈmɛksɪˌkəʊ] (American and ‘international’ pronunciation; while neutral British English would have [ˌkɜːʊ]); but what is the actual pitch of the three syllables, when the actual intonational patterns are applied? Considering three British marked intonemes only, we would hear:⁵

- *conclusive*, /'mɛksɪkσσ./ [ˈmɛksɪkσσ_], ie [[mɛk_sɪ_kσσ]];
- *interrogative*, /'mɛksɪkσσ?/ [ˈmɛksɪkσσ·], ie [[mɛk_sɪ'kσσ]];
- *continuative*, /'mɛksɪkσσ,/ [ˈmɛksɪkσσ], ie [[mɛk_sɪ_kσσ]].

4.4. In order to fully understand the mechanism of stress in Chinese, we must be able to isolate and analyze its fundamental stress units, usually called *rhythmic groups*. These do not necessarily correspond to *words*, no matter which particular language is concerned, but to syllabic strings within actual utterances, with at least one prominent syllable –bearing most of the phonatory energy– and a number of less prominent ones; all the more so in Chinese, where even the very concept of ‘word’ is particularly vague (and words of more than two syllables have more than one primary stress, although, in these cases, the dominant stress is the last one).

Let us now examine some rhythmic groups containing marked tonemes only. First of all, a disyllabic sequence most commonly follows the [\$\$] pattern: *zhīyüán* [ʧʷɿ˥˥ɿɿ] ‘employee’, *liánxì* [lʲɛ̃n˥˥ɿ] ‘to link’, *hēibǎn* [xɛ̃l˥˥pʰãn˥˥] ‘blackboard’, *jiàoshī* [ʧʷjʰãʃ˥˥ɿɿ] ‘teacher’; *wǒ shuō* [wɔ̃ʃ˥˥wɔ̃] ‘I say’, *huài hái* [xwʰãɛ˥˥kʰãɿ] ‘naughty boy’.

A useful comparison between the basic tones and three their important taxotones is clearly shown in fig 4.2: attention should be paid to [˥] /˥/ ǎ, a T-3 that closely resembles a T-2 (though not identical, as too many textbooks still describe it), and occurs only before another T-3; then, the so-called ‘half T-3’ is [˥] /˥/ ǎ, when preceding some other marked tonemes. Finally, there is even a ‘half T-4’, [˥] /˥/ à, which only occurs before another T-4. Quite evidently, the pitch contours of these taxotones make tonal adjustment easier, by avoiding inconvenient sequences; as a matter of fact, [˥] and [˥], in particular, [˥] and [˥] are nothing but ‘condensed’ versions of their corresponding full forms.

4.5. It may be argued that secondary taxotones are not fundamental, and therefore not actually necessary to foreign learners. On the other hand, if one aims at pronouncing Chinese convincingly, not just at mimicking its sounds, it is clear that four tonal contours are simply not enough. Besides, even if taxotones were not used by non-native speakers, they should anyway be able to recognize those taxotones, when listening to the Chinese. Besides, other peculiarities found in complete utterances become far less cryptical if one starts from a seven-element inventory.

fig 4.2. Tones and taxotones.

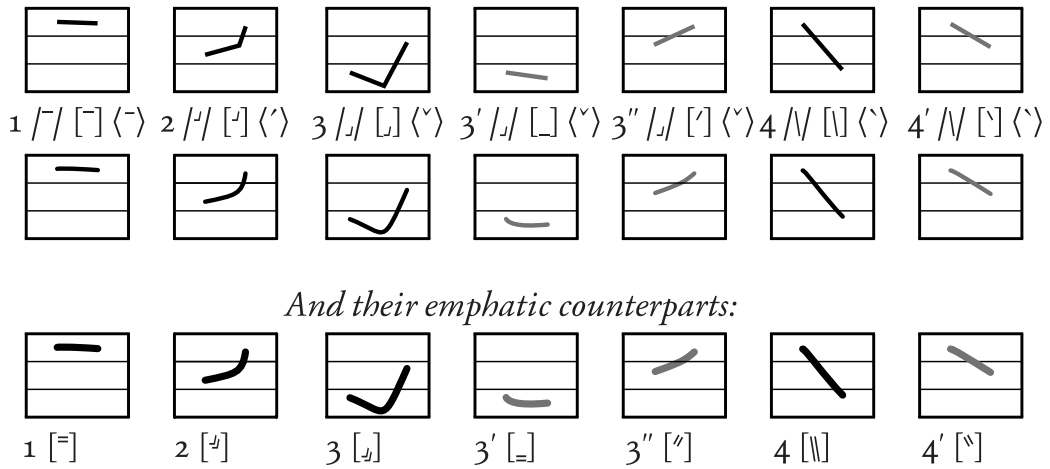
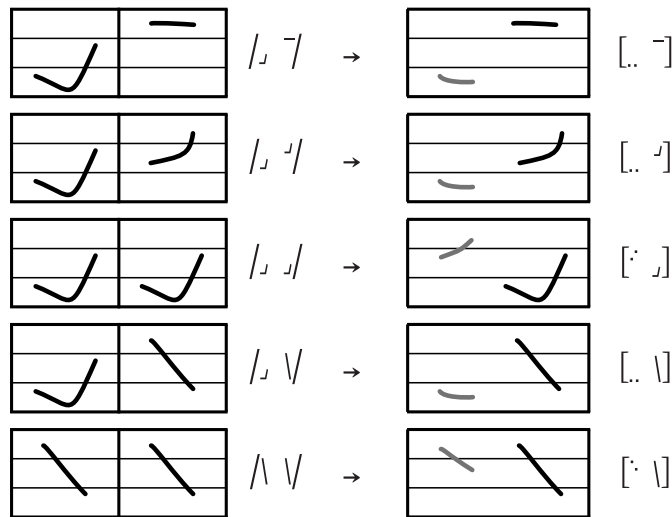


fig 4.3. Fundamental taxotones in dissyllabic sequences.



A better explanation of how the taxotones behave is provided by fig 4.3, as well as by the following examples, where secondary-stress reduced taxotones are also employed. It is clear that [˩] → [˩] (& [˩˩]) allows for a smoother transition from /˩/ to /˥/, or /˨˨˨˥/, or /˩˩˩˥/: *huǒchē* [˩˩˩˥˩˩˩˥˩˩˩˥], *děngyú* [˩˩˩˥˩˩˩˥˩˩˩˥], *kǎoshì* [˩˩˩˥˩˩˩˥˩˩˩˥]. Similarly, a sequence like /˩˩˩˥/ would be unpractical, if realized as such, or even with the taxotone just seen, ie [˩˩˩˥]; instead, [˩˩˩˥] decidedly improves the enunciation: *shǒubiǎo* [˩˩˩˥˩˩˩˥˩˩˩˥] (again, with a secondary stress and a reduced duration: [˩˩˩˥] → [˩˩˩˥], instead of [˩˩˩˥]).

In /˩˩˩˥/, there is a generally descending contour: of course, the theoretical [˩˩˩˥] would require an uncomfortable ‘zig-zag’ movement, should we actually try to

make our voice ‘climb up’ (from [.] to [ˊ]) before falling down again. Instead, with [ˊ], a very straightforward [ˊ\$] is obtained: *fàng jià* [ˊfɑŋ\|tɕjɑ]. In short, *seven tones* and as many symbols are required for [ˊ\$]: [ˊa], [ˊa], [ˊa, ˊa], [ˊa, ˊa], [ˊa, ˊa]; but only *five*, for [ˊ\$]: [ˊa], [ˊa], [ˊa, ˊa], [ˊa, ˊa] (because [ˊ] & [ˊ] → [ˊ]; [ˊ] & [ˊ] → [ˊ]); and *three*, for [ˊ\$]: [ˊa], [ˊa], [ˊa].⁶

4.6. Generally speaking, trisyllables composed by marked tonemes follow the [ˊ\$] pattern, which is substantially coherent with that of disyllables, as it can be analyzed as ‘[ˊ\$]+[ˊ\$]’. It is easy to observe that, unlike English, no syllable is entirely deprived of stress, especially if speaking at a normal speed and the word occurs in an intoneme: again, *Mexico* [ˊmɛks\|kɔ] vs *Mòxīgē* [ˊmɔˊ\|cǐˊkɛ]; moreover, both [ˊmɔˊ] and [ˊkɛ] bear a primary stress, though the former is auditorily less prominent, due to the influence of rhythm.

Other examples: *yōumògǎn* [ˊjɔ\|mɔˊkǎn] ‘sense of humor’, *féngrènjī* [ˊfɛŋ\|zɛnjī] ‘sewing machine’, *huǒchēzhàn* [ˊxwɔˊ\|tɕhɛˊtɕʌn] ‘train station’, *fùmǚshuō* [ˊfɯ\|mɯˊʃwɔ] ‘(my) parents say’. It is interesting to notice that the primitive lexemes, if taken in isolation, would switch from [ˊ\$] to the [ˊ\$] pattern: *yōumò* [ˊjɔ\|mɔ] ‘humor’, *féngrèn* [ˊfɛŋ\|zɛn] ‘to sew’, *huǒchē* [ˊxwɔˊ\|tɕhɛ] ‘train’, *fùmǚ* [ˊfɯ\|mɯ] ‘parents’.

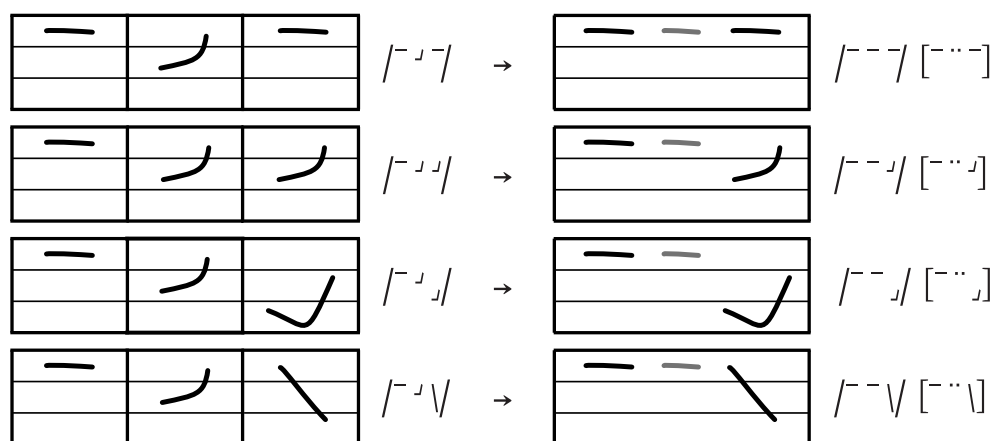
4.6.1. Besides the norms explained hitherto, which occur inevitably and nearly irrespective of enunciation speed, there are other particular cases, involving [ˊ], where a certain degree of fluency and speed is required for the tonal adjustment to take place: [ˊ] → [ˊ].

In fact, native speakers are rarely conscious of converting a T-2 or a T-3 into a T-1 (actually, the overlapping is taxotonic, not tonemic): for them, that is a perfectly natural outcome, in their speech habits; but, if asked to repeat the same sequence at a slower pace, paying attention to its taxotones, they invariably avoid [ˊ] in favor of the original taxotones.

Foreign learners must pay attention to this last point, and are even prone to omit the ‘flattening’ presented below, if they do not pronounce the language fluently enough, or are not entirely proficient in linking words together: otherwise, the use of [ˊ], though in the right place, might result puzzling and be judged (paradoxically!) as wrong.

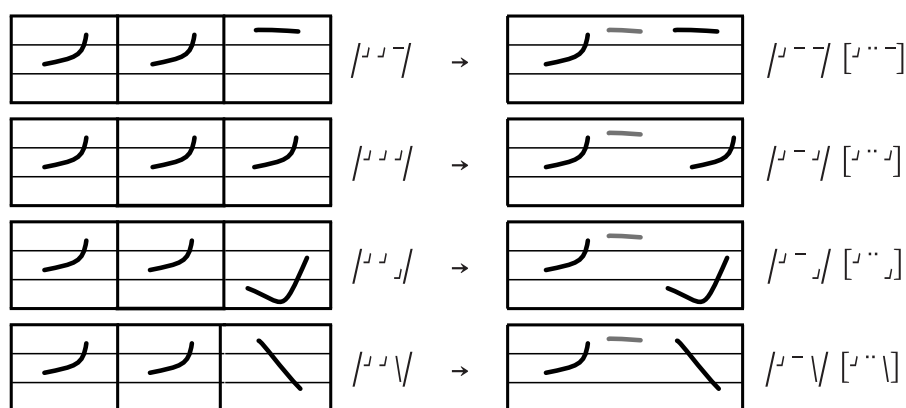
First of all, let us examine fig 4.4, presenting the behaviour of /-‘\$’\$/, ie /-‘\$’\$/ followed by any other marked toneme: [-‘\$’\$/] → [-‘\$’\$/], eg *xiānrénzhǎng* [-‘çjɛŋ˥˥ɿŋ˥˥ʂɑ̃˥˥], *dōngnánfēng* [-‘tɔŋ˥˥nɑŋ˥˥fɛŋ˥˥], *sānniánjǐ* [-‘sɑŋ˥˥nǐɛŋ˥˥tɕi˥˥], *xīhóngshì* [-‘çi˥˥xɔŋ˥˥ʂɿ˥˥]. As said, in slower speech, a slightly less spontaneous but ‘safer’ realization would resume [-‘nɑŋ˥˥-], [-‘nǐɛŋ˥˥-], [-‘çjɛŋ˥˥-], [-‘xɔŋ˥˥].

fig 4.4. Taxotones in /-‘\$’\$/+‘\$’\$/ trisyllabic sequences.



Besides, fig 4.5 presents the behaviour of /‘\$’\$/ [-‘\$’\$/] → [-‘\$’\$/], eg *Méi Lánfāng* [‘mɛi˥˥lɑŋ˥˥fɑŋ˥˥], *wánquán dòng* [‘wɑŋ˥˥tɕɥɑŋ˥˥tɔŋ˥˥], *hái méi lái* [‘kʰɛi˥˥mɛi˥˥lɑi˥˥], *xúnyáng-jiàn* [‘çyŋ˥˥jan˥˥tɕjɛŋ˥˥].

fig 4.5. Taxotones in /‘\$’\$/+‘\$’\$/ trisyllabic sequences.

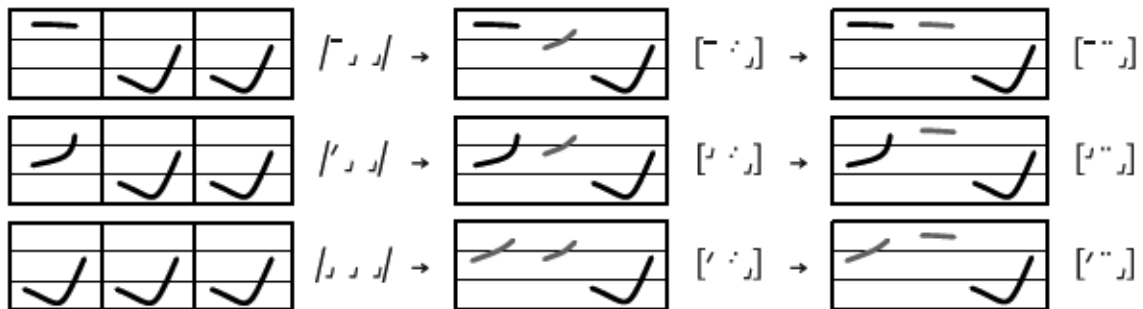


Finally, fig 4.6 describes what happens when the /‘\$’\$/ sequence follows one of the first three marked tonemes: [‘\$’\$/] → [‘\$’\$/] → [‘\$’\$/] (with [‘\$’] ≠ [˥˥]), eg *tā hěn hǎo* [‘tʰɑ˥˥hɛŋ˥˥kʰɑ̃˥˥], *hánshǔbiǎo* [‘kʰɑŋ˥˥ʂu˥˥pǐɑ̃˥˥], *wǒ yě yǒu* [‘wɔ˥˥jɛ˥˥jɔ˥˥] (be noted: [‘wɔ˥˥], not [‘wɔ˥˥]; depending on rhythm and emphasis, the alternative reading

[_wɔ̄˥˥jɛ˨˥ɔ̄˨˥ɔ̄˨˥] is possible as well). Even if a rhythmic unit includes more than three syllables with marked tonemes, it would have a primary stress on the last syllable, anyway, and another on the first one; ie with a strong tendency to keep most of the phonatory energy on the ‘outer ends’, and to reduce it in the middle. This means that, in a four-syllable sequence, the usual distribution is [‘\$,,\$,\$\$], rather than [‘\$,\$,\$\$], which would not sound very smooth; but with five syllables, there is no reason to subvert the [‘\$,\$,\$,\$\$] pattern resulting from the simple juxtaposition of [‘\$,\$,\$\$] + [‘\$,\$\$], or the spontaneous adjustment of [‘\$,\$\$] + [‘\$,\$,\$\$] (lexical factors play their role too).

A practical example: taken separately, *pīnyīn* ‘to spell words out (phon[et]ically)’ and *zìmǔ* ‘alphabet’ are naturally [‘\$,\$\$]: [ˈpʰɪŋ˥˥jɪn], [ɿ˨˥ɔ̄˨˥mɯ˨˥]. By combining them into *pīnyīn-zìmǔ* ‘phono-alphabetic script’, ‘romanization’, the theoretical (yet possible) rendering [ˈpʰɪŋ˥˥jɪn ɿ˨˥ɔ̄˨˥mɯ˨˥] normally changes into [ˈpʰɪŋ˥˥jɪn ɿ˨˥ɔ̄˨˥mɯ˨˥]. In English we have examples like *the seven-year itch* [ðəˈsevən˥˥jɪtʃ ˈɪtʃ] or *fifth-generation computers* [ˈfɪfθ ɔ̄˨˥ɛnə˥˥ɪfɪŋ kʊmˈpʰjʊtəz], which combine their plain elements: *seven* [ˈsevən], *year* [ˈjɪə], *itch* [ɪtʃ] and *fifth* [ˈfɪfθ], *generation* [ɔ̄˨˥ɛnə˥˥ɪfɪŋ], *computer* [kʊmˈpʰjʊtə].

fig 4.6. Taxotones in /\$/+/,,\$\$/ trisyllabic sequences.



It has to be underlined, both from the learner’s and teacher’s point of view, that rhythmic groups are not detached from intonation and paraphonics; instead, they influence one another: therefore, and not surprisingly, it is frequent that more than one stress arrangement is possible for similar utterances. Let us observe *Máo Zédōng zhǔxí* [ˈmɑo˥˥ɔ̄˨˥˥tɔŋ ˥˥ɔ̄˨˥˥ci] ‘chairman Mao Zedong’, with simple juxtaposition: [‘\$,\$\$] + [‘\$,\$\$]. But, in the basic sentence *tā shì zhǔxí* ‘(he/she) is (the) chairperson’, we have [‘\$,\$+\$,\$\$]: [ˈtʰa˥˥ɿ˨˥˥ ˥˥ɔ̄˨˥˥ci]; in fact, [ˈtʰa˥˥ɿ˨˥˥ ˥˥ɔ̄˨˥˥ci] (omitting sentence intonation symbols) would not be a mere variant –as in *pīnyīn-zìmǔ*– but would im-

ply that ‘he/she *is* (the) chairperson’, though with less assertiveness than [ˈtʰaʅʂuː ˌd͡ʒuˈci] ‘(he/she) *definitely is* (the) chairperson’, with an extra-strong, emphatic stress on *shì*, the copula.

On the other hand, one might say *Lǐ fùzhǔxí* [lǐ ˌfʊˌd͡ʒuˈci] ‘(the) vice-chairperson Li’, where two contiguous syllables retain a primary stress, for they logically belong to different semantic entities: rhythmically, however, they can be regarded as part of one sequence (even if –auditorily– with slightly different prominence).

4.7. A Chinese disyllable may feature a ‘zero’ toneme on its second syllable, which is therefore considered as *intrinsically unstressed*: /°\$/ [°\$]; in that case, the primary stress obviously falls on the first syllable. *Pīnyīn* systematically annotates unstressed syllables by leaving them without tonal marks: *wǒde, nàli, érzǐ, rènsǐ, luóbo, zhàngu, méiguī*, &c. Such typographical convention is mandatory, and (good) teachers would reject such misspellings as **érzǐ, *rènsǐ, and *méiguī*.

There are also some *de-stressed* syllables, /°\$/ [°\$] → [°\$], ie those with a marked toneme, which –under specific circumstances– may be ‘weakened’ up to the point of acting, *de facto*, as T-0 syllables. Though such a distinction is far from being indispensable, from a practical point of view, nonetheless it proves useful to thoroughly understand the complex mechanisms of Mandarin prosody. The official orthography is somewhat inconstant, on this respect: de-stressing is sometimes explicitly annotated, more often it is not (pretty much like tonal adjustments, which are given as implicitly understood by the reader/speaker). A few examples, duly transcribed, are given below.

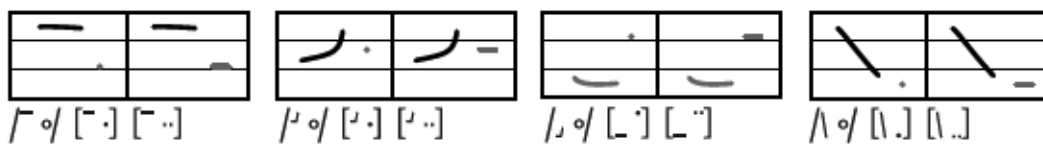
As a general rule, the actual pitch of a ‘weak’ syllable is determined by the tone of the syllable that comes before it; in fast connected speech, the pitch of the following syllable may exert a certain influence as well, but always on a secondary level. As previously said, three symbols are enough: [ˌ, ˋ, ˊ]; but, phonetically, *four* different pitches may be detected, as shown in fig 4.7: *jīzǐ* [ˌt͡ɕiːd͡ʒu] ‘chicken’, *jízǐ* [ˌt͡ɕiːd͡ʒu] ‘collected works’, *jǐzǐ* [ˋt͡ɕiːd͡ʒu] ‘muntjac’, *jìzǐ* [ˊt͡ɕiːd͡ʒu] ‘bun’.

In slow speech and in an intoneme, as said in § 4.3, the last syllable may receive a secondary stress, [ˌ\$,\$], so that it becomes slightly more prominent than it usually is: eg [ˌt͡ɕiːd͡ʒu], [ˌt͡ɕiːd͡ʒu], [ˌt͡ɕiːd͡ʒu], [ˌt͡ɕiːd͡ʒu]. The difference is noticeable, even

to an untrained ear, thus foreign learners should pay attention not to abuse of this possibility, out of the contexts in which it is allowed.

By referring to fig 4.7, again, there is clearly a mid pitch after both T-1 and T-2; more precisely, a *lower mid* pitch, after T-1, and a *slightly raised mid* pitch, after T-2: *bīpo* [ˈpī·p̄σ, --p̄hσ], *píqi* [ˈp̄hi·t̄ci, --t̄chi]. After a T-3, which is automatically ‘shortened’ to [ˌ], we have a *lowered high* pitch; finally, after a T-4, the ‘dot’ lies just in the middle of the *low* band: *bīfang* [ˌp̄ˈv̄aŋ, ˌˈv̄aŋ], *bīhui* [ˌp̄ˈh̄w̄ɤ̄ɪ, ˌˈx̄w̄ɤ̄ɪ].

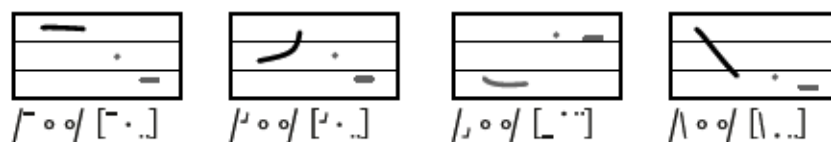
fig 4.7. Taxotones in /\$+/,\$,\$/ trisyllabic sequences.



What above mentioned, of course, works well with single rhythmic groups pronounced in isolation; but, when affected by intonation, such a kind of ‘ideal’ pronunciation is inevitably modified & compressed, and many nuances may become less peculiar, especially in current speech: even more reason, then, to ‘get by’ with just three punctiform symbols (the three-level dots we have already seen [···]).

4.8. When a marked toneme is followed by *two* unmarked ones, /\$°\$°\$, the phonatory energy is redistributed as [°\$°\$], akin to the general trend of reducing stress in the middle –as observed in /\$°\$°\$/ [°\$°\$]– while keeping most of it on the edges. As fig 4.8 clearly explains, we have *kāilái le* [ˈk̄h̄aɛ̄·l̄aɛ̄·l̄ɤ̄(ɤ̄)], *náilái le* [ˈn̄a·l̄aɛ̄·l̄ɤ̄(ɤ̄)], *zǒulái le* [ˌt̄s̄ɔ̄ɥ̄·l̄aɛ̄·l̄ɤ̄(ɤ̄)], *sònglái le* [ˌs̄ɔ̄ŋ̄·l̄aɛ̄·l̄ɤ̄(ɤ̄)], with /ˈl̄aɛ̄/ destressed to /°l̄aɛ̄/ (see below).

fig 4.8. Taxotones in /\$+/°\$°\$/ trisyllabic sequences.

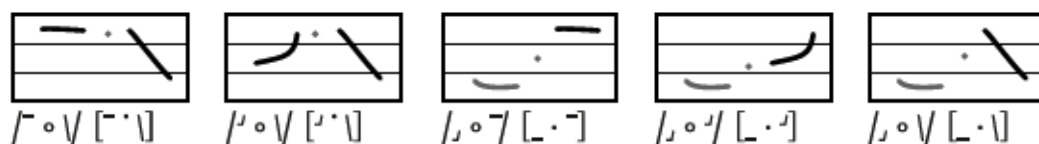


Between two marked tonemes, a T-0 (fig 4.9) is practically realized as a ‘linking taxotone’, whose pitch is directly influenced by the following tone, too, as well as by the preceding one: *chīle fàn* [ˌt̄ɕ̄h̄ūˈl̄ā{f̄āŋ}], *pá-bu-shàng* [ˈp̄h̄āˈb̄ū{ɕ̄āŋ}], *wǒde shū*

[_Lwɔ̄·dɛ̄ʒu], *liǎngge rén* [_Llǰəŋ·gɛ̄ʒɛ̄n], *kàn-de-jiàn* [_Lkʰəŋ·dɛ̄|tɕjɛ̄ŋ] (& [_Lkʰən·dɛ̄|tɕjɛ̄ŋ]).

The examples presented here lead us to say a few words more about the behaviour of ‘destressed’ syllables, which we usefully consider to be distinct from purely unstressed ones, at least conceptually: *yīnwèi* ‘because’ is regularly [_Ljɪn|wə̄j], but the alternative pronunciation [_Ljɪn·wə̄j, ·wə̄j] is permitted (with a different spelling, too: *yīnwei*). The authoritative *Contemporary Chinese Dictionary* signals these alternatives by putting a dot before a syllable with a marked toneme, as in ‘*yīn·wèi*’.

fig 4.9. Contextual pitch of a T-0 in some trisyllabic sequences.



Resultative complements are most often destressed. Taking *kàn* [_Lkʰəŋ] ‘to see’, and *qǐ* [_Ltɕh] ‘to start; to lift’, as main verbs, we can add an enclitic verb to them, like *jiàn* [_Ltɕjɛ̄ŋ] ‘to see’, or *lái* [_Llɛ̄] ‘to come’; we obtain *kànjian* [_Lkʰəŋ·dɛ̄jɛ̄ŋ, ·dɛ̄jɛ̄ŋ] ‘to see’ and *qǐlái* [_Ltɕh|lɛ̄, ·lɛ̄] ‘to rise up, to commence’ (and other meanings). But, if we insert *dé* [_Ltɕ] ‘to reach’ or *bù* [p̄u] ‘not’ between the verb and its resultative complement, a potential construction is obtained; *dé* and *bù* lose their stress, [ōdɛ̄, ōbu], while the other verb resumes the original toneme: *kàndejiàn* [_Lkʰəŋ·dɛ̄|tɕjɛ̄ŋ], *qǐbulái* [_Ltɕh|bu·lɛ̄].

Let us also compare *shōu + qǐlái + le* → *shōuqilai le* [_Lʒəu·tɕi·lɛ̄|ʒ(ɛ̄), ·lɛ̄] ‘to have collected scattered things’; if *le* comes after the main verb, the sequence *shōu + le + qǐlái* gives *shōule qǐlái* [_Lʒəu·lɛ̄ |tɕh|lɛ̄, |tɕh|lɛ̄]. A monosyllabic pronoun like *tā* ‘he, him; she, her; it’ might be inserted too: if fully stressed, we would have *shōuqi tā lái le* [_Lʒəu·tɕi·thə |lɛ̄·lɛ̄], or /·lɛ̄/ [_Lʒəu·tɕi |thə·lɛ̄·lɛ̄] (more than one arrangement is possible, depending on what the speaker wants to underline). *Tā* can be used as an enclitic too; but, since *pīnyīn* does not permit such spellings as ‘*ta*’ (unlike *lai*, *qi*, *jian*, &c), we are compelled to write *shōuqi tā lái le* anyway, for [_Lʒəu·tɕi·thə |lɛ̄·lɛ̄] / [_Lʒəu·tɕi·thə |lɛ̄·lɛ̄].

4.9. Lastly, there are some peculiar uses to be learned ‘as they are’, concerning the negative adverbs *bù* and *méi*, and the numerals *yī* ‘one’, *qī* ‘seven’, *bā* ‘eight’. Before a T-4, *bù* [p̄u] changes into /'pu/, producing a smoother combination: *wǒ bù*

qù ‘I don’t go’ /_Lwɔ̄\pu\tshy/ → /_Lwɔ̄’pu\tshy/ [Lwɔ̄’bu\tɕhý] (& [Lwɔ̄’bu\tɕhý]), with loss of stress); for a moderately stronger assertion, like ‘I do *not* go’, the spontaneous pronunciation would rather be [Lwɔ̄’pu\tɕy]. However, if the speaker wants to emphasize that he or she has ‘*absolutely no* intention to go’, *bù* may retain its toneme, possibly with two distinct rhythmic units and a potential pause, eg /_Lwɔ̄\pu. \tshy./ [Lwɔ̄\pú:: \tɕhý.] (the intonation symbols will be explained in the next chapter). The official orthography permits the written variant *bú*, as in *wǒ bú qù*, but this expedient is not systematically used.

In *x-not-x* interrogative patterns, *bù* and *méi* are generally destressed (this could be indicated by their romanization, too: *bu*, *mei*): *dǒng bu dǒng* [Lɔ̄ŋ’bu,ɔ̄ŋŋ]; *chī mei chī* [tɕhɯ·MEI tɕhɯ, ‘MEI].

The numeral *yī* ‘one’ keeps its original T-1 when the lexeme is used alone, or at the end of an utterance: *shíyī* [ʃɯ̄ji] ‘eleven’, *yī yī dé yī* [‘ji ji ·dɛ̄ji] ‘one multiplied by one gives one’ (again, please, refer to the next chapter for an explanation of our intonation symbols). Before a T-4, *yī* becomes /’ji/: *yīgòng* [‘ji\kɔ̄ŋ] ‘altogether’, *yī bàn* [‘ji\pʌŋ] ‘a half’. Before other tonemes, instead, *yī* changes into /\ji/: *yìdiǎnr* /\ji,tjəz/ [j\ɔ̄tjɛ̄ɕ] ‘a bit’, *yī tiān* [j\ɔ̄thjɛn] ‘a day’, *yī nián* [j\ɔ̄njɛn] ‘a year’. For teaching purposes, more realistic spellings are found, though they are not encountered in official orthography: *yígòng*, *yí bàn*, *yìdiǎnr*, *yì tiān*, *yì nián*.

Traditionally, *qī* ‘seven’ and *bā* ‘eight’ used to be pronounced as /’tshi, ‘pa/ before a T-4: *qīyüè* [‘tɕhì\yɛ̄] ‘July’, *bā-wèi lǎoshī* [‘bā\wɛ̄i .lɔ̄ɔ̄-ʃɯ̄] ‘seven teacher’. However, such peculiarity is now almost lost in contemporary neutral Mandarin, and could even be regarded by some as regional, or too ‘old-Pekinesque’. Official spelling does not record it, by the way, and today we only have *qīyüè* [‘tɕhì\yɛ̄], *bā-wèi lǎoshī* [‘bā\wɛ̄i .lɔ̄ɔ̄-ʃɯ̄].⁷

NOTES:

¹ See CANEPARI, L. (2007) *A Handbook of Pronunciation*, ¶ 12 ‘Japanese’

² cf HUANG, B-R. & LIAO, X-D. (2002) *Xiandai Hanyu*, p. 78-90; & LI, M. & SHI, P-W. (2004) *Hanyu putonghua yuyin bianzheng*, ¶ 9.

³ A full treatment of *qīngshēng* can be found in LU, Y-ZH. (2001) *Qingsheng he erhua*. Also, cf HE, P. (2006) *Chinese Pronunciation Course*, vol. a, ㉟ 6; & LI, M. & SHI, P-W. (2004) *Hanyu putonghua yuyin bianzheng*, ㉟ 11.

⁴ Here, too, we refer to CANEPARI, L. (2007) *A Handbook of Pronunciation*, § 11.3.3.1 *et seq*, with important enrichments and improvements.

⁵ cf CANEPARI, L. (2009) *English Pronunciation*, ㉟ 34-37.

⁶ Despite remarkable differences, in terms of tonetic notation and iconic representation of intonation, our approach has many points in common with what Cáo Wén presents in his *Hanyu Yuyin Jiaocheng*, especially § 4.4.1-4 (he even introduced the notion of ‘transitional tone’, *guòdù shēngdiào*, for the occurrences of [˙] instead of [˙]). It is sadly surprising, then, that his most recent *Chinese Pronunciation Practice* has simplified everything, ‘downgrading’ the whole treatment of tonemes and tones to nothing more than a sketchy resumé.

⁷ cf *The Contemporary Chinese-English Dictionary*, p. 26 & 1508.

5. An introduction to Chinese intonation

5.0. This chapter is presented as an *introduction* to Chinese intonation, and not as a comprehensive description of what we have observed so far. There are good reasons for keeping such a prudent approach, considering the highly peculiar mechanisms determining intonation in a tonal language. Thus, as a mark of intellectual honesty, we prefer to describe what appears to be frequent and truly useful to foreign learners, instead of diving into all the fascinating –but sometimes puzzling– phenomena of the ‘Chinese lilt’.¹

Intonation surely is what conditions the most how a language sounds, *ie* its melodic flow and its rhythm. Unfortunately, while children find it easy to pick it up just out of spontaneous imitation –sometimes even *before* mastering contours and vocoids– many adults often never attain a decent approximation to the right intonational patterns of the tongue they are learning. Obviously, there is no need to sound exactly as a native speaker, but improving intonation really makes conversation smoother and also more enjoyable, not to mention that it avoids some curious misunderstandings too, *eg* ‘are you asking me, or are you telling me?’ – ‘hey, dude, don’t speak to me *in that tone!*’ and the like.

To put it bluntly, learning intonation is a hard work. In fact, an adult brain is not as physiologically retentive as a younger one: very little can be done about that; however, a lot could be done by teachers and authors of language courses, to provide the student with –at least– a clear and reliable guide to intonation. Regrettably, a lot more should be said about how most of them parlously *fail* to do so, provided they do not omit the matter at all in the first place (which is not necessarily an evil, since a good number of handbooks just mislead the student).²

5.1. First of all, one has to bear in mind that, no matter which language we are dealing with, intonation is primarily the product of *tonality*: in other words, when the pitches of the syllables in an utterance are put together in rapid succession, a precise melodic curve is generated. Linguists analyze those curves and extract what is recurrent and typical in their shape; once conveniently normalized, then, a small

number of representative ‘shapes’ are classified as the intonational patterns of that particular language.

Intonation plays two very important roles in spoken language. First, it helps distinguish similar sentences with different communicative purposes; let us consider this minimal conversation:

– *Really?*

– *Really.*

It is clear that, from a phono-syntactical point of view, there is no difference between these two one-word sentences, apart from the way we ‘sing’ them, respectively with a rising or a falling pitch contour. According to ^{can}IPA Natural Tonetics terminology, we prefer to say that an *interrogative intoneme* [?/] and a *conclusive intoneme* [./] have been used.

Changing the pitch of a sentence does not affect its word order, ie its (*syntactical*) *linguistic arrangement* (though, in English, most interrogative sentences and their affirmative counterparts may have a different structure); however, it cannot be denied that intonation plays a ‘syntactical’ role in the aforementioned examples, as it determines the nature of a given utterance, pretty much like a phoneme is capable of determining the meaning of a single word. In this chapter, then, by saying *intonation*, we primarily mean *linguistic intonation*.

Secondly, tonality may also work to express the speakers’ feelings and mood, as well as their point of view about the matter being discussed, regardless of whether the sentence is a question or an assertion. Such a use of tonality for ‘emotional’, non-linguistic (or paralinguistic) purposes is better called *paraphonic intonation*, or simply *paraphonics* (although this last term also includes non-tonetic paraphonic features, such kinds of voice, and articulatory settings), and it is important to distinguish it from intonation proper, in order to devise a realistic and normalized picture of both. Regrettably, that is what most language handbooks tend to mix up.

All we have just said, with due differences of course, can be applied to *any* language. However, the tonal nature of Mandarin triggers an obvious question: how does intonation work in a language where most of its syllables already have a distinct pitch contour of their own? Well, *that* is where things get really tough.

5.1.1. In learning intonation, the first step is to perceive how we modulate our voice when we speak *our own* mother tongue. Surely, it would be better if everyone

knew and could use the neutral type of pronunciation, at least when it is required or advisable to do so. Anyway, the natural approach is equipped with everything is needed to analyze one's own peculiarities and compare them with the actual neutral pronunciation. For that purpose, we provide hereby the intonational patterns relevant to modern neutral British pronunciation (see below).

Clearly, this is not the easiest of tasks, since English is a worldwide, multi-centric language with more than one accepted 'norm'... not to mention those readers of ours who are not native speakers of English, and –though fluent in the language– will almost certainly speak it with a more or less clear foreign accent. In this last case, it will take longer to master neutral English intonation, but the process is substantially the same applied to native speakers.

Besides what we concisely report here, a number of useful reference books are listed in the *Bibliography*. For a thorough, clear and practical treatment of English intonation (though relatively concise, in comparison with so many whole –but incomplete– books on the subject), the most realistic and satisfactory texts are *English PronunciationS* ('EPs') and *The Pronunciation of English around the World* ('PEW', particularly, ¶ 36-49. They follow the *Natural Tonetics Method*, which is not mingled with paraphonics, by clearly separating it, though –of course they are closely related). They also include the intonation patterns of all the accents described there. It is very important to clearly know that the *NT method* has nothing to do with such unsatisfactory things as *ToBI* and other merely acoustic treatments.

Native speakers of Italian will find their language described with the same level of accuracy (and regional variants, as well) in *Manuale di pronuncia italiana*, while *A Handbook of Pronunciation* provides only seemingly less extensive analyses (because they must be considered together with § 1.1.5-8 – or, better still, with ¶ 12-14 of *Natural Phonetics & Tonetics*) also for French, German, Spanish, Portuguese, Russian, Arabic, Hindi, Japanese, and Esperanto. Finally, a large number of other languages are more concisely treated in *Natural Phonetics & Tonetics*.

5.2. Once the goal of our study has been clearly determined, an iconic representation of the intonational patterns is needed. We adopt a four-lined, three-spaced 'staff', akin to that employed in the Gregorian Chant, and based on very similar (but *relative*, not *absolute*) principles. This is called a *tonogram* (fig 5.1).

fig 5.1. A ‘graphemic-iconic’ representation of British English intonation.

| | | |
|---|--|------------------------------|
| 1 | See you on Saturday. | |
| 2 | (Will they) see you on Saturday? | |
| 3 | (If they don't) see you on Saturday... | (it'll be a total disaster.) |
| 4 | (If they don't) see you on Saturday... | (don't worry about it.) |

Each syllable with a primary or a secondary stress –and a *marked toneme*, as far as Chinese is concerned– is represented by a *line*, which, like the ancient *neumes* (ie signs), symbolizes a relative pitch and strength, according to its position, height, orientation and shape, on the tonogram. Unstressed syllables –featuring the *unmarked* toneme, in Chinese– are better represented by *dots*; though, in fact, these are to be meant as very short, nearly ‘punctiform’ segments, not extended enough to produce any peculiar contour shape.

At this early stage, however, we shall employ the alphabetic graphemes themselves, instead of line and dots, in order to make our approach even more intuitive and easy to understand. Please observe the examples in fig 5.1.

5.3. No matter which language is taken as a starting point (here, we consider English, of course), it is important that any example is repeated several times, until the speaker is able to enunciate each syllable with the right pitch. If necessary, good recordings of neutral speakers, like dubbers and experienced actors, may provide an excellent source to start from. At this point, the speaker’s own pronunciation shall be recorded and then compared to that of the neutral speakers available.

Following, one must develop a satisfactory level of proprioceptive ability. As in learning tones, ‘a good ear’ helps, but it is not indispensable, provided the learners can analyze their own voice and determine whether a certain syllable has a high or a low pitch, possibly with a falling or a rising contour. This is not only necessary, but also useful, in order to modulate one’s own pronunciation with full awareness and flexibility.

The intonation of English

5.4. From the very intuitive –but not entirely rigorous– depiction of intonation as presented in fig 5.1, let us now become familiar with more methodical tonograms: fig 5.2-3 show the four *pre-intonemes*, and the four *intonemes*, of British English (as per *EPs* and *PEW*, where neutral American English and a well-meditated ‘International’ English are described as well, together with a great number of regional accents). Each tonogram is presented twice: on the left, the strokes and dots are connected in a quite realistic depiction of the speech *continuum* (for a smoother transition from fig 5.1), while the tonograms on the right –those currently used– are more schematic, but totally sufficient for anyone who knows how to interpret them. In fact, the exact collocation of our tonographic symbols results from the ‘arithmetic mean’ of numerous utterances, once their intonational behaviour has been accurately analyzed and normalized.

fig 5.2. British English *preintonemes* (realistic and schematic).

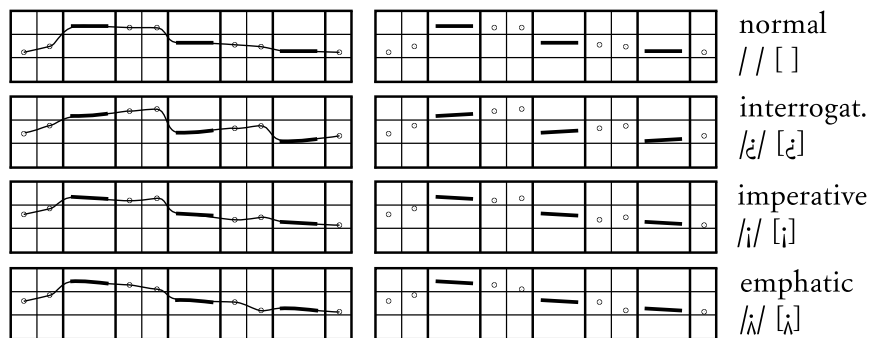
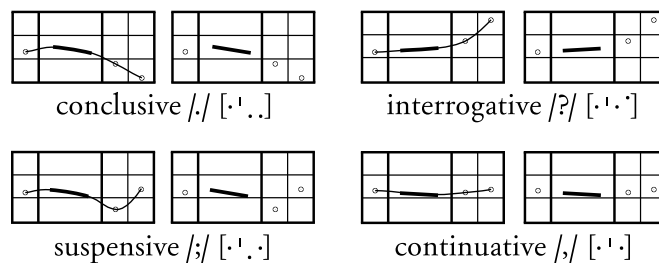


fig 5.3. British English *intonemes* (realistic and schematic).



5.5. Our approach to intonation implies a short introduction to what *intonemes* and *preintonemes* are. Roughly speaking, both can be described as the *intonational units* in which a sentence can be ‘broken down’ and analyzed, irrespective of syntactical and lexical boundaries (mirrored by traditional punctuation marks, in the

written language). A preintoneme and an intoneme taken together are usefully called an *intonation group*.

In non-tonal languages, an *intoneme* is the intonational unit whose ‘tonal core’ coincides with the pitch contour of the *last fully stressed syllable of an intonation group*, therefore called a *tonic syllable*. A number of unstressed (or weakly stressed) syllables, each with a relative height, can surround the *tonic*; in that case, stressed and unstressed syllables interact to create the *tonal contour*, distinctive of that intoneme.

An ‘ideal’ sequence would feature one *pretonic* syllable before the ‘core’, and two *posttonic* ones after it (further classifiable as *internal* and *terminal posttonics*). Of course, monosyllabic *intonemes* exist, too; if so, the *tonic* is often slightly lengthened, in order to accommodate the entire contour, which –in turn– tends to be compressed, both in its length and extension.

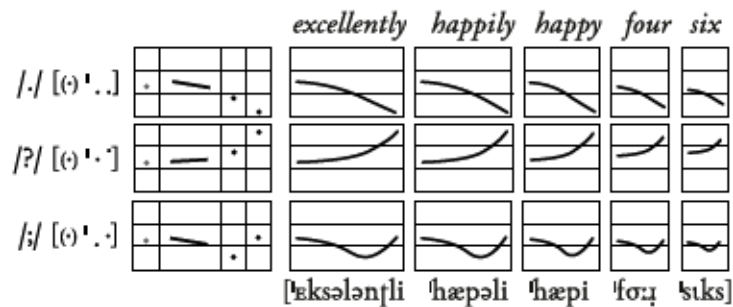
As a general rule, there are four *intonemes* in any language, whose symbols end a transcription: *conclusive* |.| – *interrogative* |?| – *suspensive* |;| – and *continuative* |,| (each will be explained below). Sometimes, and especially in tonal languages, the difference between certain intonemes can be slight, and even partially blurred by paraphonics. However, it is advisable to keep using all four intonemes anyway, since they remain *functionally* distinct, even when there may be a *tonetic* overlapping.

As its name suggests, a *preintoneme* is the intonational unit *before an intoneme in the same intonation group*, more or less with the former introducing the latter. Obviously, their symbols begin a transcription (and, as they are different from those of the intonemes, there is no risk of confusing them). For example, in English and many other non-tonal languages, the ‘foreshadowing’ nature of the *interrogative preintoneme*, |ǂ|, is well demonstrated by its shape, which appears as a compressed and shorter version of the *interrogative intoneme*, |?|. Similarly, the *unmarked preintoneme*, logically represented by a ‘zero’ symbol, | |, is somewhat reminiscent of the *conclusive intoneme*, |.|, and the *continuative intoneme*, |,| – with which it is most often associated.

Finally, there are the *emphatic preintoneme*, |ǁ|, and the *imperative preintoneme*, |ǃ|: some may argue that these are not purely intonational units, but rather something more on the paraphonic side. Surely, |ǁ|, |ǃ| are employed where paraphonics naturally plays a heavy role, yet there is an unmistakable *linguistic* feature to be considered, and that we regard as prevailing.

Now, let us examine a sentence like *his cousin's name is Bartholomew* /h[ɹ]kəzɹz 'nɛ[m ɹzbɑ:θɒləmjʊu/. The intoneme corresponds to the full name of *Bartholomew*, while the preintoneme is everything prior to it: *his cousin's name is...* If the example were *his cousin's name is Dick* /h[ɹ]kəzɹz 'nɛ[m ɹd[ɹ]k/, the intoneme would be *is Dick*. The tonic, then, would absorb part of the tonal contour that, under normal circumstances, is 'carried' by the posttonic(s). In consequence, when only one syllable is present (as in the answer to *what is his cousin's name?* – *Dick*), the result is a fusion of the expected pitch patterns, which nevertheless maintains the characteristic movements, in an attenuated form (fig 5.4).

fig 5.4. Intonemes with different number of syllables (from tonic to posttonics).



Once again, we remind our readers that the parts of an intonation group do not necessarily respect lexical or syntactical boundaries, as the following examples clearly show:

That's my favorite dictionary [ˈðæts məʃəvɛɹɪvɹɹɹ ˈdɹ[k]ʃnɹɹi.];
That patient thinks he's Giuseppe Verdi [ðæpˌphɛɹɹʃnt ˈθɹɹks ɹzɹʒɹmuˈsɛpi ˈvɛɹɹdi.].

In fact, the intonemes in these utterances are, respectively: [ɹɹɹˈdɹ[k]ʃnɹɹi.] (*-rite dictionary*) and [iˈvɛɹɹdi.] (*[-p]e Verdi*); while the preintonemes are [ˈðæts məʃəvɛɹɹv] (*My favo-*) and [ðæpˌphɛɹɹʃnt ˈθɹɹks ɹzɹʒɹmuˈsɛp] (*That patient thinks he's Giusep-*).

5.6. Let us observe, then, the sequence *see you on Saturday* in fig 5.1, where the graphemic text is arranged in the form of an intonation curve (instead of lines and dots, which will be the norm elsewhere). The *conclusive* utterance contrasts with the *interrogative* one (of a 'total' question); then, we find a *suspensive* utterance, to be opposed to a *continuative* one.

5.7. In the last two sentences, the semantic importance of what follows (given in parentheses) is fundamental, whether it is expressed out loud, or it remains implicit, instead. Evidently, the role of syntax itself is irrelevant, in this case.

The crucial point is that a *suspensive intoneme* is characterized by an element of *suspense*, as its name implies, ie a decidedly greater and more immediate anticipation; while a *continuative intoneme* precisely aims at smoothing down the transition from one unit to another. As a matter of fact, a *continuative intoneme* can be viewed as the ending part of an *unmarked preintoneme*, which is ‘promoted’ to an intonemic status for being comparatively more prominent, thanks to a slightly larger amount of phonatory energy.

5.8. In neutral British English, the *conclusive intoneme* is falling, /./ [·' .] (see fig 5.3, as well as eg 1, 3, 4, in fig 5.1): *Bartholomew* [bɑːˈθɒləmjʊ..] (which, to be meticulous, means [[·bɑː ˈθɒl ə mjʊ]]). At the very opposite, the *interrogative intoneme* is rising (/?/ [·' ·']), as in the one-word question *Bartholomew?* [ɛ̃bɑːˈθɒləmjʊː] (ie [[ɛ̃·bɑː ˈθɒl ə ˈmjʊ]]).³

The *suspensive intoneme* is falling-rising, /;/ [·' .]: *Bartholomew...* [bɑːˈθɒləmjʊː] (ie [[·bɑː ˈθɒl ə mjʊ]]). It is to be underlined that a certain resemblance between [·' ·'] and [·' .] is undeniable, as the posttonics are rising in both cases: however, this happens at different heights, and with a different contour. In fact, to have a totally similar pattern, /;/ should produce something like [· _ .] or [· \ .], which is clearly not the case. That is why we disapprove of the classical British School approach –shared by many, alas!– of reducing [·' ·'] and [·' .] to ‘two sides of the same coin’, merging both in a very broad & general ‘rising intonation unit, $\overline{\uparrow}$, or ‘/’, or ‘↑’.

5.9. fig 5.2 and the second example of fig 5.1 also feature the *interrogative preintoneme*, /ɛ̃/, whose characteristic movement anticipates (as said, in an attenuated shape) that of the following interrogative intoneme. Thus, /ɛ̃ ?/ is precisely the typical intonation pattern found in ‘total’ questions, those which one would normally reply to by simply saying ‘yes’ or ‘no’, ‘maybe’, ‘I do not know’, and the like: *Another cup of tea? – Yes, please.*

But /?/ does not occur anyway. In fact, /ɛ̃/ is peculiar enough to ‘stand alone’, as in the so-called *partial questions*, whose interrogative nature is self-evident, due to the presence of at least one ‘question word’, such as *why, when, who, how...* An an-

swer is required for that ‘part’ only, not on the whole content of the utterance. Then, by asking *What would you like to drink?*, the speaker expects us to name some kind of beverage, like *A cup of tea*, not to confirm our ‘willingness to drink’.

Generally speaking, basic partial questions end with a *conclusive intoneme*, /ǿ .;/; but, as this might sound too blunt, or even rude (depending on the communicative context, the relationship between speakers and listeners), a ‘gentler’ version exists, employing a *continuative intoneme*, /ǿ ,/ –or, more ‘colorfully’– an *attenuated interrogative* one, /ǿ ?°/.

Plain /ǿ ?/ should be carefully avoided in partial questions, as it would sound redundant, or convey another meaning, eg *How much?* /ǿhao‘mætʃ./ (or /-‘mætʃ/, /-‘mætʃ?°/; asking for a price) vs /ǿhao‘mætʃ?°/ (implying: *Are you really asking me ‘how much?’? It’s a gift!*). It is clear that the last mentioned *How much?* is a total question, after all, and such a reply as *Yes, that’s what I asked* would be perfectly normal. If said in a surprised and unbelieving way, the structure would become /ǿhao‘mætʃ?°°/ (possibly with an emphatic stress, and –no doubt– with a suitable paraphonic setting, such as *raising* or *expansion* of pitch).

A lot more should be said and explained; instead, this short –and purposely incomplete– introduction precisely aims at spurring our readers to read up more on the intonation of their own mother tongue (by referring to the various books suggested above) as we finally introduce the intonational patterns of Mandarin Chinese.

fig 5.5. Difference between total questions (1) and partial questions (2).

| | | |
|---|------------------|----------------------|
| 1 | (Will they) | see you on Saturday? |
| | | |
| 2 | (Why won’t they) | see you on Saturday? |
| | | |

The intonation of Mandarin Chinese.

5.10. At this point, we give it for granted that our readers are fully acquainted with the characteristics of the *tonemes*, and with their essential role as ‘semantically pertinent components’ of any Chinese syllable (including those with ‘no tone’, or better with a ‘zero toneme’), just like vowels and consonants. However, as in non-tonal languages, pitch modulations are also employed for intonational purposes.

This double use of tonality is actually very complicated for non-native speakers, including the Chinese whose native ‘dialect’ is remarkably different from *pǔtōnghuà*: inevitably, intonation overlaps and intertwines with the tonal flux, affecting the height and duration of the basic tones and their taxotones. More importantly, the intonation patterns can modify the actual *shape* of syllabic pitch contours, but always within an acceptable range of variation, so that their ‘tonemicity’ is preserved, and all elements remain recognizable as such. In other words, while a toneme’s *relative pitch* does not change, with respect to other tonemes in the same intonation group, its *absolute pitch* must conform to the relevant intonational ‘grid’. Such ‘grids’, which obviously correspond to Mandarin *preintonemes* and *intonemes*, are clearly depicted in fig 5.6.

However, compared to the English tonograms seen above, these ones feature a striking –and, apparently, inexplicable– difference: they are *empty*! Actually, the intonation patterns are still there, visible and distinguishable: to put it simple and straight, the intonation does not lie in the ‘content’ of the tonogram (strokes and dots, as in non-tonal languages), but in the... ‘box’ itself, so to speak. The tonograms *can* be filled with actual tonal symbols, but only when *actual tones and taxotones are determined*, each with its distinctive shape (which will be subsequently placed in appropriate height, and slightly ‘reshaped’, where necessary).

5.11. The peculiar nature of Chinese intonation, then, makes it necessary to clarify how to interpret our ^{can}IPA intonation symbols and diagrams. In non-tonal languages, they generally represent exact tonetic values, as we have shown by accurately transcribing various occurrences of *Bartholomew*, where [·] and [·] represent the actual pitch of the two posttonic syllables (respectively, in the *interrogative* and *suspensive* intonemes).

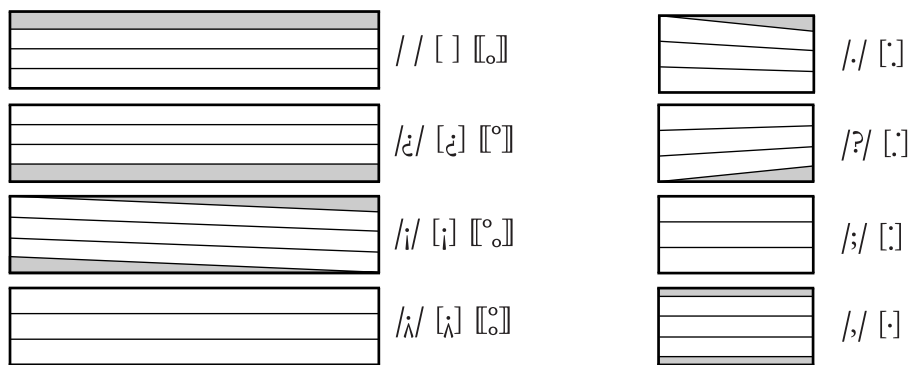
In tonal languages, instead, the intonation symbols merely suggest what sort of placement and reshaping will affect the tones and taxotones. Let us shortly review our tonemic and tonetic symbols (and the corresponding tonograms, which are more explanatory than any description): /ʔ/ [ʔ], [ʔ]; /ʔ/ [ʔ], [ʔ]; /ʔ/ [ʔ], [ʔ], [ʔ], [ʔ]; /ʔ/ [ʔ], [ʔ], [ʔ], [ʔ]; /ʔ/ [ʔ], [ʔ], [ʔ]. It is self-evident that adopting something like [·], [·], [·], [·] to show intonation would be highly confusing and misleading.

After all, what matters is having a generic indication of the pitch *range* (raised vs lowered, compressed vs expanded) ^{or} the *movement* (horizontal vs ascending vs de-

scending). Therefore, [˙] or [˘] do not refer to steep movements from up to down, and vice versa; nor should [˘] be misinterpreted as an absolute and mandatory expansion for *everything*. On the contrary, [˘] simply means that the tones and taxotones preserve their fundamental characteristics of pitch, duration and shape, for there is no compression towards the middle, as in [˘] (not to be confused with [˙\$]), a taxotone of ‘zero’ toneme, /o\$(!), nor any gradual rising, [˙], or falling, [˘].

Numerous handbooks, including many of those published in China, resort to arrow-like symbols to show intonation: ‘[↑], [↓], [↑]’; however, we are convinced that such ‘darts’ are not as intuitively self-explanatory as our [˙], [˘], [˙] (not to mention that *canIPA* better employs ‘[↑], [↓], [↑]’ for sociophonetic remarks).⁴

fig 5.6. Mandarin preintonemes and intonemes.



5.12. This being said, let us now examine the Mandarin *preintonemes*: the generic symbols employed in our phonemic transcription –/ / (for the ‘unmarked’ preintoneme) & /ɛ/, /i/, /ɿ/– can be safely employed at the phonetic level too: [], [ɛ], [i], [ɿ]. Wherever it is required and convenient to be more explicit, preintonemes can be represented by tiny circlets –actually, empty dots– in *double square brackets* (those reserved for ‘hyperphonetic’ transcriptions), in order to keep them consistent with –but also clearly distinguishable from– the black dots symbolizing intonemes.

The *unmarked preintoneme*, / / [], features a moderate compression towards the lower range of the tonogram, [˘], which makes tonal transition easier, by avoiding too large tonal ranges.

The *interrogative preintoneme*, /ɛ/ [ɛ], is compressed upwards, which means that the lower edge of the tonal band is evenly rising, [˙]. As a result, all tones will be risen too, and their typical contour will change accordingly. As in / / [], the tonal range is usefully moderated; but, thanks to a relatively (but significantly) higher

register, the listener knows it is a question, not an assertion. Indeed, the actual (ie ‘measurable’) difference can be minimal, and paraphonics may even almost nullify it; but a native speaker is still capable of telling a lowered high pitch from an elevated low one, as the human brain processes spoken language, not merely as sound waves (this being the main limit of the acoustic approach), but as the hearable counterpart of a mental act.

As far as foreign learners are concerned, no single recipe can work for everyone. Those who already speak a tonal language are accustomed to pay attention to such subtle differences in *pitch range*, while those who speak a non-tonal language tend to be more focused on *pitch movements*, as determined by a series of monosyllabic pitches (again, not just their *height*, but –more importantly– their *shape*, ie their form and orientation).

Listening to good recordings is the only and indispensable way to assimilate intonation at the passive level; but practicing intonation just through plain imitation is of little use, and potentially sterile (unless one is particularly gifted). Our tonograms represent an invaluable aid to understand *what learners should pay attention to*, when listening to Mandarin audio files, so that ‘rehearsing’ intonation may be a meditated and conscious act of speech, not a random and erratic one.

5.12.1. The *imperative preintoneme*, /_i/ [i], used in giving peremptory orders, features a gradually falling movement and a certain compression (upwards, at the beginning; then, downwards, finally): [°].

The *emphatic preintoneme* is generally not compressed at all: /_λ/ [λ] [°]; or, if we chose to consider the unmarked / / [] as the preintonemic ‘benchmark’, [°] could be regarded as an expansion of [°]. Honestly, it is nothing but a matter of point of view. What matters is that all tones and taxotones in an emphatic preintoneme will be articulated full and clear, as in single words when read aloud from a list (which is typical of the innumerable vocabulary lists that abound in Chinese language courses). Such distinctiveness is obviously coherent with the communicative purpose of emphasis.

All taxotones occurring in the various preintonemes are shown in fig 5.7, printed in gray; while the four ones printed in black, in the emphatic preintoneme, correspond to the ‘basic’ tones, as pronounced in isolation. These should be studied carefully, and taken as a reference to understand the changes affecting other taxotones in real, extended utterances.

5.13. Before introducing the four *intonemes*, we remind our readers that their tonetic notation employs a system of solid black dots, and therefore it is different from the tonemic one (while, for the preintonemes, some ‘circlets’ can be used, but are confined to extremely accurate transcriptions, as previously said).

The *conclusive intoneme* –or *assertive intoneme*, though not used exclusively in assertions– gradually compresses its tonality downwards, thus producing an overall *falling* profile: /./ [·]. In any language, that is the most frequent and versatile of all intonemes; even more so in Chinese, where concise and straightforward utterances are the norm, and intricate constructions are virtually absent. In fact, even if intonation is not directly related to syntax, it is undeniable that, the wordier a sentence is, the more frequent /,/ and /;/ will be.

5.13.1. The *interrogative intoneme* is *rising*, /?/ [·:], as its tonality is progressively raised, and the lower edge compressed upwards. As such, Mandarin Chinese conforms to a sort of ‘international’ trend, where a rising intonation is generally associated with a question, and contrasts with falling –or level– non-interrogative intonemes. (To be true, almost half of the world’s languages have *rising-falling* interrogative intonemes, /?/ [·:], where a rising part is still there, anyway;⁵ this can happen especially when the suspensive intoneme is a rising one, /;/ [·:].)

The Mandarin *suspensive intoneme* is *expanded*, without any distinctive movement: /;/ [·]. This helps express the needed *suspense*, that may be due to a sense of excitement, or fear (though not with paraphonic relevance yet), or strong opposition between two concepts, or between a condition and its outcome, and so on, depending on the context and content.

Finally, the *continuative intoneme* keeps a middle *level* tonality: /,/ [·]. It can be considered as the unmarked one (similarly to the ‘normal’ preintoneme, / / [] [·]), though on a slightly higher register, as it makes tone articulation easier, and adds no emphasis to the utterance.

fig 5.8 shows ‘isolated’ tones in black ink, as pronounced in vocabulary lists, or in /;/ [·], as well as in /;/ [·] [·]. Besides, in gray ink, all the taxotones of the four marked tonemes are shown, as they are realized in the conclusive (/./ [·]), interrogative (/?/ [·]), and continuative (/,/ [·]) intonemes. We urge our readers to carefully observe all the –often subtle– tonal alterations in their shape and height, due to different intonational constraints.

fig 5.7. Taxotones in preintonemes.

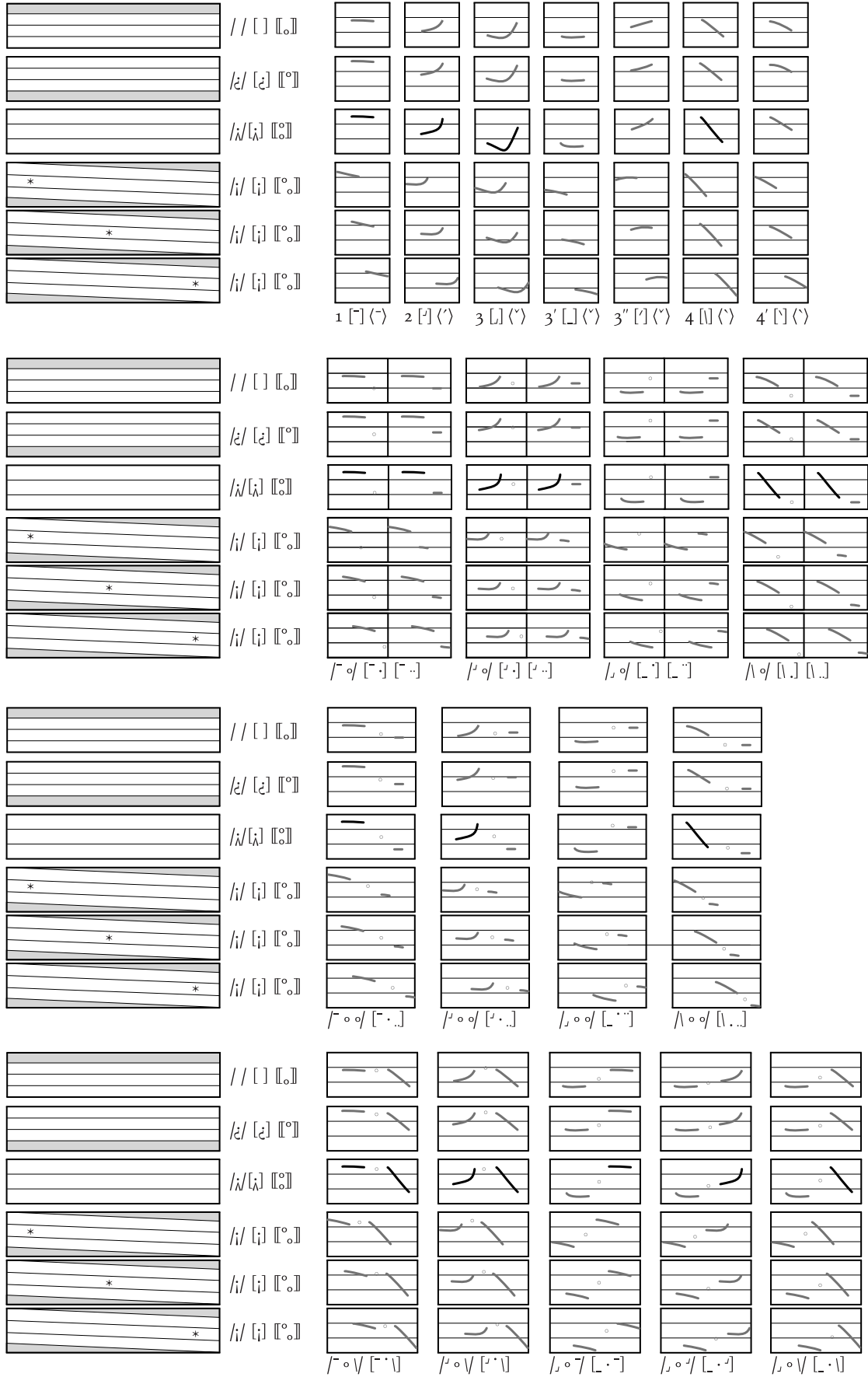
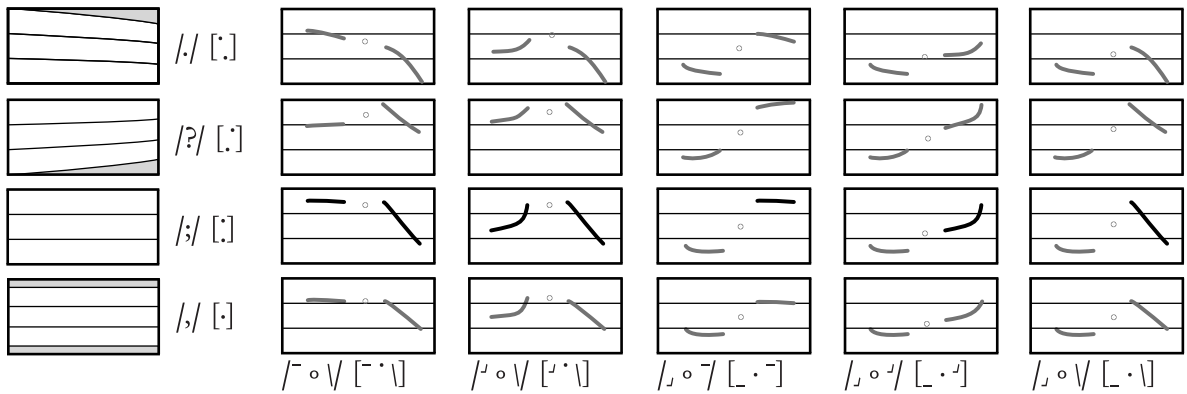
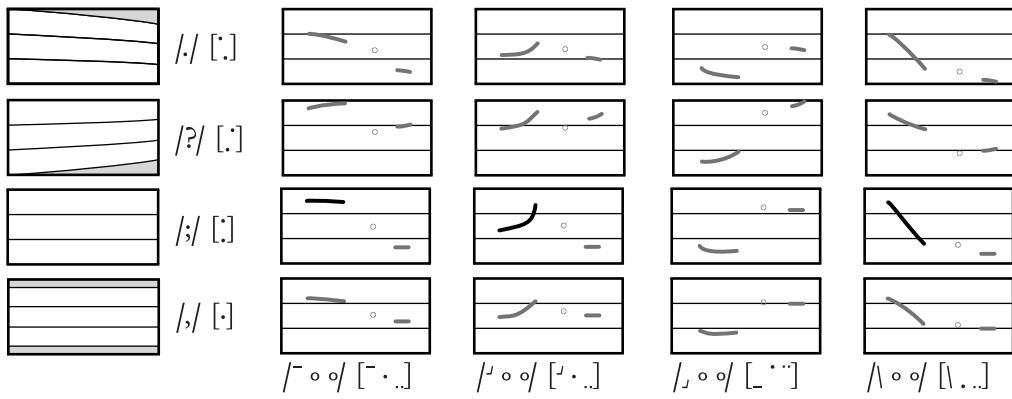
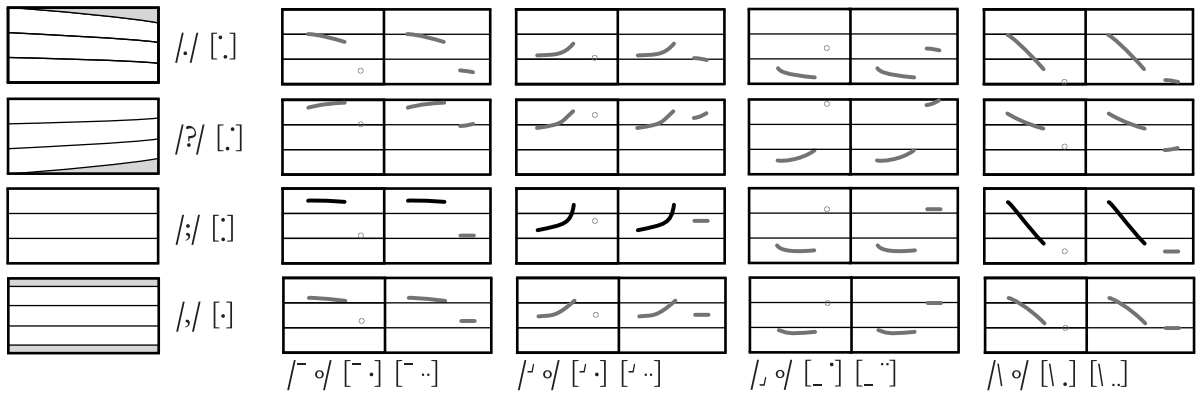
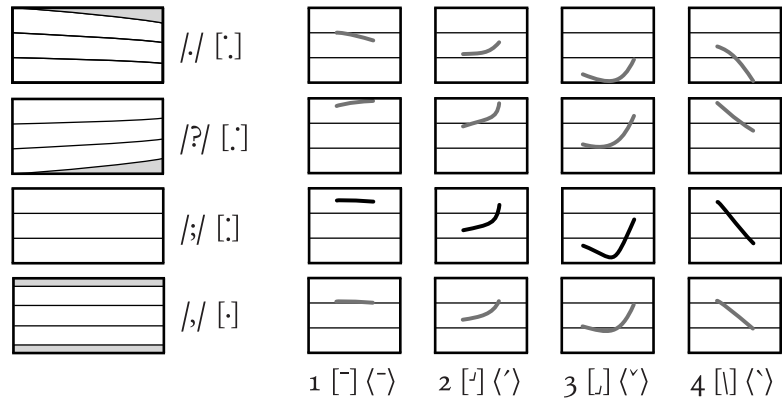


fig 5.8. Taxotones in intonemes.



5.14. It would have been pretty absurd to present the Chinese intonation structure, without describing its general nature, in the first place, and providing practical tools to reach a satisfactory awareness of the different intonational phenomena. Clearly, Chinese tonograms will mean little, at first, if not methodically compared to those of one's own language, considering some variants as well.

Undoubtedly, such comparison is easier when dealing with non-tonal languages; but, again, our tonograms offer a sure helping hand, provided there is an auditory feedback, which learners can rely on. In fact, listening to good recordings is important to get acquainted with rhythm and most recurrent syllabic structures, and intonation too, in a broader sense. But combining tones and taxotones with intonation is so delicate a task, due to continuous register changes, that it would be impossible to make any progress without the reassuring support of good native neutral voices, properly recorded. In addition to the audio quality, the quality of learning is also important: if necessary, difficult passages should be listened to, bit by bit, at lower speed, since the very beginning.

5.15. Let us deepen our understanding of how *total questions* are build up. Typically, they will feature the same syntax as that of the corresponding assertive utterances, plus an unstressed 'modal' particle, like *ma* and *ne*. While *ma* is generally quite neutral, *ne* implies that the asker is seeking confirmation to an hypothesis or, more broadly, is expecting the answerer to express an opinion. /ɿ ?/ is regularly used, but, with *ne*, /ɿ ./ (& /ɿ ,/) can be used as well, expressing different nuances:

Tā bù lái ma?

/ɿ̄˥˥˥tha|pu˥˥˥laɛ˥˥˥ma?/

[ɿ̄˥˥˥tha|b̥u˥˥˥laɛ˥˥˥ma˥˥˥]

{he/she not come *MOD}

'Is he/she not coming?'

Tā bù lái ne?

/ɿ̄˥˥˥tha|pu˥˥˥laɛ˥˥˥nɿ̄˥˥˥?/

/ɿ̄˥˥˥tha|pu˥˥˥laɛ˥˥˥nɿ̄˥˥˥./ (or /nɿ̄˥˥˥,/))

[ɿ̄˥˥˥tha|b̥u˥˥˥laɛ˥˥˥nɿ̄˥˥˥.]

[ɿ̄˥˥˥tha|b̥u˥˥˥laɛ˥˥˥nɿ̄˥˥˥.] (or [nɿ̄˥˥˥])

{he/she not come *MOD}

Tā bù lái ne? said as /ç̣⁻thə\pu¹læ^onɤ[?]/ implies a certain surprise: ‘Oh, so you mean he/she is not coming? Is that a fact?’; and it is a total question: to be precise, it is a *request of confirmation* to something that is clearly viewed as bizarre %_{or} unexpected. Instead, *Tā bù lái ne?* /ç̣⁻thə\pu¹læ^onɤ./ (or /nɤ./, with appropriate paraphonic settings), is a total question just on the surface: in fact, it is a kind of *partial* question, which we could render as ‘what (would happen) if he/she does not come?’. Clearly, we cannot simply answer that with ‘yes’ or ‘no’, but with something like ‘no big deal’, or ‘it will be a trouble, then’ &c. One could also introduce emphasis, of course: /ç̣⁻thə^opu; ç̣¹læ^onɤ./ [ç̣⁻thə^obu; ç̣¹læ^onɤ:] ‘what if he/she does not come?’; /ç̣⁻thə\pu; ç̣¹læ^onɤ./ [ç̣⁻thə\p̣u; ç̣¹læ^onɤ:] ‘what if he/she does *not* come?’.

Caution is advised when dealing with *ne*, as this particle is not primarily an interrogative one (like *ma*) and has got a wide range of possible applications. An assertive clause like *tā xúexí ne* is perfectly legitimate, but it means ‘he/she *is studying*’, *ie* an action predicated while it is happening.

The same polysemy holds true for *a* [(ʔ)a, ɥa, ɥa] (and its further euphonic variants, with a separate spelling: *ya* [ja], *wa* [wa]), and *ba*. When used in questions, these two particles imply that the answerer is supposed to express, respectively, *confirmation* or *denial* (*ie consensus* or *refusal*). Please, pay attention to how the same particle is used both in questions and in answers:

Nǐ bù lái ya?

/ç̣¹ni, ç̣\pu¹læ^oja,/

[ç̣¹n̩ ç̣\bu¹læ^oja]

{you not come *MOD}

‘So you are not coming?’

Shì a.

/ʃɥ^oa./

[ʃɥ^oɥa:]

{to-be-so *MOD}

‘Right.’

Zánmen qù ba?

/ç̣¹tsam^omɤn \tshy^opa,/

[ç̣¹tsam^omɤn \tçh̥y^obə]

{we go *MOD}

‘Shall we go?’

Qù ba.

/\tshy^opa./

[\tçh̥y^obə]

{go *MOD}

‘Let’s go.’

In *nǐ bù lái ya?* (which could be rendered also as /ç̣¹ni^opu¹læ^oja,/ [ç̣¹n̩\bu¹læ^oja]), the asker already has a rough idea about the answer, and is ‘double-checking’ it.

Zánmen qiù ba? –instead– is a question and, at the same time, a kind invitation; that is why *ba* is perfectly logical when replying too, since the answerer is not giving an order, but accepting the invitation with a mild exhortation. Surely, the sole interrogative intoneme may suffice to ask a syntactically unmarked question, especially in very colloquial speech (as we can do in English):

| | |
|--------------------|-------------|
| <i>Zánmen qiù?</i> | <i>Qiù.</i> |
| /ç'ʈsám·mɤn ʈʂhy?/ | /ʈʂhy./ |
| [ç'ʈsám·mɤn ʈʂhý:] | [ʈʂhý:] |
| {we go} | {go} |
| 'Shall we go?' | 'Let's go.' |

One should be careful not to abuse such possibility, given the intrinsic non-neutral acceptation attached to it: in fact, *zánmen qiù?* could imply 'I did not expect we were going', or 'did I misunderstand you?' (or 'are we really going?'), just to mention two interpretations; many more are possible, thanks to the multi-faceted contribution of paraphonics.

5.16. As previously said, Chinese *partial questions* often feature a different syntactic arrangement, compared to English ones. But, as in any other language, they require a selective answer to a specific topic, signaled by an appropriate interrogative word: *who?* – *what?* – *where?* – *when?* – *why?* – *how?* – *how many?* – *how much?* and so on. Clearly, these pronouns/adverbs shall receive a noticeable stress (possibly reinforced by emphasis, whenever the actual communicative purpose requires it). The 'hybrid' nature of partial question is mirrored in their intonational patterns: as in English, we have an interrogative *preintoneme*, but not a plain interrogative *intoneme*. Moreover, we find the same intonemic gradation from an informal /ç . / [ç :] (somewhat 'cold', potentially inappropriate with strangers), through /ç . ° / [ç : °] to a milder /ç , / [ç ·], and up to /ç ? ° / [ç : °] (where ° indicates the attenuation of a preceding intoneme). In the following examples, only the first one will be used:

| |
|--|
| <i>Nǐ shénme shíhòu qiù Luómǎ?</i> |
| /ç.ni 'ʂɤn·mɤ'ʂw·ʈsɔ, ç'ʈʂhy 'lwo,ma./ |
| [ç.nʈʂɤm·mɤ 'ʂw·hɔ· ç'ʈʂhý'lwo,ma:] |
| {you which? time go Rome} |
| 'When are you going to Rome?' |

Nǐ gēn shéi qù Luómǎ?

/ç̣.ni ˦kʰn ˦ʂei, ç̣\ʈshy ˦lwσ,ma./

[ç̣.n̩kʰnʂei \ʈçh̥y˦ ˦lwσ,ma˥˥]

{you follow who? go Rome}

‘Whom are you going to Rome with?’

5.17. The same is true for *exclusive* (or *alternative*) *questions*, which require a global reply, but focus their actual ‘interrogativeness’ on one portion only, the ‘x-not-x’ structure; from the intonational point of view, then, exclusive questions act pretty much as partial questions do. Therefore, an interrogative intoneme is not needed, sufficing an interrogative preintoneme. As usual, the intoneme can vary from /./ to /?◦/. Be noted that the negative adverbs *bù* and *méi* are often reduced to *bu* and *mei*:

Nǐ qù bu qù Luómǎ?

/ç̣.ni\ʈshy◦pu \ʈshy, ç̣˦lwσ,ma./

[ç̣.n̩\ʈçh̥y˦.bu \ʈçh̥y˦ ç̣˦lwσ,ma˥˥.] (& [·bu], ‘attracted’ by the following fourth toneme)

{you go not go Rome}

‘Are you going to Rome (or not)?’

Nǐ zuò mei zuò zuòyè?

/ç̣.ni\ʈswσ◦mei \ʈswσ, ç̣\ʈswσ\jɛ./

[ç̣.n̩\ʈʂwσ˥.mei \ʈʂwσ˥ ç̣˦ʈʂwσ\jɛ˥.] (& [·mei])

{you do not do homework}

‘Have you done your homework (or not)?’

5.18. In addition, there is also a kind of *selective questions*, made up by two (or more) clauses connected by *háishì*/*háishǐ* ‘or...?’. In fact, we are not dealing with ‘individual questions’ proper, but with complex sentences whose interrogative nature derives from the opposition between different choices. Depending on the actual meaning, there can be ‘total’ and ‘partial’ selective questions; respectively: ‘x, or not-x?’, and ‘x, or y?’.

Either way, the first part usually starts with an *interrogative preintoneme*, and ends with a *continuative intoneme*, /ç / [ç ·], while the second part ends with a *conclusive intoneme*, /./ [·]. To express a stronger participation, we can have /ç ;/ [ç ·], and even /ç ?/ [ç ·] –thus, separating more clearly the two parts– regularly followed by /./ [·].

With more than two alternatives ('x, or y, or z, or...?'), the same scheme repeats before any *háishǐ*, which itself normally bears a continuative intoneme, /ʰkʌeŋɕu,/. Depending on the speaker's attitude, one alternative can be clearly underlined by ending with /;/ [·], while other alternatives (but the last one) keep /./ [·]. Examples:

Nǐ qù háishǐ bú qù Luómǎ?

/ç.ni|tshy, ʰkʌeŋɕu, ʰpu|tshy,; ʰlwɔ,ma./
 [ç.n|tçhý· ʰkʌɛŋɕu· ʰpu|tçhý· ʰlwɔ,mʌɑ:]
 {you go or not go Rome}
 'Are you going to Rome (or not)?'

Nǐ qù Bālí háishǐ qù Luómǎ?

/ç.ni|tshy ˌpa'li,| ʰkʌeŋɕu, |tshy ʰlwɔ,ma./
 [ç.n|tçhý· ˌbɑ'li· ʰkʌɛŋɕu· |tçhý ʰlwɔ,mʌɑ:]
 {you go Paris or go Rome}
 'Are you going to Paris, or Rome?'

Nǐ qù Bālí háishǐ qù Luómǎ, háishǐ qù Běijīng?

/ç.ni|tshy ˌpa'li,| ʰkʌeŋɕu, |tshy ʰlwɔ,ma;| ʰkʌeŋɕu, |tshy ˌpɛiˊtsin./
 [ç.n|tçhý· ˌbɑ'li· ʰkʌɛŋɕu· |tçhý ʰlwɔ,mʌɑ:] ʰkʌɛŋɕu· |tçhý ˌbɛiˊtɕɪŋ·
 {you go Paris or go Rome or go Peking}
 'Are you going to Paris, or Rome, or Peking?'

5.19. For a full treatment of Chinese intonemes, one final remark has to be made about the *continuative intoneme*, /./ [·], which has two apparently opposite functions: (1) to draw less attention than a *suspensive intoneme* does; and (2) to highlight a word, a concept or an entire clause, even without employing any emphasis.

Should we decide to renounce any continuative intoneme would result in much longer strings, with most words bound together in almost endless preintonemes:

apart from being unpractical, auditorily, it would produce a flat and ‘litany-like’ effect, animated only by marked intonemes. An even more unsatisfactory result would be produced if we decided to use /./ or /:/, wherever /./ is normally encountered. Not that this is wrong *in se*, but it would sound quite unnatural and heavy, unless the speaker intentionally implies a stronger focus on every and each segment of the utterance. Let us examine the following example:

Ānnà zǎoshang qǐle chuáng xǐ-zǎo, chī diǎnr zǎofàn zài chū-ménr.

/ʔan\na,| ,tsa^oʃan _tshio^lx^ʻtshwan ,si,tsa^o.| ˉtshu^ltjaaʒ ,tsa^o\fan, \tsaE⁻tshu^ʻmʃʒ./

[ˈ(ʔ)an\na| _tʃʌʌʃan _tʃhɿ^lx^ʻtʃhwan ˈci,tʃʌʌʃʌ| ˉtʃhu^ldjʌʒ ˌtʃʌʌʃanˈtʃʌʌʃ ˈtʃhu^ʻmʃʒ.]

{Anne morning rise*MOD bed have-shower eat a-bit-of breakfast then exit-door}

‘As she gets up in the morning, Anne has a shower, eats a small breakfast and then gets out’.

Surely, /ʔan\na,| ,tsa^oʃan _tshio^lx^ʻtshwan ,si,tsa^o.| ˉtshu^ltjaaʒ ,tsa^o\fan, \tsaE⁻tshu^ʻmʃʒ./ –with /./ both after *Ānnà* and *xǐ-zǎo*– would be correct, as well; but /ʔan\na,| ,tsa^oʃan _tshio^lx^ʻtshwan ,si,tsa^o.| ˉtshu^ltjaaʒ ,tsa^o\fan. \tsaE⁻tshu^ʻmʃʒ./ would not be so (not to mention /ʔan\na;| ,tsa^oʃan _tshio^lx^ʻtshwan ,si,tsa^o;| ˉtshu^ltjaaʒ ,tsa^o\fan; \tsaE⁻tshu^ʻmʃʒ./).

5.20 Hereby we present a number of examples featuring various intonemes and preintonemes, depending on actual communicative purposes. We recommend our readers to examine these examples with all due attention, and to methodically compare the phonemic notation with the phonetic one.

Nà gūniang, jiùsuàn bù héshàn wǒ yě huì ài tā.

/\na⁻ku^onjan,| ˉtʃʃu^o\swan^opu^ʻx^ʻ\ʃan;|| ,w^o.jɛ\ɰwei, ʔaE⁻tha./

[\na⁻ku^o-njan| ˈtʃʃu^o\swan^opu^ʻx^ʻ\ʃan;|| ˈw^o.jɛ\ɰweiː \tʃʌʌ.thaː]

(as usual, with the possible variants: [ta, ˌtha, ˉtha])

{that girl even-if not good I also can love she}

‘Even if that girl were not good, I would love her nonetheless’.

Wǒ zhēn xīwàng néng xué-hǎo hànyǔ.

/,w^oˉtʃʃɛn ˉsi\wan, ˈnɛn^ʻsɰe,ɰa^o \ɰan,ɰy./

[,w^oˉtʃʃɛn ˈci\wɛnː ˈnɛn^ʻɰe,ɰa^o \ɰanːɰyː]

{I true hope can learn-good Chinese}

‘I truly hope I can learn Chinese well’

Wǒmen shì lǐjiě nín de kànfǎ | kànfǎ.

/ʋwσ◦mɤn |sɥ ʎi,tsjɛ: ˈninˈtɤ |khan,fa./, /|khan◦fa./
 [ʋwσˈmɤn |sɥʎi,tsjɛ: ˈninˈtɤ |kʰɤŋ,fa˥˥:], [|kʰɤŋ,va˥˥:]
 {we be-so understand you-Sir-^{*}DE point-of-view}
 ‘We surely understand your point of view, Sir’.

Tóngshìmen dōu fēicháng gǎnxiè nín.

/ˈthσm|sɥ,; ˌtσu ˌfɛiˈtʃhɤn, ʎan|sjɛˈnin./
 [ˈthσm|sɥ,mɤn: ˌtσu ˌfɛiˈtʃhɤn ˌŋɤŋ|tʃjɛˈnin:]
 {colleagues with-no-exception extremely thank you-Sir}
 ‘All colleagues thank you very much, Sir’.

By underlining *fēicháng* and *nín*, for example, the entire tonal structure should be reorganized around the syllables with stronger prominence:

/ˈthσm|sɥ,; ˌtσu; ˌfɛiˈtʃhɤn ʎan|sjɛ, ˈnin./
 [ˈthσm|sɥ,mɤn: ˌtσu; ˌfɛiˈtʃhɤn ˌŋɤŋ|tʃjɛˈnin:]
 ‘All colleagues thank *you* very much, Sir’.

Lián nǐ de qīzī yě bú huì zuò fàn.

/ˈljɛn,ni◦tɤ ˌtʃhi◦tsɥ; ʎjɛˈpu|ɤwei, |tswσ|fan./
 [ˈljɛn_p|ɔdɤ ˌtʃhiˈdzu; ˌjɛˈbu|ɤwɛi˥˥ ˌdʒwσ|fɤŋ˥˥:]
 {even your wife also not can cook}
 ‘Even your wife cannot cook’.

Mǎkě jǐ-yuè jiù huí-guó?

/ç̣˥ma,khɤ:; ç̣˥tsi|tɥ◦tsjσu ˈɤwei˥˥kwσ./ (less spontaneously: /|tsjσu/)
 [ç̣˥ma,khɤ˥˥(ɤ): ç̣˥dʒ|tɥ˥dʒjσu ˈɤwɛi˥˥kwσ.] (& [ˈdʒjσu])
 {Mark how-many? month then return-country}
 ‘In which month is Marco going back to his country?’

Wèishénme bú dào Běijīng zhǎo gōngzuò?

/ç̣˥wɛi˥˥ɤn◦mɤ ˈpu|taσ ʎpei˥˥tsin, ç̣˥tʃaσ˥kσm|tswσ./
 [ç̣˥wɛi˥˥ɤm˥˥mɤ ˈpu|tqσ ˌbɛi˥˥tɥn ˌç̣˥tʃaσ˥ŋσm|tʃwσ˥˥:]
 {why? not go Peking look-for job}
 ‘Why not looking for a job in Peking?’

Jiǎrú wǒ bú gào su nǐ, nǐ jiù zěnme le?

/ʃtsjaʹzu ʃwσ'pu|kaσosu ʃni;: ʃni|tsjσu, ʃtsɣnσmxσlxʃ/ (better: /ʃwσσpu/)

[.dʒjaʹzu ʃwσ'bu|kqσ.zu ʃn|]: ʃn|dʒjσu ʃtsɣɣm·mx'lx.]

{if I not tell you you then how? *MOD}

'If I don't tell you, what would you do?' (ie 'what would you *dare* to do?')

Jiǎrú nǐ bú gào su wǒ, yě méi shénme.

/ʃtsjaʹzu ʃni'pu|kaσosu ʃwσ;: ʃjE'mei'ɣɣnσmx./

[.dʒjaʹzu ʃn|bu|kqσ.zu ʃwσσ;: ʃjE'mei'ɣɣm·mx.]

{if you not tell I also not something}

'If you do not tell me, no big deal'.

Guóhuì bù bāng wǒ tōngguò zhèxiē fǎlǜ, jǐngchá jiù yào dài bǔ wǒ!

/'kwσ|xlwei, |pu'pan ʃwσ ʃthσn|kwσ |tɣɣ'sje ʃfa|ly;| ʃsin'tɣha |tsjσu|jaσ, |tai,pu ʃwσ./

[·g'wσ|xlwɣ; |b'u'pan..wσ ʃthσn|kwσ |dʒɣ'cje ʃfa|ly'| |cjɣm'pɪn ʃdʒjσu|jaσ |tɣɣ'bu,wσσɣ.]

(or [|tɣɣ.zjɣ ʃfa|ly|.], & [|dʒɣ'pu.wσ.].)

{Parliament not help me pass these law police then arrest me}

'If Parliament does not pass these laws for me, the police will arrest me!'

Guóhuì bù bāng wǒ tōngguò zhèxiē fǎlǜ, xiànbīng yě méiyǒu zhèngjù dài bǔ wǒ.

/'kwσ|xlwei, |pu'pan ʃwσ ʃthσn|kwσ |tɣɣ'sje ʃfa|ly;| ʃsin'tɣha ʃjE'mei'jσu |tɣɣn|tɣ;,
|tai,pu ʃwσ./

[·g'wσ|xlwɣ; |b'u'pan..wσ ʃthσn|kwσ |dʒɣ'cje ʃfa|ly'| |cjɣm'pɪn ʃjE'mei'jσu ʃtɣɣn|tɣ;.

|tɣɣ'bu,wσσɣ.] (or [|tɣɣ.zjɣ ʃfa|ly|.], & [|dʒɣ'pu.wσ.].)

{Parliament not help me pass these law police also not have proof arrest me}

'Even if Parliament does not pass these laws for me, the police has no proof to arrest me, anyway!'

5.21. Let us compare also the following occurrences of the same sentence, with emphasis on different words. Double-line symbols, [=, ʹ, ʹ, ʹ, =, ||, ʹ], very intuitively represent the reinforced intensity of emphatic taxotones, with respect to 'normal' [=, ʹ, ʹ, ʹ, =, |, ʹ]). We chose to employ *Canepari* in these sentences, especially because it is a four-syllable surname, rendered as /,kha'ni'pha|li/.

We omit the phonemic transcription, in this case, as it is of very little use to highlight emphatic phenomena; besides, it is a good opportunity for our readers to check their ability to derive a phonemic transcription from a phonetic one:

Kǎnǐpālì jiàoshòu zài Wēinísī dàxué jiǎng yǎnyīnxié.

{Canepari professor to-be-in Venice university explain phonetics}

[kʰǎnǐpha|l| ˈdʒjaɔ|sɔ̌y| ˈdʒǎʃ˦wəǐnǐsu ˈdʒǎʃɥɛ| ˈtɕjaň ˌɥ̌jɪn˨˨ʃɥɛ.]

‘It is professor *Canepari* that taught *phonetics* (not another subject) at Venice University’.

[kʰǎnǐpha|l| ˈtɕjaɔ|sɔ̌y| ˈdʒǎʃ˦wəǐnǐsu ˈdʒǎʃɥɛ| ˈtɕjaň ˌɥ̌jɪn˨˨ʃɥɛ.]

‘*Professor Canepari* (not a generic *Mr Canepari*) taught *phonetics* (not another subject) at Venice University’.

[kʰǎnǐpha|l| ˈdʒjaɔ|sɔ̌y| ˈdʒǎʃ˦wəǐnǐsuː ˈdʒǎʃɥɛ| ˈtɕjaň ˌɥ̌jɪn˨˨ʃɥɛ.]

‘Professor *Canepari* taught phonetics at *Venice University*’.

[kʰǎnǐpha|l| ˈdʒjaɔ|sɔ̌y| ˈtɕǎʃ˦wəǐnǐsu ˈdʒǎʃɥɛ| ˈtɕjaň ˌɥ̌jɪn˨˨ʃɥɛ.]

‘Professor *Canepari* taught phonetics *at Venice University*’.

[kʰǎnǐpha|l| ˈdʒjaɔ|sɔ̌y| ˈdʒǎʃ˦wəǐnǐsu ˈtɕǎʃɥɛ| ˈtɕjaň ˌɥ̌jɪn˨˨ʃɥɛ.]

‘Professor *Canepari* taught phonetics at *Venice University*’.

[kʰǎnǐpha|l| ˈdʒjaɔ|sɔ̌y| ˈdʒǎʃ˦wəǐnǐsu ˈdʒǎʃɥɛ| ˈtɕjǎʃ˦ː ˌɥ̌jɪn˨˨ʃɥɛ.]

‘Professor *Canepari* taught phonetics at *Venice University*’.

[kʰǎnǐpha|l| ˈdʒjaɔ|sɔ̌y| ˈdʒǎʃ˦wəǐnǐsu ˈdʒǎʃɥɛ| ˈtɕjaň ˌɥ̌jɪn˨˨ʃɥɛ.]

‘Professor *Canepari* taught *phonetics* at *Venice University*’.

Parentheses & quotations

5.22. Lastly, fig 5.9 shows the diagram of *parenthetic phrases*, or simply *parentheses* (either *low*, [l ˩]; or *mid*, [˨ ˨]), and of *quotations*, [˨ ˨]. In the following sentence, each of them occurs once:

First of all –he said– ‘let’s consider “Natural” Phonetics, as it’s properly called’.

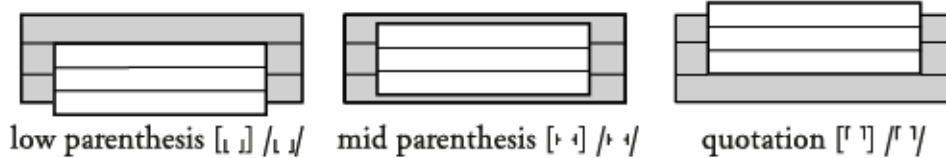
[fɛ˨˥˥ʂf əv˨˥˥tʰi˨˥˥sɛ˨˥˥dʰi˨˥˥ ˨˥˥ts kʷɪn˨˥˥sɔ̃˨˥˥ ˨˥˥nætʃjɛ˨˥˥] fə˨˥˥nɛf[kɛs˨˥˥ː ɪəz[kɛ˨˥˥pʰɪpəli ˨˥˥kʰo˨˥˥tɕ˨˥˥].

Parentheses typically feature an overall reduction of their stress and an increase in the rate of speech, while the pitch is compressed in the low –or mid– range of the tonogram. *Quotations*, instead, are quite the opposite, as their role is –precisely– putting one or more words in full evidence, by means of a slightly louder and distinct enunciation: thus, their stress is increased, the speech rate is reduced, and the pitch is raised (*without* compression).

Marking such prosodic subtleties in phonetic –let alone *phonemic*– detailed transcriptions is neither necessary nor recommendable. The symbols [l ɿ], [t ʰ], [ʰ ˊ] are more than sufficient to bear in mind all these differences, with respect to ‘normal’ utterances. Then, even with regard to stress (and taxotones, as far as tonal language are concerned) there is no need to deviate from the usual notation.

Quotations must not be confused with *direct speech*. Let us go back to *First of all –he said– ‘let’s consider “Natural” Phonetics, as it’s properly called’*. In that sentence, only *he said* should be excluded, because all the rest –and what may follow– is direct speech, indeed.

fig 5.9 Parentheses and quotations.



Chinese *parentheses* are low, [l ɿ], after a conclusive intoneme; and mid, [t ʰ], after other intonemes; quotations, as said, are high, [ʰ ˊ]. Such intonational superstructures are quite useful in tonal language, thanks to small variations in pitch (small, yet perceptible by native speakers, but not so much –at the very beginning– by foreign learners, who have already a hard time struggling with tones and taxotones in ‘normal’ utterances). A few examples:

Wǒ ài nǐ, qīn'àide. Nǐ yě ài wǒ; name, zánliǎ hái děng shénme?

[wɔ̃˨˥˥ʔaɛ˨˥˥ni˨˥˥. ɿ˨˥˥tʃɪn[ʔaɛ˨˥˥tʃɪn˨˥˥] ɿ˨˥˥ni˨˥˥jɛ˨˥˥ ʔaɛ˨˥˥wɔ̃˨˥˥. | tʰɿ˨˥˥nɑ˨˥˥mɛ˨˥˥, | tʰɿ˨˥˥tsan˨˥˥lja˨˥˥, tʰɿ˨˥˥ʔaɛ˨˥˥tʃɪn˨˥˥ ˨˥˥ʃɛ˨˥˥m˨˥˥mɛ˨˥˥, /
[.wɔ̃˨˥˥ʔaɛ˨˥˥nɿ˨˥˥] | ɿ˨˥˥tʃɪn[ʔaɛ˨˥˥dʃɪn˨˥˥] | ˨˥˥ni˨˥˥jɛ˨˥˥ ʔaɛ˨˥˥wɔ̃˨˥˥. | tʰɿ˨˥˥nɑ˨˥˥mɛ˨˥˥ | tʰɿ˨˥˥dʒal˨˥˥lja˨˥˥. tʰɿ˨˥˥ʔaɛ˨˥˥tʃɪn˨˥˥ ˨˥˥ʃɛ˨˥˥m˨˥˥mɛ˨˥˥.]

{I love you beloved you also love I then the-two-of-us yet wait what?}

‘I love you, my dear, and you love me. So, what are we waiting for?’

Jìrán zánliǎ xiāng'ài, qīn'àide, wǒ xiǎng gēn nǐ jié-hūn.

[tsi'zan 'tsan.lja 'sjan|pae; 'tshin|pae'otx, | 'wɔ'sjan 'kɣn.ni 'tsje'ɣwɣn./

[dʒɔ'zan 'tsal.lja 'cjan|pae'otx; 'tchin|pae'dɣ'otx | 'wɔ'cjan 'gɣn.ni 'dʒje'ɣwɣn.]

{given-that the two-of-us love-each-other beloved I want be-with you marry}
 'Since we love each other, my dear, I want to marry you'.

Tā sī-kai shàngbian xiězhe «Mǎkě shuàigē shōu» de xìnfēng yī kàn, xīnli jiù xiězhe:
«Wǒ yǐjīng jié-hūn le, huitóu zài shuō».

[ˈthaˈsuːkhae, | ʂanˈpjen ˌsjeˈotx, | ˈma.khɣ ʂwaɛˈkɣ, ˌʂuːotx, | ʂinˈfɣn ˌjiˈkhan; |

ʂinˈli ʂjɔˌsjeˈotx, | ˈwɔˌjiˈotxin 'tsje'ɣwɣn.lɣ. || ˈɣweiˈtʂou ʂsaeˈʂwɔˈ.]

[ˈthaˈsuːkaɛ | ʂan.bjɣn ˌcjeˈdɣ'otx | ˈma.khɣ ʂwaɛˈkɣ ˌʂuːdɣ'otx | ʂiŋˈfɣn ˌjiˈkhaŋ | ʂiˌl |

dʒɔˌcjeˈdɣ'otx | ˈwɔˌjɔˌdɣ'otx | ˌɣweiˈtʂou ʂsaeˈʂwɔˈ.]

{he rip upper-side being-written Mark handsome-guy receive envelope letter-inside
 being-written I already marry *MOD next-time talk-again}

'He opened the envelope headed to handsome Mark, and the letter said I am married. Maybe next time'.

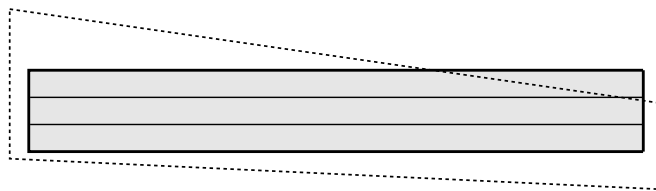
5.23. We think it is useful to present a comparison between *natural perception* and the *acoustic measurements* of intonation: fig 5.10 shows a horizontal tonogram (in Chinese, it could be a / / [] [°]), or a /λ/ [λ] [°]), and –superimposed in dotted line– the tonal range which would be produced by such acoustic softwares as *Praat*.

The natural approach, as previously said, implies a useful normalization of numerous occurrences of the same phenomenon; the human ear is a selective organ, and only a small portion of the actual sound waves of speech (the one that ‘matters’) is eventually sent to the brain, which –in turn– discards the incidental and the non-typical. This is done, in part, by the language-oriented nature of human beings; in part, by the fact of their having pronounced tens of thousands of utterances, and listened to even much more of them.

However, no matter whether we are native or (fluent) non-native speakers, almost no-one can be confidently sure of being capable of actually describing intonation units, by simply looking at one of the particular tracings produced by acoustic softwares, if they are not accurately normalized. These tracing are of little practical use, because, paradoxically, they are too *finicking*. A machine will present a melodic curve *as it records it*, plain and simple, with everything put on the same level, in-

cluding what the human brain would downgrade at the paraphonic level, or even ignore altogether, for being nothing more than an automatic adjustment. In fact, in fig 5.10, the much broader range at the beginning, and the gradual frequency reduction towards the end, are absolutely normal: that is why our brain perceives a level intonation, instead of a steeply falling one.

fig 5.10. Natural (tonogram) and acoustic (dotted line) depictions of intonation.



NOTES:

¹ The complex mechanism of intonation is described in depth in CANEPARI, L. (2007) *Natural Phonetics & Tonetics*, § 13 et seq.

² None of the books that we have examined so far offers a really convincing treatment of Chinese intonation, and that explains why the present chapter, though very complex and detailed, practically features no notes. Interestingly, it is an old-fashioned handbook, such as the 1981 edition of the *Practical Chinese Reader*, that has captured our attention: its pragmatic approach to intonation is reminiscent of the old British school scheme, where /;/ and /?/ are often merged, and paraphonics is currently mixed up with linguistic intonation; however, the *Reader* has the merit of getting straight to the point, by coupling many sample sentences – basically, but clearly, annotated – with a number of relevant audio files.

Surely, that is not the most scientific way of analyzing pronunciation, but –at least– it does not muddle up the readers with a bunch of mysterious symbolic conventions, as those adopted in CAO, W. (2002) *Hanyu yuyin jiacheng*, p. 140-148 (which, frankly, seems quite random and empirical, whereas a good annotation system should always be founded on scientific and consistent elements, like ours *preintonemes* and *intonemes*).

³ Among the *very few reliable* works on the matter, we would suggest KINGDON, R. (1958) *The Groundwork of English Intonation*, although not completely free from paraphonic mixes.

⁴ We will note –*en passant*– the existence of a particular kind of pronunciation handbooks, those compiled for the Chinese themselves who aim at a post where ‘standard pronunciation’ is required: we have found one *Henan sheng putonghua shuiping ceshi zhidaoshu* ‘A Guide to the Henan Province Common Language Proficiency Test’, where intonation is shown, quite erratically, with a very similar system of dashes and arrows.

⁵ For example, Swedish and Vietnamese feature /?/ [.:].

6. Phono-tonetic analysis of extended texts

The North Wind and the Sun

6.1. In this chapter, what we have explained so far will be summarized and put in practice, by accurately transcribing two extended texts. Customarily, many linguists choose the Aesopian tale *The North Wind and the Sun*, as a sort of ‘universal specimen’ for phonetic analysis. Actually, a tale is not always a good choice, since it usually contains a lot of descriptive passages, accompanied by a moderate amount of *direct speech*. The latter, in our view (and according to the natural approach itself), is *the* most natural and representative form of ‘spoken language’; in a tale, however, direct speech is often reduced to very short sentences, offering little room for complex intonational patterns and paraphonic features. Hence the need for descriptive passages, where longer utterances are more common, provided they are... *said* by the speaker, not mechanically read aloud, like a press release.

In fact, reading aloud is nothing but the *phonic rendition of a written text*. Good writers –journalists above all– aim at being as convincing and understandable as possible, while (hopefully!) striving for conciseness, in order to keep their readers from falling asleep. Natural speech, instead, obeys to different dynamics, paying much more attention to rhythm and ‘flow’, rather than simply delivering information. When these two different ways of using language are forced to coexist, results are rarely satisfactory.

In reading aloud –for example, from a newspaper, or a textbook– everyone must have experienced that uncomfortable feeling of ‘something missing’, as if the text lacked strength and balance. Typically, passionate and eloquent speakers have a hard time coordinating the pauses perceived by them as spontaneous and necessary, with those found in the written page. The main reason is that traditional punctuation works –almost exclusively– as a typographical aid to highlight the *syntax* of complex utterances, and only secondarily to mark expressive pauses and emotional features.

As such, punctuation evidently fails at reproducing the colourful richness of real speech, with all its changes in pitch, speed, and paraphonic nuances. That is why ‘verbatim’ transcriptions –like parliamentary and judiciary records– look so redun-

dant in some parts, and desperately elliptical in others. Omitting altogether what readers could not reconstruct by themselves is exactly the price to pay, in order to make written language understandable.

6.1.1. Our version of *The North Wind and the Sun* is a decent compromise, in this regard; yet it remains a step behind what would be judged as an ‘ideal target’ for our ^{can}IPA phono-tonetic notation. Besides, we added two total questions in *direct speech*, at the end of the tale proper, as it would have been otherwise impossible to introduce a sample of the relevant /¿ ?/ intonation.

The Mandarin text is presented, first, in *pīnyīn* romanization: punctuation and hyphenation are somewhat ‘richer’ than what is found in common usage (as *pīnyīn* is usually subordinated to traditional Chinese script, which is very parsimonious with punctuation marks, and –of course– uses no hyphens between Chinese characters). However, such ‘enrichments’ will certainly help our readers parse longer sentences with ease, and will offer a useful reference to underline similarities and differences between grapho-syntactical punctuation and a true intonemic notation.

We provide quite a literal English translation, *not* the English version of the tale. However, those who know Mandarin should better avoid looking at any translation at all, since a residual cross-language interference is inevitable anyway. In order to render the communicative ‘logic’ of the original in a genuine way, any addition is put in square brackets. Verbal persons and tenses are not explicitly marked in Chinese, which rather focuses on distinguishing verbal aspect. We assume the traveller is a male, in order to get rid of such masterpieces as *he or she, him or her, his or her...* The Wind and the Sun, though ‘personified’ here, will be referred to as *it*.

6.1.2. *Yǒu yī-huì, Běifēng gēn Tàiyáng zhèng zài nàr zhēnglùn shéide běnshì dà. Shuō-zhe shuō-zhe, lái le yī-ge zǒudàorde, shēnshang chuānzhe yī-jiàn hòu páozǐ. Tāmen liǎ jiù shāngliang-hǎo le shuō: «Shéi néng xiān jiào zhè-ge zǒudàorde bǎ tāde páozǐ tuō-xiàlai ya... jiù suàn shéide běnshì dà».*

Hǎo... Běifēng jiù shǐ-qi dà jìn r lái jǐnguā jǐnguā; kěshì tā guā de yuè lǐhai, nà-ge rén bǎ páozǐ guǒ de yuè jǐn. Dào mǒliǎor, Běifēng méile fázi, zhǐhǎo jiù suàn le. Yīhuìr Tàiyáng jiù chūlái rěrěrde yī shài, nèi zǒudàorde mǎshàng jiù bǎ páozǐ tuō-le-xiàlai. Suǒyǐ, Běifēng bù néng bù chéngrèn, dào dǐ háishì Tàiyáng bǐ tā běnshì dà.

Nǐ xǐhuān zhè-ge gùshì ma? Wǒmen zài jiǎng yī-biàn ma?

That is:

Once upon a time, [the] North Wind and [the] Sun were there disput[ing] whose skill was great[er]. [While still] talking [and] talking [again], a traveller came [along], wearing a thick cloak on [his] body. The two, having discussed, said then: ‘[The one] who can first make this traveller take his cloak off... let us consider its skill [as] great[er]’

OK... [The] North Wind, then, began us[ing] great energy [and] repeatedly blew hard; however, the more violently it blew, the more closely [did] that man fold [his] cloak [around him]. Finally, [the] North Wind had no way out, and was forced to give up. In a while, [the] Sun came out [and, as soon as it] shone scorchingly, that traveller took off his cloak immediately. Consequently, [the] North Wind could not help admitting [that], in the end, [the] Sun had [a] great[er] skill than it did.

[Do] you like this tale? [Shall] we tell [it] once again?

A more natural translation:

Once upon a time, the North Wind and the Sun were disputing over which was the stronger, when a traveller came along, wearing a thick cloak. The two made deal and said: ‘Who first succeeds in making this traveller take off his cloak... shall be declared the stronger’.

OK... The North Wind, then, began blowing with full power; however, the more violently it blew, the more closely did that man fold his cloak around him. Finally, the North Wind was left with no other choice but giving up. In a while, the Sun shone out warmly, making the traveller took off his cloak immediately. Consequently, the North Wind could not help admitting that, in the end, the Sun was the stronger of the two.

Did you like this tale? Shall we repeat it?

6.1.3. And now, let us examine the *photonemic* transcription: this does not represent the actual pronunciation, but it is a reasonable approximation to it. As a matter of fact, native speakers and L2 speakers, who already know the language, should be able to read our phonotonic symbols and mentally ‘translate’ them into their phonetic equivalents, pretty much as they do when reading aloud a passage in *pīnyīn* (which, more or less, can be viewed as a non-IPA phonemic transcription *in se*). With respect to *pīnyīn*, however, our phonemic transcription provides some extra –but, decidedly, not at all secondary– information.

First, it clearly indicates any necessary tonal variations, like that of *yī* ‘one’: *yī-huí* /jɿ˥˥wɛi/, *yī-ge* /ʃji˥kɤ/, *yī-jiàn* /ʃji˥tsjɛn/, *yī-biàn* /ʃji˥pjɛn/. Secondly, it avoids ambiguities in rendering *érhuà*, as in *yīhuìr* /jɿ˥˥wɤʒ/ (with *two* tonal changes, too, typical of idiomatic use; while, in a more formal delivery, /ʃji˥˥wɤʒ/ would be preferred). Finally, and more importantly, our readers will find a secure guide to the pronunciation of Mandarin with the right intonation: preintonemes, intonemes, parentheses and quotations (/ ɿ /, / ʃ /), pauses (of different duration: / | /, / || /, / ||| /), and *paraphonics*, as well.

Regarding this last point, we have decided to be quite ‘sober’. We could not omit the *raised larynx*, ⟨.:⟩, and the *backing of the body of the tongue*, ⟨+⟩, so typical of a genuine ‘Mandarin voice’ (but not necessarily of all Chinese languages); and we put them together, ⟨.: +⟩, *before* the actual transcription, in order to mean that they are referred to the whole text (or, in this case, to *any* text in Mandarin Chinese). Then, we simply added ⟨⟩, indicating that the following unit (here, a monosyllable: /ɿ˥˥wɤʒ/) is uttered *more slowly* than usual, to increase the suspense; and ⟨»⟩, implying that the entire sequence /ɿ˥˥wɤʒ˥˥mɤn˥˥tsɛ˥˥, ɿ˥˥tsjɛn˥˥ʃji˥˥pjɛn˥˥mɑ˥˥?/ is uttered *more quickly* instead (of course, the paraphonic symbol must be repeated before any new relevant rhythmic group).

But paraphonics can be much more intricate than that, being able to represent the *lowering* of the habitual pitch extent of a given speaker, ⟨◊⟩; or its *raising*, ⟨°⟩; or an overall *expansion*, ⟨◌◌⟩; or, finally, a *compression*, ⟨◌◌⟩. Of course, these pitch settings can be employed for proper intonational purposes –ie *linguistic* intonation, between []– as well, and that is why mere acoustic measurements are of no scientific relevance, if not combined with the discerning ability that only the human brain possesses.

Moreover, one speaker may feature a particularly nasal voice; or an exaggeratedly acute pitch, commonly said *falsetto*; and many more peculiarities. Some of them are typical of a certain language, and thus generally used; others may be limited to one single speaker, or are intentionally employed in *that particular utterance*, for emotional –if not ‘theatrical’– purposes. Multiple combinations are practically endless. As can be seen, the subject is exceptionally vast and multi-faceted, and we strongly suggest our readers to refer to *Natural Phonetics and Tonetics*, and to *Phonetic Notation*, for a more detailed treatment of paraphonics.

⟨.: +⟩ /ɿ˥˥wɤʒ˥˥wɛi;|| ɿ˥˥pɛi˥˥fɤn, ˥˥kɤn˥˥thɛ˥˥ʃan;: |tsɤn˥˥tsɛ˥˥nɑʒ ˥˥tsɤn˥˥lwɤn.| ʃɛi˥˥tɤ
ɿ˥˥pɤn˥˥ʃw |ta.|| ɿ˥˥ʃwɤʒ˥˥tɤ ˥˥ʃwɤʒ˥˥tɤ;|| ɿ˥˥lɛ˥˥lɤ ʃji˥˥kɤ ɿ˥˥tsɤ|taʒ˥˥tɤ.|| ˥˥ʃɤn˥˥ʃan, ˥˥tɤhwan˥˥tɤʒ

ʹji|tsjɛN \kʰouʹphaoʰtsu. || ʰthaʰmɤN|ja: |tsjɔuʰʂAN|jan ʰkʰaʰoʰlɤ ʰʂwɔ, || ʰʂɛi ʰnɤNʰsjɛN
|tsjaʰo: |ʰʂɤʰkɤ|tsou|taʰʂoʰtɤ, | ʰpaʰthaʰoʰtɤ ʹphaoʰtsu, | ʰthwɔ, |sjaʰlaɛʰja, || |tsjɔu|swan,
ʹʂɛiʰoʰtɤ ʰpɤNʰʂɤw |ta. ||

ʰ ʰkʰaʰoʰ; || ʰpɛiʰfɤN, |tsjɔu|ʂwʰoʰtʂi ||ta, |tsjɤʂʹlaɛ, ʰtsinʰkwa, ʰtsinʰkwa. || ʰkʰɤ|ʂwʰtha; ||
ʰkwaʰoʰtɤ |ʰɛ|liʰkʰaɛ, || |naʰkɤ ʹʂɤN; | ʰpaʰphaoʰtsu, ʰkwoʰtɤ |ʰɛ, |tsin. || |taʰo |mwɔ, |ljaʰʂ; ||
ʰpɛiʰfɤN ʹmɛiʰoʰlɤ ʹfaʰoʰtsu. || ʰʂwʰkʰaʰoʰ, |tsjɔu|swanʰoʰlɤ. || |ji, |kʰwɤʂ; || |thaɛʹjan
|tsjɔuʰʂhuʰlaɛ, || |ʂɤʰʂɤʰoʰtɤ ʰji|ʂaɛ. || |nɛi|tsou|taʰʂoʰtɤ, || ʰma|ʂAN, |tsjɔu|paʰphaoʰtsu,
ʰthwɔʰoʰlɤ |sjaʰlaɛ. || ʰswɔʰji, | ʰpɛiʰfɤN, || |puʰnɤN|pu ʹʂɤʰɤN|ʂɤN; || |taʰoʰti, |
ʹkʰaʰʰoʰʂw|thaɛʹjan, | ʰpiʰtha, ʰpɤNʰʂɤw ||ta. ||

ʰni, |siʰkʰwan |ʰʂɤʰkɤ |kuʰʂwʰoʰmaʰ? || » ʰwɔʰmɤN|tsaɛ, » ʰtsjan ʹji|pʂɛNʰmaʰ? |||/

6.1.4. It is true that, for such languages as Italian and Spanish, the distance between the phonemic theory and the phonetic reality is so small, that a good phonemic transcription is already enough for a convincing –yet non-neutral– pronunciation. Thus, for Italian *dice* /'ditʃe/ ‘(he or she) says’, one could actually say [diʃe], instead of the expected [diʃɛ], without any fear of being misunderstood or harshly criticized. In Chinese, the same distance is more relevant, and a word like *jièjie* /tʂjɛʰtʂjɛ/ ‘younger sister’ should not be articulated exactly as [tʂjɛʰtʂjɛ]; native speakers would certainly understand it, but would not accept it even as a ‘passable’ alternative to the possible neutral realizations: [tʂjɛʰdʒjɛ] (& [dʒjɛ, dʒjɛ], in fast speech, and in a preintoneme).

In other words, if *pīnyīn* can largely suffice to ‘transcribe’ Mandarin phonemically (provided intonation, pauses and paraphonics are adequately registered too, possibly via our own *canIPA* symbols), a more detailed *phonetic* transcription must be provided anyway, in aid to anyone who uses it to learn the actual pronunciation. We recommend to read through the following transcription with utmost caution and attention: this concise phonotonic *summa* contains a very large portion of what has been explained so far, in the former chapters. There are multiple examples of vocalic and consonantal taxophones, and as many taxotones. There is more than one occurrence of *érhuà* and *qīngshēng*. Intonemes and preintonemes are duly represented, in combination –as said– with a small, but meaningful, number of paraphonic traits.

We remind, once again, that the consonantal gradation, /C/ [ʰC, ʰC̥, ʰC̣] and /Ch/ [ʰC̥h, ʰC̣h, ʰC̣h], represents the taxophonic *extremes* reachable in fast, fluid and connected speech, but oscillations are absolutely normal. In particular, for /ʰC, ʰCh/,

[_oC̣, _oČ̣] are more common in preintonemes; while, in an intoneme, /_oC, _oCh/ [_lC̣, _lCh] (or [_oC̣, _oČ̣Ch], and even ‘full’ [_lC̣, _lCh]) are statistically frequent. Thus, for better realism, we will precisely adopt [_lC̣, _lCh] in intonemes, and [_oC̣, _oČ̣] elsewhere.

(∴ ∴) [l-jɔʋ|j|ɕkwəŋ|] ..bɛi-fɛn· gən|thɛʃʃjan: |ʈʂɛn·dʒaɛ|nɐʃʃ |dʒɛn|lwɜŋ| |ʃɛi-dɛ
 ·pʃɛŋʃɛŋ |tʃɑ:| |l-ʃwɔ-dʒɛ |ʃwɔ-dʒɛ·| |lɛɛ·lɛ |ʃi-gɛ..dʒɔʋ|tʃɔŋ.dʃ| |ʃɛŋ-ʒɛn· |tʃhwɛŋ-dʒɛ·
 |ʃi-dʒɛn |ɕɔʋ'pɦɔɔ-dʒɛ·|| |thɛ·mɛl_ljɛ· |dʒɔʋ-ʃɛn·ljɛn..hɔŋ|lɛ |ʃwɔɛ|| |ʃɛi |nɛn·cɛɛn
 |tʃjɔŋ: |tʃɛ.gɛ..dʒɔʋ|tʃɔŋ.dʃ| |bɛ-thɛ-dɛ |'pɦɔɔ-dʒɛ·| |tʃwɔ· |cɛjɛ.lɛʃ.jɛ|| |dʒɔʋ|swɛŋ:
 |ʃɛi-dɛ |pʃɛŋʃɛŋ |tʃɑ:| |

· |ɕkwɔŋ|| ..bɛi-fɛn· |dʒɔʋ-ʃɛŋ|tʃi |tʃɑ: |tʃjɛʃʃ·lɛɛ· |dʒiŋ-kwa· |dʒiŋ-kwa|| |khɛʃɛŋ-thɛ||
 |kwa-dɛ |ʃɛ|ɕkwɛʃ|| |nɛ.gɛ |ʒɛn: |bɛ'pɦɔɔ-dʒɛ· |kwɔ-dɛ |ʃɛ.tʃiŋ|| |tʃɔŋ |mɔŋ.ljɛɛŋ||
 ..bɛi-fɛn |mɛi·lɛ |fa-dʒɛ·|| |dʒɛ·ɕkwɔ· |dʒɔʋ|swɛŋ.lɛ·|| |j|ɕkwɛʃʃ|| |thɛʃʃjan
 |dʒɔʋ-tʃhu·lɛɛ|| |ʈʂ-ʈʂɛ-dɛ |ʃi|ɕkwɛʃ|| |nɛi..dʒɔʋ|tʃɔŋ.dʃ|| |mɛʃɛn· |dʒɔʋ.bɛ'pɦɔɔ-dʒɛ·
 |tʃwɔ·lɛ |cɛjɛ.lɛʃ.jɛ|| |l-ʃwɔ·j| |bɛi-fɛn·|| |bɛ'pɦɔɔ· |tʃhɛn|ʈʂɛŋ|| |l-dʒɔŋ-t|
 |lɛɛ-ʒɛŋ|thɛʃʃjan· |b|thɛ· |pʃɛŋʃɛŋ |tʃɑ:| |

ɛ'ni_c|kwɛŋ |tʃɛ.gɛ |kɛ.ʒɛŋ-mɛ|| » ɛ.wɔ-mɛn|tʃɛʃ » ɛ-tʃjɛŋ |ʃi|pʃɛŋ.mɛ|| |

6.2. We propose another short text, taken from Hé Píng’s *Chinese Pronunciation Course* VOL.2, lesson 1: «*Liúyǒu yúdi*», which can be rendered as ‘to preserve a leeway’. In this case, our translation will be less literal, as some passages would be simply meaningless in English otherwise.

Liúyǒu yúdi.

Yàmǎxiùn Hé rèdài yǔlín zhōng yǒu yī-zhǒng jiào yǐxióng de dòngwù. Zhè-zhǒng yǐxióng shì shìjiè-shàng zuì dà de shìyǐshòu, měi-tiān yào chī 1.6-wàn-zhī mǎyǐ.

Rán'ér, ràng rén dàwéi jīngtàn de shì yǐxióng yǒu yī-zhǒng xíxìng: tā chī mǎyǐ shí juébù zhǎn-jìn-shā-jué, dāng tā páo-kāi měi yī-ge yǒu chéng-qiān-shàng-wàn zhī mǎyǐ de yǐxiúé shí, tā jiù zhǐ chī qízhōng yī xiǎo bùfen, zuìduō wéi 500-zhī, tā duì shèngxià de mǎyǐ bù-wén-bù-wèn, jìngzhǐ qù xúnzhǎo xià yī-ge mùbiāo. Kějiàn, yǐxióng qīngchu de zhīdao, yào shǐ zìjǐ de zhōngqún shēngcún-xiàqu, jiù bìxū yǒu mǎyǐ-jiāzú de shēng-shēng-bù-xī.

Yǐxióng shàngqiě zhīdao liúyǒu yúdi, « jihuà » shǐyòng zīyúán, rénlèi bù gāi sān-sī-ér-hòu-xíng ma?

Always leave yourself a leeway

In the Amazon tropical rain forests, there is a kind of animal, called an *anteater*, which is the biggest myrmecophagid in the world, and eats up to sixteen thousand ants per day.

However, what is most surprising is that the anteater has a peculiar habit: when eating ants, it does not simply ‘wipe them out’; as it digs out an anthill with tens of thousands of ants inside, it eats but a small portion of them, no more than five-hundred, totally ignoring the others, and goes straight looking for the next target. Evidently, the anteater clearly knows that, in order for its own species to survive, it is necessary that the ant family lives on.

If even anteaters know how to leave themselves a leeway, and ‘plan’ their use of resources, should men not think thrice before acting, too?

6.2.1. For *The North Wind and the Sun*, we provided a normalized transcription ‘as it should be’, after collecting a number of recordings from various informants, who, though speaking Mandarin natively, presented different levels of proximity to neutral pronunciation. For *Always leave yourself a leeway*, instead, we referred to a single audio file in a pronunciation course, publicly available on the market; this obviously bound us to adopt a more realistic transcription, where oscillation from the ‘norm’ will emerge quite plainly.

Let us start, as usual, with the phonemic transcription, before presenting the ‘normalized’ phonetic one. It goes without saying that the reader should pay attention to the differences (often very minute) between the romanization and our *canIPA* transcriptions

<.: +> /ʰljou.jou˨˥y˨˥ti.∥
 |ja.ma|syn˨˥x̥ |ʒx̥|tae ˨˥y˨˥lin˨˥t̚sɔŋ,∥ ˨˥jou|ji˨˥t̚sɔŋ, ˨˥|tsjaɔ,| ˨˥ji˨˥sɔŋ˨˥t̚x̥ |tɔŋ|wu.∥∥
 |t̚sɔŋ˨˥t̚sɔŋ˨˥ji˨˥sɔŋ,| |s̚u,ː |s̚u|tsjɛ˨˥s̚aŋ,| |tswei|ta˨˥t̚x̥, ˨˥ˈs̚u˨˥ji˨˥|s̚u.∥∥ ˨˥mei˨˥t̚hjen |jaɔ˨˥t̚shu;∥
 ˨˥ji˨˥t̚jaʒ, |ljou|wan˨˥t̚s̚u ˨˥ma.ji.∥∥
 ˨˥ʒan˨˥ʔx̥ʒ;| |ʒan˨˥ʒx̥n, ∥ta˨˥ˈwei˨˥ t̚sin|than˨˥t̚x̥|s̚u.∥ ˨˥ji˨˥sɔŋ;| ˨˥jou |ji˨˥t̚sɔŋ˨˥si|sin.∥∥ ˨˥tha˨˥t̚shu
 ˨˥ma.ji ˨˥ˈs̚u;∥ ˨˥tsɛ|pu,| ˨˥ʒan|tsin ˨˥s̚a˨˥tsɛ.∥∥ ˨˥tan˨˥tha ˨˥phaɔ˨˥khae ˨˥mei˨˥ji˨˥k̚x̥,| ˨˥jou ˨˥t̚shx̥n-
 ˨˥t̚hjen|s̚aŋ|wan˨˥t̚s̚u ˨˥ma.ji˨˥t̚x̥, ˨˥ji˨˥sɛ˨˥ˈs̚u;∥ ˨˥tha|tsjou˨˥t̚s̚u˨˥t̚shu ˨˥t̚hi˨˥t̚sɔŋ, |ji˨˥sjaɔ-
 |pu˨˥f̚x̥n.∥∥ ˨˥|tswei˨˥t̚wɔ˨˥ː; ˨˥ˈwei˨˥ː ˨˥wu˨˥pai˨˥t̚s̚u.∥∥ ˨˥tha|twei, |s̚x̥n|sja˨˥t̚x̥ ˨˥ma.ji˨˥ː; |pu˨˥w̚x̥n-
 ˨˥pu˨˥w̚x̥n.∥∥ |tsin˨˥t̚s̚u,ː |t̚shy˨˥syn˨˥t̚s̚aɔ, |sjaɔ˨˥ji˨˥k̚x̥ |mu˨˥pjaɔ.∥∥ ˨˥k̚x̥|tsjen.˨˥ ˨˥ji˨˥sɔŋ,

ˊtsin˥˥shu˥˥tɕ ˊtɕw˥˥tao˥˥.|| |jao ˊɕw |tsu˥˥tsi˥˥tɕ ˊtɕson˥˥tshyn, ˊɕxN˥˥tshwɕN\sjao˥˥tshy;||
|tsjɔu\pi˥˥sy, ˊjɔu ˊma˥˥ji˥˥tsja˥˥tsu˥˥tɕ,| ˊɕxN˥˥ɕxN\pu˥˥si˥˥.|||

ˊji˥˥sjɔN; |sɕan˥˥tshje ˊtɕw˥˥tao˥˥, ˊjɔu˥˥jɔu˥˥y\ti˥˥.|| |tsi˥˥\kwa, ˊɕw\jɔN ˊtɕw˥˥yɛN;|| ˊʒɛN\lei˥˥|
ʒ\pu˥˥kae˥˥san˥˥su˥˥;| ʒ˥˥xʒ\kɔu ˊsin˥˥ma˥˥.|||/

<.: +> [ˊjɔu˥˥jɔu˥˥y\ti˥˥.|||

|ja˥˥ma˥˥ɕY˥˥N˥˥ɕɕ ˊʒx\taɕɕ ˊy˥˥liN˥˥tɕON˥˥|| ˊjɔu˥˥jɔu˥˥tɕON˥˥ ˊtɕjaɔ˥˥| ˊjɕɕjON˥˥dɕ ˊdON˥˥wɕ˥˥.|||
|tɕɕ.dɕON˥˥jɕɕjON˥˥| |sɕy˥˥: ˊɕw\tɕje˥˥.ʒɕN˥˥| ˊdɕweI\taɕ.dɕ ˊɕw\jɕɕy˥˥.||| ˊmeI˥˥tshjeN ˊjao˥˥tɕhu˥˥||
ˊji˥˥djɛɕ ˊjɔu˥˥wɕN˥˥dɕy˥˥ ˊma˥˥j|||

ˊʒAN˥˥ʒx˥˥| ˊʒɕN˥˥ʒɕN˥˥ ˊta˥˥weI ˊdɕIN˥˥\thɕN˥˥dɕ ˊɕw˥˥.|| ˊjɕɕjON˥˥| ˊjɔu˥˥ |jɕɕdɕON˥˥ɕi\ɕY˥˥.|||
ˊtha˥˥tɕhu˥˥ ˊma˥˥j˥˥ɕw˥˥.|| ˊdɕye\pɕ˥˥| ˊdɕɕN˥˥\tɕɕN˥˥ ˊɕa˥˥tɕye˥˥.||| ˊdan˥˥tha ˊphɕO˥˥khaɕ ˊmeI˥˥ji˥˥-gɕ˥˥|
ˊjɔu˥˥ ˊtɕhɕN˥˥tɕhjeN˥˥sɕAN˥˥wɕN˥˥dɕy˥˥ ˊma˥˥j˥˥dɕ ˊji˥˥ɕye˥˥ɕw˥˥.|| ˊtha˥˥dɕjɔy˥˥.dɕy˥˥tɕhu˥˥ ˊtɕhi˥˥tɕON˥˥
|jɕɕɕjaɔ˥˥pɕ˥˥vɕN˥˥.||| ˊdɕweI˥˥twɕ˥˥: ˊweI˥˥: ˊwu˥˥bɕɕtɕw˥˥.||| ˊtha˥˥\twɕ˥˥: ˊɕxN˥˥\ɕja˥˥dɕ ˊma˥˥j˥˥:
|pɕ˥˥-wɕm˥˥bu˥˥wɕN˥˥.||| ˊdɕY˥˥tɕw˥˥: |tɕhɕ˥˥ɕY˥˥tɕɕ˥˥ |ɕja˥˥jɕɕgɕ ˊmu˥˥pˊjao˥˥.||| ˊkhɕ\tɕjeN˥˥ ˊjɕɕjON˥˥
ˊtɕhIN˥˥tɕu˥˥dɕ ˊtɕw˥˥dɕO˥˥.||| |jaɔ ˊɕw ˊdɕy˥˥tɕ˥˥dɕ ˊdɕON˥˥tɕhY˥˥N˥˥ ˊɕxN˥˥tshwɕN\ɕja˥˥tɕy˥˥||
|tɕjɔy˥˥|b˥˥ɕy˥˥ ˊjɔu˥˥ ˊma˥˥j˥˥dɕja˥˥tsu˥˥dɕ˥˥| ˊɕxN˥˥ʒɕN˥˥bɕ˥˥ɕi˥˥.|||

ˊjɕɕjON˥˥: |sɕɕN˥˥tɕhje ˊtɕw˥˥dɕO˥˥ ˊjɔu˥˥jɔu˥˥y\ti˥˥.||| ˊdɕi˥˥\kwa ˊɕw\jɔN ˊdɕw˥˥yan˥˥.|| ˊʒɕN˥˥\lei˥˥|
ʒ\pɕ˥˥gɕɕ ˊsan˥˥su˥˥.| ʒ˥˥xʒ\kɔy ˊɕIN˥˥ma˥˥.|||]

6.2.2. The actual audio file, instead, presented some interesting realizations: each will be presented along with its ‘normalized’ counterparts. Sometimes, the phonetic reality contrasts quite noticeably with what one would expect: thus, where necessary, a concise explanation will be provided.

Starting with ‘*yī-zhǒng jiào yíxióng de dòngwù*’, the speaker produces [jɔy˥˥jɕɕdɕON˥˥], instead of [ˊjɔy˥˥jɕɕtɕON˥˥], underlining the numeral *yī* /ji/ ‘one’; this implies a general reorganization of the rhythmic unit, [ˊ\$,\$,\$] → [ˊ\$,\$,\$], also under the influence of the primary stress of [tɕjaɔ˥˥] nearby.

In *yíxióng de dòngwù*, the audio file presents [ˊʒɕjON˥˥dɕ |tɕN˥˥wɕ˥˥], with [ʒ] before a front vowel (a permitted alternative to [j]), and the primary stress on *yǐ*, as this is the lexeme –meaning ‘ant’– that distinguishes a *yíxióng* ‘anteater’ (lit. ‘ant-bear’) from any other kind of *xióng*. Besides, note [dɕ], instead of [-dɕ], as [ˊ] allows for an easier transition from [ˊ] to [l]. Finally, [lton˥˥wɕ˥˥] reveals a semi-neutral pronunciation, *dòngwu*, which is nevertheless used by professional voices as well; the unvoiced, non-lenited [t] is as normal a taxophone, in [ˊ\$], as [t, d] would be, too.

7. Phonetic transcriptions in current handbooks.

7.1. We are perfectly aware that such a highly detailed phonotonic transcription, as that presented in this work, would be quite challenging for beginners with no previous training in linguistics and phonetics. It must be remembered, by the way, that learning Chinese implies learning Chinese characters, too, and that requires a huge amount of time and energy on the part of the learners. It is not surprising, then, that so many students do not even think about studying Chinese pronunciation in a systematical way: in their view, the task is too complicated and demanding, to be undertaken since the very beginning, when *any* other aspect of the language (its script, as said, but also its grammar and vocabulary) is to be learnt from scratch.

We shall reply to that by saying, first of all, that being able to *write* Chinese quick and swift, while *speaking* it so bad that even native speakers do not understand what we say, is simply pointless, and disrespectful as well. Secondly, in our opinion, most handbooks are not very well balanced, with respect to the space devoted to the basics of pronunciation, though many ‘pronunciation drills’ are generally presented throughout the first one or two volumes of the series (as in the *Practical Chinese Reader* and similar contemporary textbooks).

The ordinary approach is that of letting students improve their pronunciation simply out of mere imitation and repetition of audio files. With few exceptions, *that alone does not work*. Hence the need of specifically prepared pronunciation handbooks, for those who want a more complete perspective, and aim at acquiring a native –or nearly-native– pronunciation.

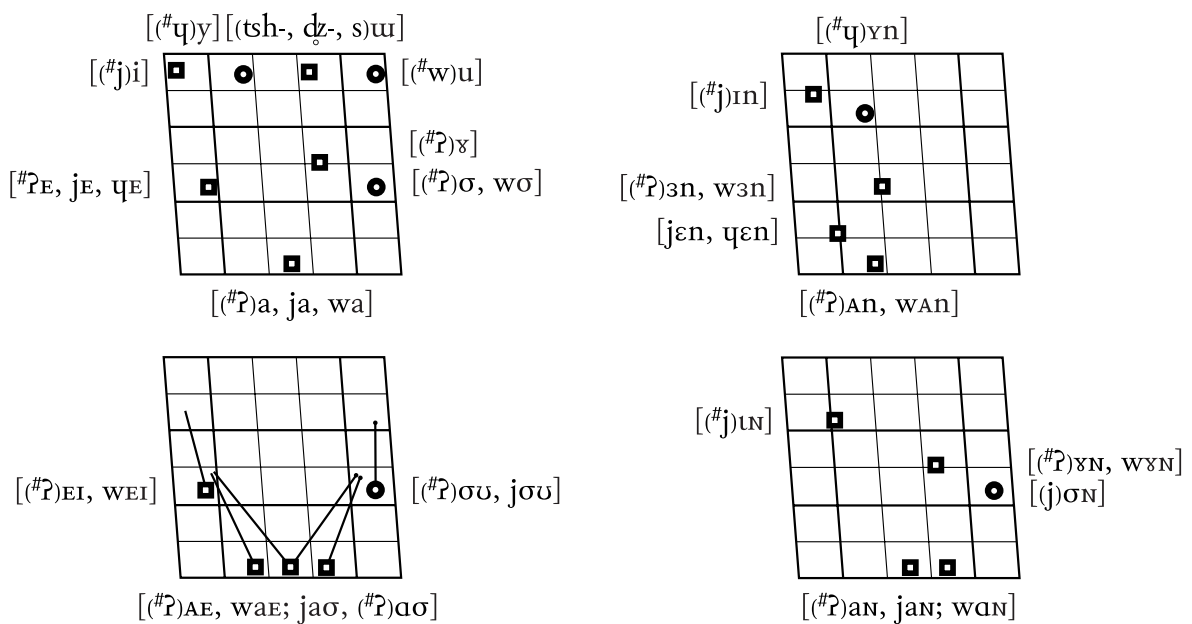
This work, in fact, can be certainly used as a pronunciation handbook, as it provides a detailed description of all phonotonic phenomena that are relevant to the Mandarin pronunciation system. Of course, in an ordinary language course, such an abundant amount of information could not be easily condensed and presented in a readable form. Beginners would invariably find that cumbersome, and not as useful as it can actually be, after a certain degree of fluency and familiarity with the overall ‘sound’ of the language has been attained.

Therefore, a slightly simplified –but not inaccurate– version of our ‘normalized’ pronunciation is the most advisable and viable solution, for teaching purposes. This ‘soft’ approach has another advantage, that is *leaving no student behind*, even the least interested in phonetics, and pronunciation in general; while the more enthusiast –and gifted– ones will find what is needed to build a strong basis, which will enable them to deal with more complex analysis, with relative ease and proficiency.

7.1.1. Our proposal is simple and effective: for each group of [C] taxophones, just *one* contoid is used, thus avoiding the aggravation of rigidly respecting the three-way consonantal ‘gradation’, that is natural for native speakers, but not necessarily for foreign learners; namely, our choice fell on [C_h, C̣], for stops and stop-strictives, and on [C̣], for constrictives. Here follows a complete list:

| | | | | |
|-----------------|-----------------|---------------|---------------|--------------------|
| <i>b</i> [ḅ] | <i>p</i> [ph] | <i>f</i> [f] | <i>m</i> [m] | |
| <i>d</i> [ḍ] | <i>t</i> [th] | | <i>n</i> [n] | <i>l</i> [l] |
| <i>z</i> [dẓ] | <i>c</i> [tsh] | <i>s</i> [s] | | |
| <i>j</i> [dʒ̣] | <i>q</i> [tʃh] | <i>x</i> [ç] | | <i>y</i> [j] & [ɥ] |
| <i>zh</i> [dʒ̣] | <i>ch</i> [tʃh] | <i>sh</i> [ʃ] | | <i>r</i> [ʒ] |
| <i>g</i> [g̣] | <i>k</i> [kh] | <i>h</i> [x] | <i>ng</i> [ŋ] | <i>w</i> [w] |

fig 7.1. Simplified vowel inventory, for teaching purposes.



7.1.2. For [(j)V, (j)VV, (j)VN], a similar criterion is followed, but without re-nouncing realism: thus, a diphthong like [σʊ] will not be simplistically reduced to ‘[σu]’, even though [ʊ] is not used individually elsewhere. In other words, simplicity must be combined with realism (fig 7.1).

After a well meditated consideration, we have found that fourteen vocoids are ideal for a very convincing approximation to neutral Chinese pronunciation: [i, ɪ, ɪ; ɛ, ɛ; ʌ, a, ɑ; ɜ, ʌ; ʊ; u, ʊ; σ]; to these, we add [ʒ, ʒ̣, ʒ̥] (for ⁿ[ʊ] and the various occurrences of *érhuà*). We keep [i, ɪ, ɪ], [ɛ, ɛ] and [ʌ, a, ɑ] distinct, as well as [ɜ, ʌ] and [u, ʊ], instead of simply [i, ɛ, a, ʌ, u], because certain combinations (especially those with [-n, -n]) are actually *easier* to reproduce, if the inventory available is just a bit more complex than what would be phonemically indispensable.

| | | | | | | | |
|---------------|----------|-------------|----------|------------|-------|-------------|-------|
| A | | | | | | | |
| <i>a</i> | [a] | | | | | | |
| <i>ai</i> | [ʌɛ] | <i>uai</i> | [wʌɛ] | <i>ao</i> | [ɑσ] | <i>iao</i> | [jaσ] |
| <i>(w)an</i> | [(w)ʌn] | <i>ian</i> | [jɛn] | <i>üan</i> | [ʏɛn] | | |
| <i>(i)ang</i> | [(j)ʌn] | <i>uang</i> | [wʌn] | | | | |
| E | | | | | | | |
| <i>e</i> | [ɜ] | <i>ie</i> | [jɛ] | <i>üe</i> | [ʏɛ] | <i>ei</i> | [ɛɪ] |
| <i>en</i> | [ɜn] | <i>er</i> | [ɜʒ] | <i>eng</i> | [ɜn] | <i>ueng</i> | [wɜn] |
| I | | | | | | | |
| <i>i</i> | [i, #ji] | <i>ï</i> | [ʊ, ʒ̣] | <i>ui</i> | [wɛɪ] | | |
| <i>in</i> | [(#j)ɪn] | <i>ing</i> | [(#j)ɪn] | | | | |
| O | | | | | | | |
| <i>o</i> | [σ] | <i>ou</i> | [σʊ] | <i>uo</i> | [wσ] | | |
| <i>ong</i> | [σn] | <i>iong</i> | [jσn] | | | | |
| U | | | | | | | |
| <i>u</i> | [(#w)u] | <i>ü</i> | [(#ʏ)ʏ] | <i>iu</i> | [jσʊ] | | |
| <i>un</i> | [wɜn] | <i>ün</i> | [(#ʏ)ɜn] | | | | |

It is not hard to notice that our selection tries to incorporate as many vocoids as possible that are in common with the two major accents of English: modern British English, and the so-called ‘General American’ English. Of course, both native and non-native English speakers will use only one of the two accents (or even a different

one, from Australian, or Canadian, to Irish and Scottish, and so forth, not to mention the huge number of regional US accents other than the ‘General’ one). However, what matters is trying to keep certain sounds distinct.

For instance, neutral British speakers will certainly be able to produce a perfect [ɛ], but may find that their [æ] is a more familiar replacement for [ɛ]: let it be so (many Chinese themselves, after all, say [jæ˥n] instead of [jɛn]). Similarly, while [ɜ] is common in British English (eg *fur* [ˈfɜː]), most Americans could not be able to distinguish it from their regular [ə]: no panic, let them use [ə], at first; listening to audio files *and* practicing kinesthesia, with a constant ear feedback, will enable those really interested in acquiring [ɜ] –and, at a more advanced stage, also [ɛ]– to succeed. The same story goes, within certain limits, for those who speak languages admitting even fewer distinctions: for instance, neutral Italian and Spanish have only [a], but not [A], nor [ɑ]. Articulating [a] everywhere is fine, and phonemically correct; yet it is preferable for the transcription itself to keep using three symbols anyway, as there will be some students who, being more sensitive than others to the quality of vowel sounds, want to know *where* [A] and [ɑ] should be used, in case they manage to acquire them. In other words, this ‘simplified’ transcription is the ideal intermediary between the essentiality of *pīnyīn*, and the full realism of the neutral ‘normalized’ pronunciation.

Consequently, what we propose here is not the ‘international’ pronunciation suggested in § 2.8.3, on the basis of that originally presented in our *Pronuncia cinese per italiani*, which was intended –primarily– for someone who simply wants to be understood, irrespective of how foreign his or her accent may sound. Here, instead, we are dealing with something more sophisticated, and more suitable for those who learn Chinese... the ‘professional’ way, so to speak.

7.1.3. As far as *érhuà* is concerned, we have devised a reasonable scheme, derived from that proposed in § 2.7, with just minor adjustments. Then, we have:

- A. [aʔ]: *huà-huà* [ˈɕwa\ɕwaʔ] (ⁿ[ˈɕwa\ɕwəʔ]);
 [jaʔ]: *chuāngliánr* [ˈtʃhwan˥˥jaʔ] (ⁿ[ˈtʃhwan˥˥jəʔ]);
 [ɥaʔ]: *yuánquānr* [ˈɥɛn˥˥ɥaʔ] (ⁿ[ˈɥɛn˥˥ɥəʔ]);
 [waʔ]: *hǎowánr* [ˈxɑŋ˥˥waʔ] (ⁿ[ˈxɑŋ˥˥wəʔ]).

- B. [ɣʌ]: *bèikér* [ˈbɛi˥˥kʰɣʌ] (ⁿ[ˈbɛi˥˥kʰɣʌ]);
 [jɣʌ]: *shǐjìnr* [ʃɿ˥˥dʒɣʌ] (ⁿ[ʃɿ˥˥dʒɣʌ]);
 [ɸɣʌ]: *qiūqūr* [ˈtɕʰy˥˥tɕʰɸɣʌ] (ⁿ[ˈtɕʰy˥˥tɕʰɸɣʌ]);
 [wɣʌ]: *kǒuwèir* [ˈkʰɔu˥˥wɣʌ] (ⁿ[ˈkʰɔu˥˥wɣʌ]).
- [jɤʌ]: *shǐjiēr* [ʃɿ˥˥dʒɤʌ] (ⁿ[ʃɿ˥˥dʒɤʌ]);
 [ɸɤʌ]: *mùjüér* [ˈmʊ˥˥dʒɸɤʌ] (ⁿ[ˈmʊ˥˥dʒɸɤʌ]).
- C. [σʌ]: *tǔdòur* [ˈtʰu˥˥dσʌ] (ⁿ[ˈtʰu˥˥dσʌ]);
 [jσʌ]: *zhīliúr* [ˈdʒʰi˥˥lʃσʌ] (ⁿ[ˈdʒʰi˥˥lʃσʌ]);
 [ɑσʌ]: *nǎimáor* [ˈnɑi˥˥mɑσʌ] (ⁿ[ˈnɑi˥˥mɑσʌ]);
 [jaσʌ]: *màimiáor* [ˈmɑi˥˥mjaσʌ] (ⁿ[ˈmɑi˥˥mjaσʌ]);
 [wσʌ]: *shānpōr* [ˈʃɑm˥˥pʰwσʌ] (ⁿ[ˈʃɑm˥˥pʰwσʌ]).
- D. [āʌ]: *piānchǎngr* [ˈpʰjɛŋ˥˥tɕʰhāāʌ] (ⁿ[ˈpʰjɛŋ˥˥tɕʰhāāʌ]);
 [jāʌ]: *dòujiāngr* [ˈdɔu˥˥dʒʰjāʌ] (ⁿ[ˈdɔu˥˥dʒʰjāʌ]);
 [wāʌ]: *tiānchuāngr* [ˈtʰjɛŋ˥˥tɕʰhwāʌ] (ⁿ[ˈtʰjɛŋ˥˥tɕʰhwāʌ]).
- E. [ǎʌ]: *bǎndèngr* [ˈbɑn˥˥dǎʌ] (ⁿ[ˈbɑn˥˥dǎʌ]);
 [jǎʌ]: *qímíng* [ˈtɕʰi˥˥mǎʌ] (ⁿ[ˈtɕʰi˥˥mǎʌ]);
 [wǎʌ]: *xiǎowèngr* [ˈɕjɑσ˥˥wǎʌ] (ⁿ[ˈɕjɑσ˥˥wǎʌ]).
- F. [ōʌ]: *tōukòng* [ˈtʰɔu˥˥kʰōʌ] (ⁿ[ˈtʰɔu˥˥kʰōʌ]);
 [jōʌ]: *kūqióng* [ˈkʰu˥˥tɕʰjōʌ] (ⁿ[ˈkʰu˥˥tɕʰjōʌ]).

Needless to say, the learners can choose to make a step forward, and acquire the full inventory of native *érhuà* instead: actually, in the course of using ‘simple’ [-ʌ], many speakers spontaneously will end up articulating [Y], due to the effect of co-articulation and articulatory economy, which obviously give preference to what makes pronunciation easier and more natural.

7.2. When the present ‘simplified’ phonetic transcription is used in ordinary handbooks, and in extended texts, no phonemic transcription is needed, as *pīnyīn* certainly suffices to give a rough idea of how the sound system of Mandarin Chi-

nese works. Instead, tones and intonational units shall always be presented in full notation, as any simplification could easily jeopardize mutual intelligibility. Here is an example of connected speech, transcribed according to the simplified pronunciation described above:

<.: +> [ʔljɔu.jɔu'ɥy\di.].||
 \ja.ma\cYN'ɣɣ 'ɰ'ɰ\daE ..ɥy'liŋ~dɰON|| ~jɔu\ji.dɰON· \dɰjaɔ· |.ji'cɰON·dɣ 'dɰON\wu.].||
 \dɰɣ.dɰON .ji'cɰON·| \sɰ': 'sɰ'\dɰjE..sAN,| 'dɰWEI\da.dɣ· 'sɰ'ji\ɣɔu.].|| ..MEI~thjɛN 'jaɔ~tɰhɰ'.||
 ~ji·dɰjaɰ· 'ljɔu\wan.dɰ' 'ma.ji.].||
 'ɰAN'ɰɰ'ɰ' | 'ɰAN'ɰɰN· \|da'WEI 'dɰIN\than.dɣ 'sɰ'.|| .ji'cɰON·| jɔu \ji.dɰON·ci\cIN.].||
 ~tha'tɰhɰ' 'ma.ji 'sɰ'.|| 'dɰɥE\bu·| ..dɰan\dɰIN 'sA'dɰɥE.].|| 'dan~tha 'phɔɔ~khAE ..MEI'ji·gɣ·|
 ~jɔu 'tɰhɣN'tɰhjɛN'sAN\wan.dɰ' 'ma.ji·dɣ· |ji'cɰɥE'sɰ'.|| ~tha\dɰjɔu.dɰ'~tɰhɰ' 'tɰhi~dɰON·
 \ji.cɰjaɔ\bu.fɰN.].|| 'dɰWEI~dɰɔ'': 'WEI': 'wu.bAE~dɰɥ.].|| 'tha\dɰWEI· 'sɣN\cɰja.dɣ· 'ma.ji':
 \bu·wɰm~bu\wɰN.].|| 'dɰIN~dɰ'': |tɰhy 'cYN~dɰɔɔ· |cɰja.ji·gɣ 'mu~bjɔɔ.].|| ..khɣ\dɰjɛN·,
 .ji'cɰON· ~tɰhɰN~tɰshu·dɣ ~dɰ'~dɰɔɔ.].|| \jaɔ sɰ' |dɰɥ~dɰ'~dɰ' ..dɰON'tɰhɣŋɔ ~sɣN'tshwɰp·
 |cɰja.tɰhy.].|| \dɰjɔu\bi~cɣ· ~jɔu 'ma.ji'dɰja'dɰu·dɣ·| ~sɣN'sɣN\bu~ci.].||
 .ji'cɰON· |sAN~tɰhjɛ ~dɰ'~dɰɔɔ· 'ljɔu.jɔu'ɥy\di.].|| 'dɰi\ɰwɰ'· .sɰ'ljɔN 'dɰɥ~ɥɛN.].|| 'ɰɰ\lɛI·|
 ɰ'bu'gAE 'sAN~su.].| ɰ'ɰɰ\ɰɔu 'cIN~ma.].||]

7.3. As we have just demonstrated, introducing phonetics in a common Chinese course is not that complicated, nor confusing, provided the teacher knows how to present the subject in a professional and interesting way. If phonetic symbols are inserted, here and there, just for the sake of it, without a consistent method, they will be certainly ignored, if not frowned upon.

The Chinese teachers *must* know (at least) the basics of their students' mother tongue, including its phonological and phonetic system, and should be able to distinguish all the relevant vocoids and contoids. Sadly, most teachers from China have a certain command of English, but not a scientific knowledge of its pronunciation; rarely do they speak any other foreign language, *before* going to the country where it is spoken, and working there for a while.

When we started learning Chinese at Rome University, back in 1998, one of the first 'oddities', from the point of view of an Italian native speaker, was to determine the exact value of *p*, *t*, *k*, as opposed to *b*, *d*, *g*. Reading the three couples according to their Italian pronunciation –[p, b], [t, d], [k, g]– they sounded identical to our

Chinese teacher. He could not help insisting on the ‘aspiration’ factor, for *p, t, k* /ph, th, kh/, totally ignoring the fact that [Ch] do not belong to neutral Italian, unless one is using emphasis, and *randomly* articulates a *slightly* ‘aspirated’ contoid, which nevertheless *has no phonemic relevance*. It was a really frustrating situation, with both parts stuck in a deadlock, and unable to tell what was wrong in the first place. One can easily figure out what invariably happens, when foreign students try to ‘refine’ their tones and, even, dare to learn Chinese intonation, and their teacher has nothing more to offer, but going on and on repeating the sentences contained in the textbook, ‘hoping for the best’. All this, of course, because no reliable and scientifically coherent notational system for tones or intonation is available.

7.3.1. The general lack of interest in the potential of phonetics, now apparently prevailing among most teachers and the authors of language courses, does not prove that this has always been the case. In fact, as far as Chinese is concerned, it has to be remembered that one of the first and most prestigious textbooks ever published (besides Zhào Yüánrèn’s *Mandarin Primer*) is the well known *Chinese*, by the Linguaphone Institute in London, whose first edition dates back to almost a century ago. This brilliant product is *entirely* transcribed in IPA (also because *pīnyīn* was yet to come, in the Thirties), including quite an extensive tonetic notation, with thick symbols for the tones in [ˆ\$], and thin ones for the tones in [˘\$].

It is worth noticing that the phonetic notation adopted in *Chinese* is surprisingly very similar to the ‘simplified’ one presented above, notwithstanding some peculiar choices, mainly due to typographical limitations (and the general ‘sobriety’ of IPA in those –otherwise glorious– pioneering days). Here is a complete list of the [C] in the two systems:

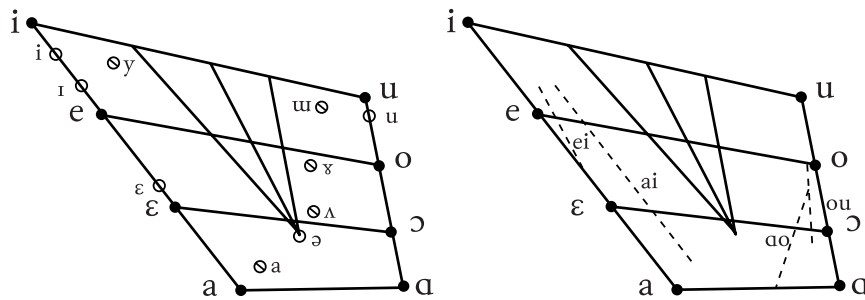
| | | | |
|---------------------|---------------------|-----------------------|-----------------------|
| <i>can</i> [ḅ, ph] | <i>off</i> [ḅ, ph] | <i>can</i> [ḍz, tsh] | <i>off</i> [ḍz, tsh] |
| <i>can</i> [ḍ, th] | <i>off</i> [ḍ, th] | <i>can</i> [ḍʒ, tʃh] | <i>off</i> [j̣, ch] |
| <i>can</i> [g̣, kh] | <i>off</i> [g̣, kh] | <i>can</i> [ḍʒ, tsh] | <i>off</i> [ḍʒ, tʃh] |
| <i>can</i> [x̣] | <i>off</i> [x̣] | <i>can</i> [ʒ̣] | <i>off</i> [ɹ̣]* |
| <i>can</i> [ʃ̣] | <i>off</i> [ʃ̣] | | |
| <i>can</i> [ç̣] | <i>off</i> [ç̣] | | |

*described as a constrictive: *can*[□]

[V] are treated in a less similar way, yet the two systems remain largely coherent and –so to speak– *mutually intelligible* (fig 7.2). Please observe the following list:

| | | | |
|--------------------------------|---|-------------------|-------------------|
| <i>can</i> [jɛ, ɸɛ] | <i>off</i> [iɛ, yə] | <i>can</i> [(j)a] | <i>off</i> [(V)a] |
| <i>can</i> [ɤ] | <i>off</i> [ɤ]; after [ḡ, kh, x]: [ʌ] | | |
| <i>can</i> [wσ] | <i>off</i> [uɔ]; after [ph, b, m, f]: [ɔ] | | |
| <i>can</i> [u] | <i>off</i> [u] | <i>can</i> [ɿ] | <i>off</i> [ɿ] |
| <i>can</i> [(w)ɛɪ, (w)ɛɛ] | <i>off</i> [(u)ɛɪ, (u)ɑɪ] | | |
| <i>can</i> [ɑσ, jaσ, (j)σʊ] | <i>off</i> [(i)ɑσ; ou, iu] | | |
| <i>can</i> [ɪn, ʏn, ɯn] | <i>off</i> [ɪn, ʏn; ɪŋ] | | |
| <i>can</i> [jɛn, ɸɛn] | <i>off</i> [iɛn, yan] | | |
| <i>can</i> [(w)an, (j)an, wɑn] | <i>off</i> [(u)an, (i)ɑŋ, uaŋ] | | |
| <i>can</i> [(w)ɜn] | <i>off</i> [əɪ, un; ʌŋ] | | |
| <i>can</i> [(j)σn] | <i>off</i> [(i)uŋ] | | |

fig 7.2. Vowel diagrams as presented in *Chinese*, by the Linguaphone Institute.



7.3.2. Clearly, this analysis was produced when no highly technological equipments were available to phoneticians. They simply had their ears, their brains, some quite rudimental phonographs, and a limited number of Chinese informants at hand. They employed such ‘antediluvian’ devices as the *artificial palate*, modeled out of fragile resin, or even papier-mâché; the palate had to be covered in chalk powder, or cocoa powder, then inserted in the mouth, and immediately removed, after the relevant phone had been uttered, to determine its articulatory details.

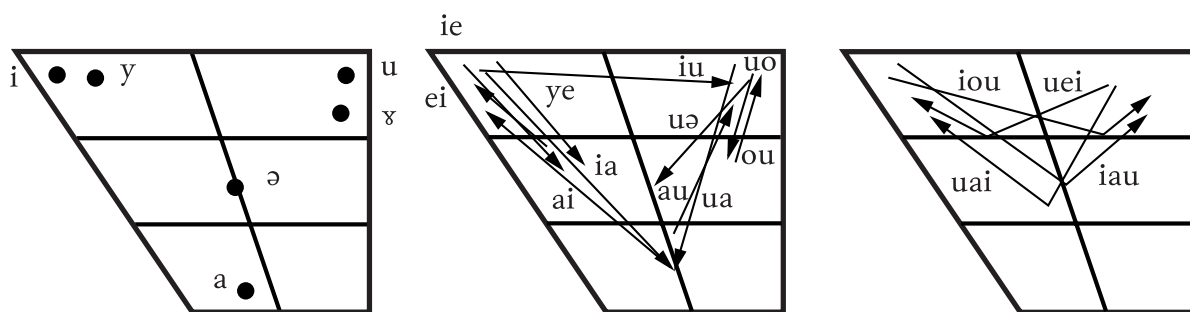
Vocoids were even harder to analyze: a thin metallic chain, often with a small lead ball in the middle, had to be kept on the tongue surface, refraining oneself from spitting it out, or swallowing it; at the same time, the phonetician had to pronounce a vocoid, or a diphthong, as clearly and naturally as a normal human being –in such a tight spot– can do, while an X-ray apparatus would take multiple photo-

graphs of the phonatory act (with a much higher exposure to radiations than in modern X-ray procedures). Nonetheless, those pioneers would do their best, not with the purpose of a mere display of scholarly expertise, but to provide their readers with something that could be really useful to learn languages better.

The idea that *IPA* should be limited to serious articles, essays and books ‘for insiders only’ (which no-one else would ever read) belongs to more recent generations of phoneticians. On the contrary, the International Phonetic Alphabet was created, primarily, as a practical tool for teaching languages *to everyone*. A radiant example of that is precisely what we find in the first Linguaphone courses (on 78 *rpm* records, and in their second editions, on 45 *rpm* records); unfortunately, the succeeding editions –on audiocassettes, and CD’s– reduced the phonetic sections to one plain and cold page, or completely abolished any *IPA* symbols..

7.3.3. Given such premises, it is quite hard to understand why the *Journal of the Phonetic Association*, as recently as in 2003, sanctioned an *official* description of Mandarin Chinese, by LEE, W-S. & ZEE, E., *Illustrations of the IPA – Standard Chinese (Beijing)*, which is embarrassingly inferior to what was available decades before. Just have a look at the ‘official’ vowel diagrams (fig 7.3).

fig 7.3. Vowel diagrams as presented in *Journal of the Phonetic Association* (2003).



Apart from the questionable collocation of [ɤ] in the first diagram, the other two trapezia mix up true diphthongs, like ‘[ei, ai, au]’ (^{can}[EI, Aɛ, ɑO]), and what should be treated not as diphthongs, or even ‘triphthongs’, but as [JV(V)] sequences, eg ‘[ye, uei, iau]’ (^{can}[ɥɛ, wɤI, jaO]). Regarding ‘[uai, iau]’, we cannot help criticizing the lower ‘angle’, which should indicate ‘[a]’, but –clearly– rather lies in the area of *off*[ə] (for ‘[uai]’, the angle might be rendered as *off*[ɜ], too, but nothing lower than that; not even *off*[ɐ], which would appear less absurd).

The officially sanctioned symbols for Mandarin [C] are not ‘wrong’ *in se*, but are certainly far from being the ideal choice. For one thing, *postalveolar* /tʂh, tʂ, ʂ, ʐ/ appear as ‘[tʂ^h, tʂ, ʂ, ʐ]’, maybe to avoid that someone would interpret ‘[tʂ^h, tʂ, ʂ, ʐ]’ as the corresponding *apicopalatal* contoids, used –for instances– in many Dravidian languages (^{can}[tʂ^h, tʂ, ʂ, ʐ]). We shall not repeat here what we have said about the official treatment of /ʍ/ [ʍ, ʍ̩]. Besides, the authors do not even mention the possibility of some kind of consonantal gradation, nor do they draw any clear line between phonemics and phonetics. Tones and taxotones are treated vaguely and clumsily, while intonation is... *not discussed at all*. If the members of the International Phonetic Association really accept those FOUR pages as *the* official description of one of the most important –and most studied– languages in the world, we shall ‘let silence reign’, instead of speaking frankly, and hurting somebody’s feelings.

7.3.4. There is, nonetheless, something that must be said, loud and clear. What is most annoying in many official analyzes is not the use of *offIPA in se*, but rather its *misuse*. Undeniably, *offIPA* is structurally simpler than *canIPA*: from one point of view, this is a practical advantage, but also a functional limitation. In other words, when one is dealing with just a few dozens of symbols, and is not forced to be precise, a parlous mistake is less likely to occur (hence our harsh criticism of the slovenly treatment reserved to ‘[uai, iau]’ in the *JIPA* diagrams); *canIPA* symbols, on the contrary, number in the hundreds, and that obviously requires a good deal of care and perfectionism. However –be noted– this greater effort eventually pays off, whenever a transcription must be accurate and detailed, but should not lose its readability and –why not?– even its ‘typographical appeal’.

Since the introduction of *IPA* to date, AD 2011, many *generic* analyzes have been realized for most languages. The only reason why someone may be interested in writing another essay on the pronunciation of a language is trying to do something better than it has been so far; possibly, with a better system. Invariably, *canIPA* works more efficiently than its official ‘competitor’, when the complexity of certain languages (including Chinese) poses so demanding challenges, that excessive simplicity would result in a waste of time. A more flexible system is, of course, harder to use *properly*, and we do not deny the risks that this aspect involves.

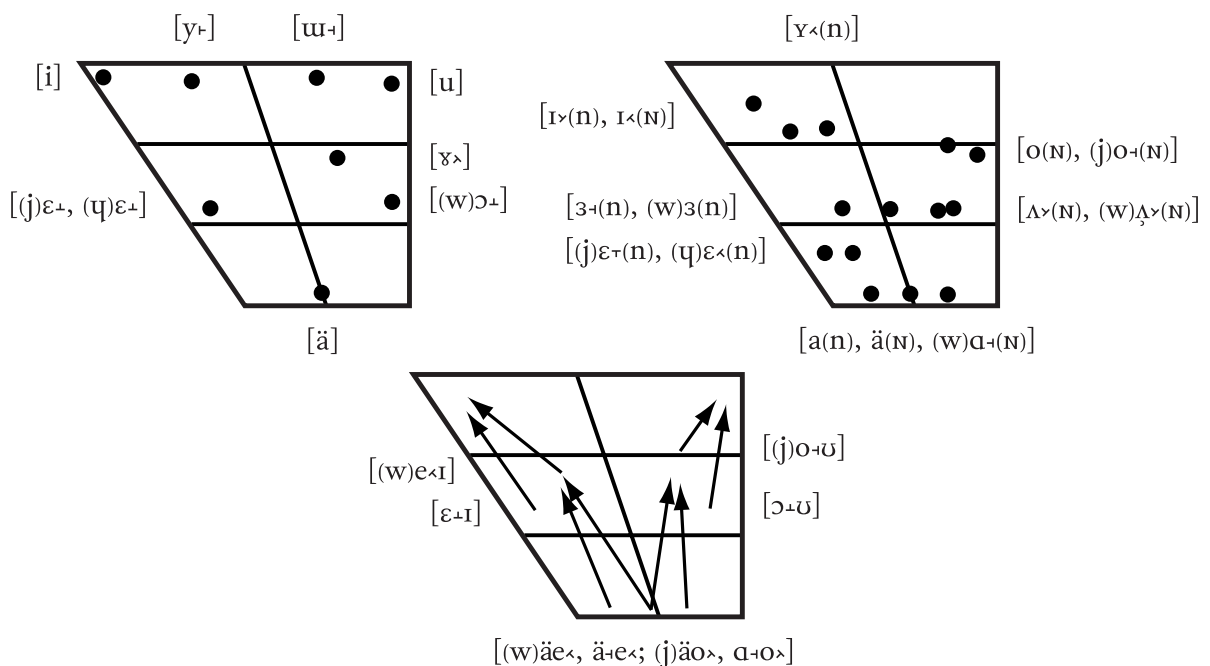
At the same time, we would happily expect the same conscientiousness on the part of those who remain loyal to the old school. If a transcription aims at being

simple, let it be simple... but not sloppy; consequently, if the generic symbols adopted may have numerous values, an author is *compelled* to pay the utmost attention to what we call *vocograms* and *orograms*: if these are drawn correctly, an otherwise ambiguous transcription regains reliability and, ultimately, its scientific relevance. Whether this was the case with *JIPA's Standard Chinese*, we will let our readers judge by themselves.

7.3.5. More important than that, we are interested in proving that *offIPA* can be used with appreciable results, provided everything is done to compensate for its transcriptional limitations. First of all, no transcription should ever become an ugly conglomeration of diacritics. Sure, there are some *canIPA* symbols with *three* diacritics, usually representing peculiar articulations *or* co-articulations, as in $[\tilde{x}]$: *nasalized, laterally contracted, and laryngealized*. But, in most cases, unitary symbols are available, and no diacritic is necessary at all.

When using *offIPA*, instead, diacritics are inevitable almost everywhere: the point is deciding whether all of them should be printed –or handwritten– at all times, or should they be presented only once, in conjunction with the relevant illustrations. In *fig 7.4*, we give a practical example: only official symbols are used, apart from *can* $[\pm \prec \succ \tau \times \dot{\succ}]$, whose value is intuitively self-evident.

fig 7.4. Official vowel diagrams of Chinese ‘normalized’ pronunciation (with *offIPA* symbols).



The diacritics help understand which symbol corresponds to which dot (or arrow); but, once the exact value of a certain vocoid has been understood, thanks to the diagrams, the diacritics could be gradually abandoned, until only primary symbols are used. That would work acceptably well, if we stick to a teaching-oriented kind of transcription. But, as we know, the *real* pronunciation of neutral Chinese implies frequent cases of *laryngealization*: [Ṽ, ṅ, Ṇ, ḗ, ḥ]; *nasalization* is quite common, too, eg [Ṽ̃, ḗ̃]. Vowels may be *partially devoiced*, [Ṽ], but also *totally devoiced*, [Ṽ̃]. We might have to indicate more than one peculiarity at the same time, as in [ḗ̃̃], and –adding *partial labialization*– [ḗ̃̃̃]. And if we were to annotate all the subtleties involving Chinese [C], some symbols could feature even more diacritics!

Now, it is clear that doing the same in *offIPA* would not be viable, especially in continuous transcriptions. The only option would be that of skipping all basic diacritics, whenever more specific ones have to be inserted. Frankly, that does not seem very satisfying, nor entirely scientific; rather, it proves –once again– that nobody should expect to ride a bike and travel faster than sound: that is what supersonic jets have been invented for! However, a bike still can be a good bike, if we do not try to use it as a jet. The bottom line is: let us keep *offIPA* true to itself.

7.3.6. In this paragraph, we present an example of how *offIPA* can express a good deal of phonetic information, after having duly considered its typographical restraints and the purposes for which it was originally invented: learning the pronunciation of foreign languages. As such, we will not attempt to transcribe neutral Chinese, with all its taxophones and other peculiarities, but the slightly simplified pronunciation described in § 7.1.1-2. Where necessary, the exact value assigned to an official symbol will be specified between ‘[[]]’, by means of diacritics. Let us start from the consonantal inventory:

| | | | | |
|-----------------|------------------|---------------------|---------------|---------------------|
| <i>b</i> [b] | <i>p</i> [ph] | <i>f</i> [f] | <i>m</i> [m] | |
| <i>d</i> [d] | <i>t</i> [th] | | <i>n</i> [n] | <i>l</i> [l] |
| <i>z</i> [d͡z] | <i>c</i> [tʃh] | <i>s</i> [s] | | |
| <i>j</i> [d͡ʒ] | <i>q</i> [t͡ʃh] | <i>x</i> [ç] | | <i>y</i> [j] & [ɥ] |
| <i>zh</i> [d͡ʒ] | <i>ch</i> [t͡ʃh] | <i>sh</i> [ʃ] | | <i>r</i> [ɹ] [[ɹ̄]] |
| <i>g</i> [g̃] | <i>k</i> [kh] | <i>h</i> [χ] [[χ̄]] | <i>ng</i> [ŋ] | <i>w</i> [w] |

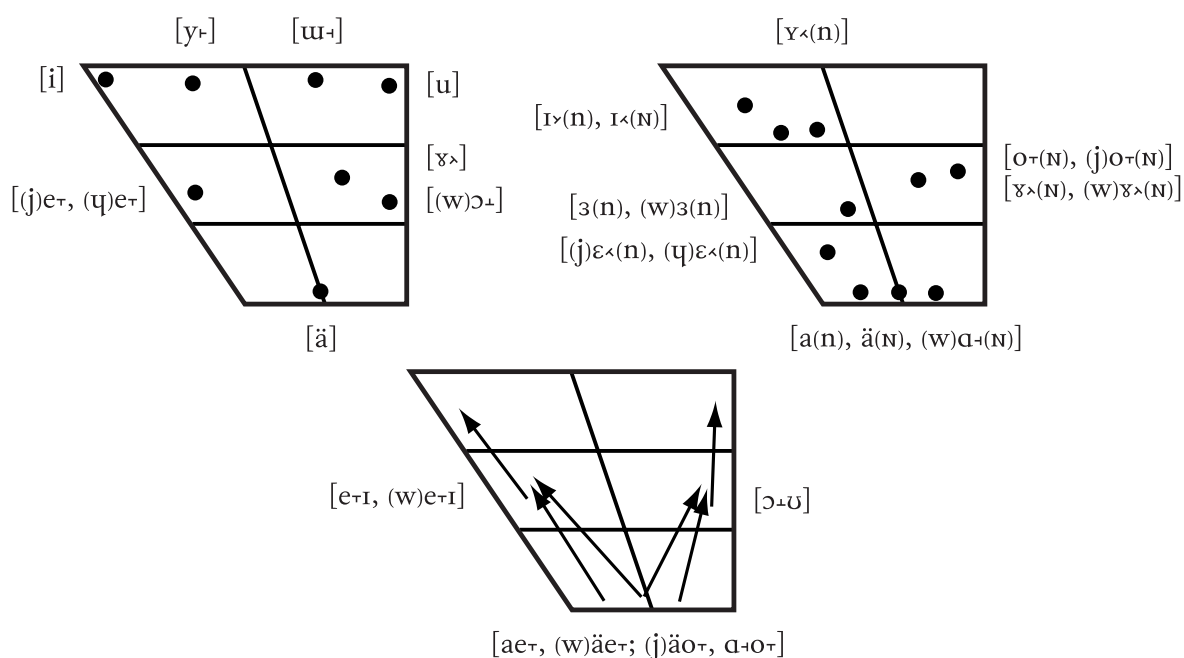
Of course, all the relevant articulatory illustrations (orograms, labiograms and laryngograms) must be provided, as in fig 3.2-7. As far as the actual symbols are concerned, *off*[C̣] may correspond to *can*[C̣, C̣̰]; this can be positively exploited, after all, by presenting *off*[C̣] as ‘synthetic’ symbols, capable of expressing a certain free oscillation. Similarly, even if *off*[h] is primarily given as a constrictive, or ‘fricative’, in *off*[Ch] it should be articulated, rather, as an approximant: *off*[[h_τ]] (*can*[h]).

For /t͡ʂ(h), ʂ, ʐ/, the corresponding ‘retroflex’ *off*[d͡ʂ, t͡ʂh, ʂ, ʐ] are fine, but the relevant orograms should help the students distinguish Mandarin postalveolars from true ‘retroflex’ contoids (*canIPA*: *apicopalatals*). The notation ‘[[_τ]]’ dispels any doubt about the exact articulation of [ʐ].

Since our font contains unitary symbols for all stopstrictives, we obviously use them: [ts, d͡ʒ; tɕ, d͡ʒ; t͡ʂ, d͡ʒ̰]. Most *Unicode* typefaces still have ‘[ts, d͡ʒ; tɕ, d͡ʒ; t͡ʂ, d͡ʒ]’, but the general trend, nowadays, is expressing [KS] as [KS] everywhere, as if they were sequences of a stop plus a constrictive, instead of single articulations. Font editing is now a relatively easy and inexpensive task, and some minor enrichments to the typefaces adopted in a textbook should not really pose any difficulty. However, if there is no choice but following the official recommendations, up to the last, the correct renderings should be [ts, d͡ʒ; tɕ, d͡ʒ̰; t͡ʂ, d͡ʒ̰], avoiding sad things like ‘[t͡ʂ, t͡ʂ]’.

Regarding the vocalic inventory, there is little to say, more than what has been clearly explained in § 7.1.2. Please refer to the official trapezia in fig 7.5, and compare them with our *canIPA* vocograms in fig 7.1.

fig 7.5. Official vowel diagrams of Chinese ‘simplified’ pronunciation (with *offIPA* symbols).



A

a [a] [[ä]]*ai* [ae] [[eː]]*uai* [wae] [[äeː]]*ao* [ao] [[aːoː]]*iao* [jao] [[äoː]]*(w)an* [(w)an] [[ä]]*ian* [jɛn] [[ɛː]]*üan* [ɥɛn] [[ɛː]]*(i)ang* [(j)an] [[ä]] *uang* [wan] [[aː]]

E

e [ɛ] [[ɛː]]*ie* [jɛ] [[eː]]*üe* [ɥɛ] [[eː]]*ei* [eɪ] [[eː]]*en* [ɛn]*er* [ɛʒ] [[ɛː]]*eng* [ɛŋ] [[ɛː]]*ueng* [wɛŋ] [[ɛː]]

I

i [i]*ï* [ɯ, ɨ] [[ɯː]]*ui* [weɪ] [[eː]]*in* [ɪn] [[ɪː]]*ing* [ɪŋ] [[ɪː]]

O

o [ɔ] [[ɔː]]*ou* [ɔʊ] [[ɔː]]*uo* [wɔ] [[ɔː]]*ong* [ɔŋ] [[oː]]*iong* [jɔŋ] [[oː]]

U

u [u]*ü* [y] [[yː]]*iu* [jɔʊ] [[ɔː]]*un* [wɜn]*ün* [ɥn] [[yː]]

For /Çw/, as said, we use [ç], ie *off*[ç], and the same contoid can be employed in the articulation of *érhuà*: [Vɿ]. For transcribing more genuine articulations, those with [Y], the official inventory offers no unambiguous solution. Clearly, we cannot employ *off*[ç] to render our *can*[ɥ], a *postalveolarized prevelar semi-approximant* (with lateral contraction, see § 2.3); the only official alternative is *off*[ɿ]. In *canIPA*, the same symbol is –more conveniently– employed for the most typical realization of American English *r*, a *postalveolarized prevelar slightly rounded approximant* (with lateral contraction, as well), which we could render as ‘[ɥɿ^w]’ in *offIPA*: [ɥ] is quite advanced, [[ɥː]], and the [ɿ] diacritic –generically described as implying ‘rhoticity’– apparently is the only passable solution to mark lateral contraction. (Again, here, [ɿ] is a simple postalveolar approximant: [[ɿː]], or [[ɿ]].)

However, since *off*[ɿ] is usually employed as an *alveolar* approximant, a notation like ‘[Vɿ]’, or ‘[Vʲ]’, could lead to legitimate criticism. As a practical solution, ‘[ɥɿ]’ might work well; for *can*[Y], then, we would have *off*[ɥr, ɣ, ɔ, ɔ̣] &c. Anyway, *offIPA* always leaves room for ‘personal preferences’! Therefore, when using a very small type size, we suggest ‘[Vɿ]’ as less hard to decipher than the tiny ‘rhoticity hooklet’.

8. Peking.

The Peking accent vs the Peking dialect

8.1. We have carefully considered whether to include analyses of regional pronunciations in a thesis that specifically focuses on *neutral* Chinese pronunciation. That would be the logical continuation of this work, and –in fact– we have already started collecting material for a thorough examination of the most important regional accents, from both the People’s Republic of China and other parts of the vast Chinese-speaking world.

To this end, we have prepared a detailed pronunciation questionnaire (enclosed in a separate appendix, at the end of this thesis). Based on the recordings of several informants and numerous other audio files, we are collecting a rich sound database, which –we hope– will be useful also to other researchers interested in the subject.

For the purposes of this dissertation, nonetheless, we think it is advisable not to venture into some descriptions that may certainly be interesting, but would also require a lot of space, to be presented in full (certainly much more than what would be advisable in a monographic essay, like the present one). Otherwise, we should limit ourselves to scattered considerations, without actually providing accurate transcriptions nor illustrations. Quite evidently, such a reading could be rather disappointing.

However, Peking certainly deserves an exception, given the close relationship between the Common Language and the dialect spoken in the capital. But it is important to point out that we will deal with the *Pekingese accent* of Mandarin Chinese, not the *Peking dialect* proper, as the latter is an independent language, with its phonological, lexical and grammatical peculiarities. Surely, there is no sharp boundary between the *accent* and the *dialect* of Peking, but rather a *continuum* of more or less marked variants, from a kind of almost neutral pronunciation, with very few regionalisms, down to a sort of perfect *diglossia*, where the speaker unconsciously oscillates, even within the same utterance, depending on the context and the linguistic register.

8.2. Peking is a huge city with about twenty million inhabitants, many of whom were not born in the capital. The vast number of immigrants from other parts of China, who clearly do not speak Pekingese natively, as well as the influence of mass media and the increasingly higher literacy rate, have deeply changed the linguistic situation in the area. Many dialects coexist with different degrees of proficiency in speaking *pǔtōnghuà*. Moving from one part of the city to another, nowadays, is a frequent phenomenon, as social and job mobility has increased dramatically in recent years. Under such circumstances, pronunciation will evolve more quickly, making it harder to define a univocal pattern for describing what a ‘Peking accent’ sounds today.

The old-fashioned pronunciation, consecrated by *xiàngshengr* crosstalk masters as Hóu Bǎolín and Guō Qǐrú,¹ is no longer the only one used by the locals; indeed, what many call ‘Pekingese’ now is a largely stereotypical relic, confined to local comedy and a small number of (very) aged speakers. From a phonemic and phonetic point of view, traditional Pekingese largely coincides with Mandarin. The paraphonic settings, instead, are remarkably different: for example, Hóu Bǎolín’s Pekingese-accented *pǔtōnghuà* sounded high-pitched and crisp, with clearly uttered tones and a rapid flow; this is precisely how proper *xiàngshengr* delivery style is supposed to be. By contrast, Hóu’s stooge, Guō Qǐrú, would intentionally speak with a less refined –yet understandable– pronunciation, which was decidedly more on the side of the *dialect* proper, while interacting with Hóu, to convey the theatrical idea of a *naïve* character.

Video-sharing websites, like *YouTube* and its Chinese *alter ego*, *Yōukù*, provide quite a rich selection of *xiàngshengr* performed by Hóu and Guō, which is highly advisable to carefully listen to, in comparison with neutral voices from contemporary Chinese courses audio files. Also the recordings attached to the old Linguaphone Institute *Chinese Course* offer abundant examples of Peking-accented pronunciation.

Therefore, and contrary to popular belief, the ‘traditional’ pronunciation described above has little to do with the ‘lazy mutter’ most often –but unjustly– associated with speakers from the *hútònggr* alleys.² This is infamously known for its anarchical use of *érhuà* (see G2), and the frequent reduction –if not ‘annihilation’!– of constrictive and stop-strictive contoids. Foreigners, as well as many Chinese themselves, would typically find it nearly unintelligible, but also fascinating, and would

even try to imitate it, in order to ‘sound local’. We cannot help discouraging such a practice: true locals will hardly appreciate a bad imitation of something that they consider exclusively their own, while people from other parts of China could wonder why Westerners should waste their time learning how to *mispronounce* the Common Language.

8.3. Upon suggestions from our informants, we managed to determine three examples of what can be honestly targeted as a ‘modern’ Peking accent (again, not *dialect*). First of all, Mr Jiāng Wén (1963-), a well known movie star and director; then, Ms Yáng Lán (1968-), a successful entrepreneur and talk show hostess; finally, Mr Zhāng Bīn (1969-), a sport anchorman from CCTV5.

We chose to select them for three good reasons: (1) all of them are about fifty years old, that is the most representative age range for an accent (with ‘mature’ people, although not old; then, all things considered, they are still ‘young’ speakers, but free from temporary phonic habits so typical of real young people); (2) their pronunciation is recognisably Pekingese, but still within a flawless spoken *pǔtōnghuà*, being the classical example of a ‘mediatic pronunciation’ of the capital city; (3) all of them are nationally well-known, therefore one can easily find plentiful audio and video material on the Web, ranging from plain interviews to the presentation of nation-wide programs.

Anyone interested in gaining an insight into the differences between a (‘clean’ Pekingese) mediatic *accent* and a proper Pekingese *dialect* (too marked to be accepted as *pǔtōnghuà*) can easily succeed. It is sufficient to compare the pronunciation of the three above-mentioned people with that of Mr Jiāng Kūn (1950), a master of *xiàngsheng*, who, on the contrary, habitually speaks using the real Pekingese dialect (even if it is no longer that of Hóu Bǎolín).

Fundamental traits

8.4. Impressionistically, contemporary Peking-accented Mandarin sounds rather annoyingly ‘buzzing’ in comparison with its neutral counterpart. However, that does not imply an overall tenser articulation. On the contrary, our recordings, coupled with articulatory kinaesthesia, suggest a *laxer* articulation of both vocoids and contoids, mainly as a result of the same reduction and attenuation phenomena described in § 1.5.1-3, 3.4.2-3, and 3.7: primarily, /_ots(h), _os/ [z] (& [j]) and /_otʂ(h), _oʂ/ [ʒ], in unstressed syllables with ‘zero’ tonemes, not in an intoneme. This is not sur-

prising, and actually happens in more colloquial varieties of English, too, as anyone will easily notice, by comparing the language spoken daily among relatives, friends and colleagues, with that adopted in dubbing foreign movies and documentaries.

The typical *buzz* accompanying Pekingese Mandarin is rather a consequence of the behaviour of /((Ç)w/ and *érhuà*. Depending on the speaker and the elocution speed, [w̥] may feature a noticeably *stronger lateral contraction* than in neutral pronunciation; frequently, it is realised as a postalveolar intense approximant [ʝ], possibly ‘coloured’ by the addition of a certain degree of lateral contraction (representable –if really needed– with ‘[ʝw̥]’); finally, a more marked variant is [wʝ] (with a laxer and less advisable alternative, [wʝ̥]).

The massive use of *érhuà* is perhaps the most stereotypical characteristic of Peking-accented Mandarin. However, so far we have not found two single speakers from Peking who would totally agree on which words should take the ‘extra’ /-ʝ/ (ie one not strictly requested in neutral pronunciation) and which ones should not. If a clear rule existed, we might treat the behaviour of *érhuà* in Pekingese pronunciation not simply as a morpho-lexical phenomenon; but, until solid evidence of a certain ‘logic’ is found, we prefer not to reproduce here a boring list of words, as those invariably presented in self-styled ‘serious’ essay.³ As far as our field of study is concerned, the mere frequency of *érhuà* is of little importance, compared to its true phonetic nature, which shows oscillations similar to that just described for /((Ç)w/. Therefore (beside neutral or nearly neutral [ŷ]), [Vʝ, Vʝw̥] are also possible (including their nasalized counterparts).

Modern Pekingese also features a distinctive realisation of /[#]w/ as [v], a labiodental approximant (or [v], a semi-constrictive), except in /[#]wɔ/: *wànwéiwǎng* ‘World Wide Web’ *p*[\u0280\u0280\u0280\u0280\u0280\u0280, -\u0280\u0280-] (with further variants for /Ei/; see below), but *tā wèn wǒ* ‘he/she asks me’ *p*[\u0280\u0280\u0280\u0280\u0280\u0280\u0280\u0280, -w\u0280\u0280(\u0280)] (obviously, other stress patterns are possible). Sometimes, but not systematically, partial labialization can occur: [v̥, V]; yet, [v] can be safely employed in a ‘broad Pekingese’ kind of normalized transcription.

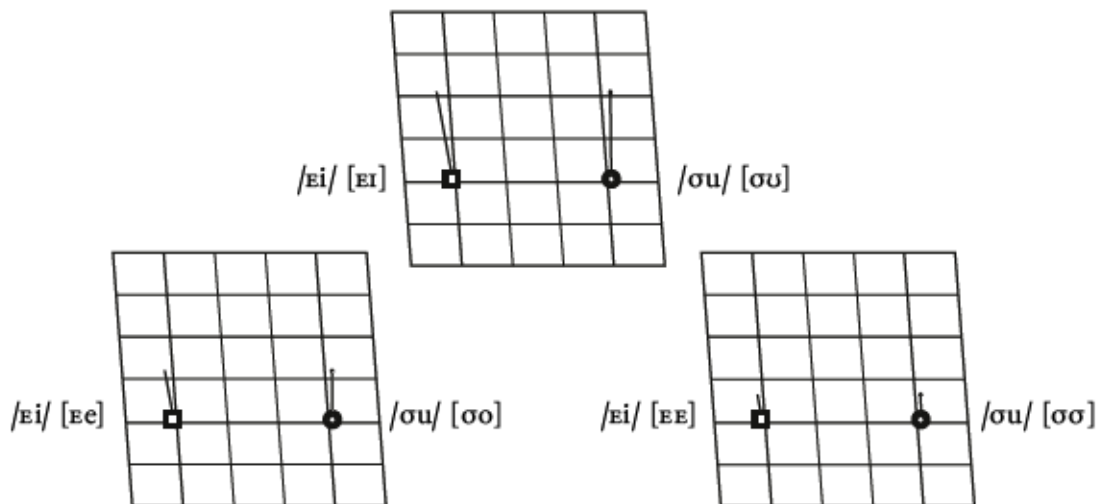
Such variants are gradually gaining ground among youngsters, while being far less frequent, if not virtually absent, among most elders. In their early attempts to learn Chinese, foreign students whose native language (including English) treats /w/ and /v/ as distinct phonemes are easily puzzled by that ‘v-sound that should not be there’. Some of them will actually use their own [v] instead, in search of a more

genuine pronunciation, only to be mildly reprimanded by their teacher, who may reinterpret a fully constrictive [v] as a badly pronounced /f/. Therefore, unless one is able to distinguish [v] from [f] (as many Dutch speakers are, for example), /w/ [w] is the best and only choice for teaching purposes.

Another trend –popular among young Peking speakers– involves articulating /ei, ou/ as narrower diphthongs: [ɛɛ, ɔɔ] (fig 8.1). In quicker speech and in preintonemes, even monotimbric realizations are possible, [ɛɛ, ɔɔ] (particularly in /o\$/).

As far as stress and intonation are concerned, there are no significant differences with respect to neutral Chinese. However, a relatively large number of common words ending with a marked toneme, /(\$)\$\$/, feature a ‘zero’ toneme in Peking-accented Chinese: *pángxié* ‘crab’ [‘pʰan˥˥xiɛ], *wéiqún* ‘apron’ [‘uɛi˥˥tɕyn, ‘uɛɛ], *nángua* ‘pumpkin’ [‘nɑŋ˥˥gwa]; these coexist with their ‘normal’ realizations, as presented in most dictionaries: *pángxiè* [‘pʰan˥˥xiɛ], *wéiqún* [‘uɛi˥˥tɕhyn, ‘uɛɛ], *nánguā* [‘nɑŋ˥˥kwa]. Similarly, post-verbal monosyllabic pronouns *wǒ*, *nǐ*, *tā* are more systematically reduced to /o\$/ than in neutral pronunciation (unless emphasis is added): *wǒ wèn nǐ* ‘I ask you’ [ˌwɔʃ˥˥vɜ̃n˥˥nɿ], *gēn tā zǒu* [‘gɛn˥˥ta˥˥tɕɔʃ˥˥, -tɕɔʃ˥˥].

fig 8.1. Diphthong reductions in Pekingesé accent, especially among younger generations.



NOTES:

¹ *Xiàngshengr* is a traditional Pekingese comic performance, featuring jokes and funny dialogues arranged in a crosstalk pattern. It requires an exceptional mastering of the language – traditionally, the Peking dialect, rich in puns and popular jargon– and a good deal of dry wit.

Hóu Bǎolín (1917-1993) is still remembered as an outstanding *xiàngshengr* performer. He was most revered for his creativity and ability to play with language, but also for his high-profile approach to *xiàngshengr*: he actively promoted the use of *pǔtōnghuà* in performing modern crosstalk, both to spread the Common Language throughout China (thanks to radio and –later– tv broadcasting), and to make a quintessentially Pekingese art become popular across the country and beyond.

Always assisted by his lifetime friend and colleague Guō Qǐrú (1900-1969), he even dedicated some famous *xiàngshengr* to the relationship between dialects and what he believed should be the ‘national’ language of all Chinese. Such famous pieces of crosstalk as *Pǔtōnghuà yǔ fāngyán* [Common Language and Dialects], and especially *Běijīnghuà* [The Peking Dialect], remain as valuable testimonies of how the delicate matter of language evolved in China in the past decades, besides being a precious collection of authentic Pekingese pronunciation.

² The *hútònggr* are ancient narrow streets, or alleys, most commonly associated with the oldest and most traditional parts of Peking. In that city, *hútònggr* are formed by several lines of *sìhéyüànr*, the traditional Pekingese square courtyard residences; many neighbourhoods have been formed by joining one *sìhéyüànr* to another to form a *hútònggr*, and then joining one *hútònggr* to another. The word *hútònggr* is also used to refer to such neighbourhoods.

Since the mid-20th century, the number of *hútònggr* has dropped dramatically in Peking, as many of them have been demolished to make way for new roads and buildings. More recently, some *hútònggr* were designated as protected areas, in an attempt to preserve this aspect of Chinese cultural history.

³ cf also LU, Y-ZH.(2001) *Qingsheng he erhua*, p. 151-189; and HE, P. (2006) *Chinese Pronunciation Course*, p. 159-162.

9. ‘*Praateries*’

Usefulness of computer-aided analysis applied to phonetics and tonetics

9.1. As far as our field of study is concerned, and coherently with the ‘natural’ approach that we chose to adopt, no particular equipment is absolutely necessary, apart from a good recording apparatus (whether a classic tape-recorder, or a more modern digital device) and a high quality headset. Our ears and brain will do the rest. That does not mean that speech analysing softwares –such as the easily available *Praat*–¹ are useless to ‘natural’ phoneticians; just the opposite is true, but the user must be aware of the technical limitations that all similar computer-aided research protocols –so far– have been unable to overcome.

The bottom line is that instrumental data will bear fruit only when supported by, and submitted to, the phonetician’s own scientific reasoning, which –in turn– should be mediated by the attentive listening of well selected audio samples. It must be remembered that machines will record and analyse *everything*, without previously filtering out what a human being would spontaneously discard, for being fortuitous and therefore not representative, such as background noises, or voice distortions due to contingent factors. Yet, such a remarkable degree of (unnecessary) comprehensiveness, paradoxically, fails to provide a full account of what a diligent phonetician really needs.

As a blatant demonstration of our point, we invite anyone to download the latest version of *Praat* from Boersma and Weenink’s webpage. We have chosen *Praat* precisely because it is readily available and is free of charge, in addition to being constantly updated both by the authors and by a large community of users worldwide. Once installed, load the software and select an appropriate audio file, for example a .WAV clip (note that .MP3’s are not supported) from any Chinese course at hand. A wide range of choices are available to the user, including the possibility of ‘drawing’ vectorial images for successive editing (the three illustrations enclosed hereby precisely originate as *Praat*-generated .EPS vectors).²

9.2. With little effort, it is easy to determine where the vocal chords are vibrating and how intensely –within certain limits– making it possible to tell fully voiced phones from fully voiceless ones. But, as soon as one tries to analyse mixed phonation contoids, or partially devoiced vocoids (among others), what appears on the screen is often very elusive and ambiguous.

In a similar way, constrictive articulations are not hard to find, as they produce a unique, densely ‘indented’ contour. However, it is practically impossible to determine the exact place of articulation of a given contoid, apart from very loose distinctions, such as ‘front’ (like [s]) vs ‘back’ (like [x]). Nor is it possible to univocally distinguish ‘slit’ and ‘grooved’ contoids, or certain constrictives and their corresponding approximants, which even untrained persons may be able to recognise on the spot. Interestingly, *Praat* appears to be less reliable precisely where the human ear does its job with more ease!

The only reasonable way to detect the characteristics of a given contoid through *Praat* and other ‘*Praat*-esque’ softwares, would be to compare the tracings and diagrams extractable from the informant’s recording with those from ‘specimen’ audio files, *ie* audio samples of (generally single) phones whose articulation is certain and ‘certified’. Typically, a ‘specimen’ should feature the voice of a neutral speaker, preferably a professionally trained one, like a dubbing voice actor; or, the phonetician’s own voice, provided the articulation is performed after careful phonokinaesthesia.

Even in that case, however, instrumental phoneticians cannot confidently assert ‘this phone is...’, but only ‘based on *what I see on the screen*, this phone *looks* more or less similar to...’. At the end of the day, not an exceptional performance, unless the ultimate evidence is provided: an X-ray of the phonatory organs while the phone is being uttered, or a sequence of photographs –if not an X-ray film!– for dynamic articulations, *eg* to distinguish [i:] and [j:].

Now, it goes without saying that such a purely instrumental approach would be ‘mathematically correct’, for being detached from the human element –which is subjective and relatively fallible– but nothing proves that it could also be resolute, in terms of scientific reliability; or, of any practical use, for example to those who develop an interest in phonetics to improve their language learning skills. Even more importantly, the instrumental approach is simply not viable, when a phonetic survey is to be conducted on a large basis of informants, and an expensive and

cumbersome X-ray machine... happens not to be near at hand! Not to mention the non-secondary fact that the amount of radiation to be administered to each informant would be more than enough to cause severe consequences.

9.3. With regard to vocoids, *Praat* clearly shows their *quantity* (: duration), obviously in terms of seconds, or fractions of a second, according to a mental category readily understandable by human beings: time. But, as far as vocalic *quality* is concerned, the software can simply report a number of ‘representative frequency values’, or *formants*, expressed in hertz (: cycles per second), which –on the contrary– humans cannot measure naturally.³ It is the coexistence of the right formants *as a whole* that allows the hearer to disambiguate vowels; or, in the case of language teachers, to tell whether their students have learned how to pronounce the vowel sounds which do not exist in their mother tongue(s), or continue to employ their native phones as a substitute. In this respect, mere frequency values are of little use.

From the point of view of a ‘natural’ phonetician who wants to prepare the *canIPA* vocogram, or the *offIPA* trapezium, of a given language (or just of a certain speaker), computer-measured formants may be worth the effort of being collected and analysed only if the same microphone and the same recording device are invariably used, without altering any of the settings, and always under the same environmental circumstances (humidity, pressure, sound insulation degree &c). Considering how variable the situation can be, when a phonetician records different informants, or when they produce an audio file by themselves, and then send it to the researcher, anyone can easily imagine why we are far less enthusiast than ‘acoustic’ phoneticians are about dealing with those ‘cryptical’ Hz values.

It must be remembered that the same vocoid, eg [a], can be produced by different formant combinations, depending on the speaker’s sex, age and physical characteristics, too; this happens even within the same speaker, depending on whether one is whispering, or roaring out loud in a stentorian way, since also voice intensity plays a certain role. Finally, of course, vocalic formants are –slightly– influenced by surrounding contours, too, which makes the scenario even more intricate. Quite enough to look at the ‘prodigies’ of computer-aided phonetics with a bit of prudent scepticism.

Therefore, before recording a pronunciation test, every informant should provide some basic data to ‘calibrate’ the software, so that it is possible to treat the formant

values relevant to a single utterance not as an absolute, mathematically measurable item unrelated one another, but as something comparable with a general ‘grid’, where both a five-year-old boy’s [a] and that pronounced by his grandmother can be collocated according to a universal criterion.

Such preliminary data should include, at least, a perfect rendition of the four extreme vocoids [i, æ, α, ɯ], and the corresponding rounded ones, [y, œ, ɔ, u], whose formant values will constitute the fundamental ranges which all other vocoids shall be compared to. Unfortunately, apart from being a tedious and time-consuming process, that would require the uncommon blessing of finding a non-phonetician informant who perfectly knows what to do and how, to articulate unfamiliar vocoids, including the theoretical ones not attested in any living language, eg [ɯ, œ]. The inevitable question is, should we really bother going through so much trouble to meet a machine’s needs, when our ears and brains can do the job for us just fine? As long as voice-analysing softwares lack the ability to automatically accommodate to different voices without the tiresome pre-setting procedures described above, the ‘natural’ approach will remain the winning choice.

9.4. In our view, a software like *Praat* really offers phoneticians a helping hand only in studying tonality, which is very intuitively represented by a series of irregular segments, or –when these happen to be tightly connected– by a solid line. The pitch contours relevant to complete sentences, if considered as a whole, can be a good hint to the intonational patterns of non-tonal languages; while the segments corresponding to single syllables reveal how taxotones behave in tonal languages.

In Chinese, the two layers of intonation and tonemicity are tightly intertwined, which makes it difficult to interpret *Praat*-generated tracings, if one does not know what to look for in the first place, and how. But, no matter how careful and rigorous one can be, yet there is little hope to truly understand Chinese intonation merely by means of instrumental measurements; even more so, for those who actually have to *learn* it, to improve their pronunciation of the language. To provide a clear example of our point, we have extracted the tonal contour of six short sentences submitted to a reliable informant from Peking, and recorded in a nearly sound-proof environment (of course, the recordings can be made available to anyone who may be interested in double-checking what we report hereby).

The original tracings produced by *Praat* were somehow too ‘coarse’ to be used as such, as in some parts no tonal contour appeared at all (mainly due to an excessive

laryngealization), while in others the same contour was clearly incoherent with what could be heard from the earphones. Consequently, fig 9.1-2-3 must be considered as the result of a patient and meticulous editing process, whose length and error-proneness *in se* should be exasperating enough to warn wannabe instrumental phoneticians against any form of ‘easy optimism’.

All four fundamental intonation units, / . , ; ?/, are duly represented. We intentionally selected sentences where the first toneme abounds, given that the slightly deformations produced by intonation are far easier to observe, when the toneme’s theoretical profile is a horizontal one. Please, be noted that tonality is indicated by the solid thick line, while the dotted line represents intensity (which must be considered, due to its influence on tonality, at least at the perceptive level).

The grey areas suggest the corresponding intonational bands: the criterion we followed is the ‘grapho-arithmetic mean’ of the actual height of the various taxotones, compared to the height that we would expect to find in /;/, when the tonal band is clearly expanded, and the four tonemes are realised with their ‘fullest’ taxotones, ie those typically used at the beginning of all Chinese courses.

9.5. The six sentences are presented hereby in *pīnyīn* romanization first, then in *canIPA* transcription, and are followed by their English translation:

Mā dānxīn jiā-bān ma?

/ǐ̇ma ˋtanˋsinˋtsjaˋpanˋmaˋ?/

[ǐ̇ma ˋtʰɑŋˋˈt͡ɕiŋˋˈd͡ʒjaˋpʰɑmˋma.]

‘Is Mom worried about working overtime?’

Mā hěn dānxīn jiā-bān.

/ma ˋhɛnˋdānxīnˋjiāˋbānˋ./

[ma ˋhɛŋˋˈtʰɑŋˋˈt͡ɕiŋˋˈd͡ʒjaˋpʰɑnˋ.]

‘Mom is very worried about working overtime.’

Tā shì nǐ māma ma?

/ǐ̇thəˋʃuˋniˋmaˋmaˋmaˋ?/

[thəˋʃuˋˋnɿˋmaˋmaˋmɿˋ.]

‘Is she your mother?’

Tā bú shì wǒ māma?

/thə ˋpʉˋʃuˋwɔˋmaˋmaˋ./

[thə ˋpʉˋʃuˋˋwɔˋmaˋmaˋ.]

‘She is not my mother.’

Nǐ kàn nǎ-bě'n shū?

/ɿni\kʰan ˌna.pɕn̄˥˩su/

[ɿni\kʰaŋ ˈna.b̚ɕŋ̚˥˩su]

‘Which book do you read?’

Nǐ ài kàn nǎ-bě'n shū jiù kàn nǎ-bě'n shū.

/ˌni\ʔaɛ\kʰan ˌna.pɕn̄˥˩su;| \tɕjɔu\kʰan ˌna.pɕn̄˥˩su./

[n̄˥˩ʔaɛ\kʰaŋ ˈna.b̚ɕŋ̚˥˩su;| ˈtɕjɔu\kʰaŋ ˈna.b̚ɕŋ̚˥˩su]

‘Read whatever book you like.’

fig 9.1.

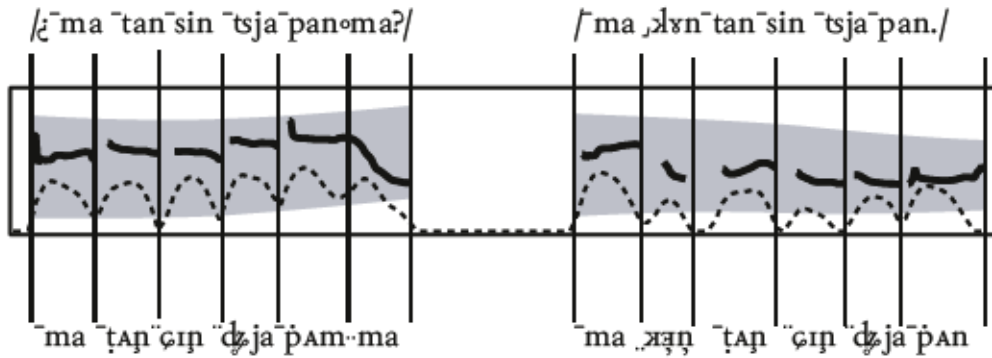


fig 9.2.

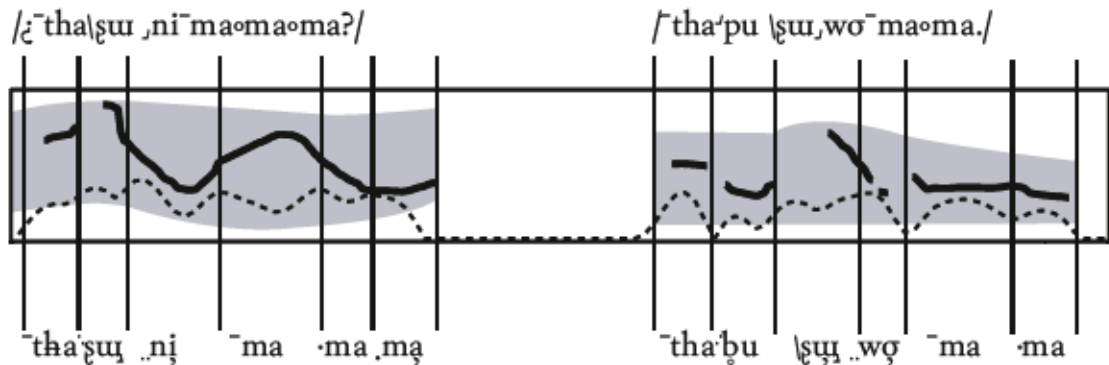
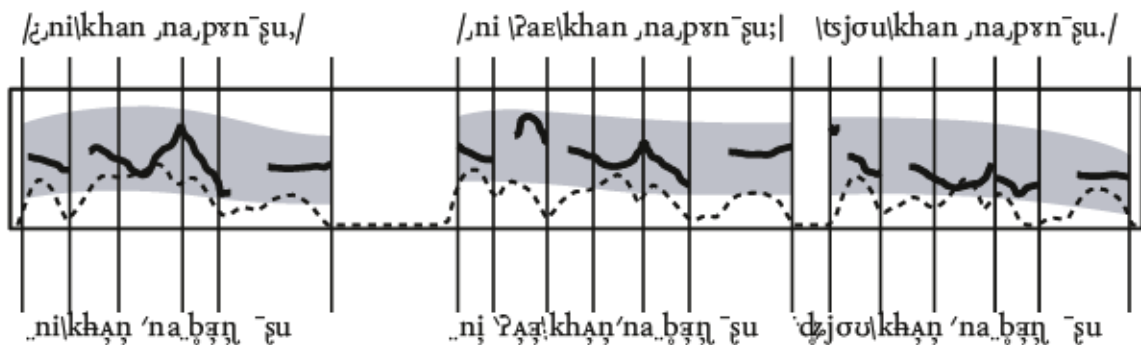


fig 9.3.



NOTES:

¹ *Praat* (also the Dutch word for ‘talk’) is a free software program for the analysis of speech in phonetics. It has been designed and continuously developed by Paul Boersma and David Weenink of the University of Amsterdam. It can run on a wide range of operating systems, including various Unix versions, Mac OSX and Microsoft Windows (95, 98, NT4, ME, 2000, XP, Vista). The program also supports speech synthesis, including articulatory synthesis.

² In this case, we used *Praat* 5.2.03.

³ For a detailed explanation of how formants are generated, and to produce the distinct ‘timbre’ of the various vocoids, please refer to a reliable book or, with extreme caution, to a similar website.

10. Conclusions

We will not try to summarise in a few lines the considerable amount of data presented in this work, as every aspect of Mandarin Chinese pronunciation has been described in details, but we also did our best to avoid any unnecessary rigmarole. The purpose of these conclusions is highlighting the main reasons why –in our view– this analysis is worth reading.

Phonology in general

Though *pīnyīn* is, all in all, a fair representation of the essential components of *pǔtōnghuà* pronunciation, yet it cannot be taken as a true *phonemic* transcription. It is surprising, then, that the matter is not addressed effectively by most of the analyses that we have consulted. Many linguists feel content with a simplified kind of *phonetic* transcription, whose *offIPA* symbols are so generic that –apparently– a separate phonemic transcription would be almost superfluous, also because the *-emic* and the *-etic* layers are often mixed up.

That is why we regarded it important to draw a clear line between them, and devised a thoroughly meditated *phonotonic* scheme. In it, we made it clear that the ‘aspirated consonants’ are simply */Ch/ sequences*, and there is no real need for an extra ‘*/tʰ, tʰh, ʃ/*’ phonemic set, as */ts, tsh, s/* can safely include the taxophones represented by *j, q, x* in *pīnyīn*. Instead, the ‘mythological apical vowels’ have been re-conducted to their genuinely vocalic nature, by adopting ‘*/u/*’, irrespective of the numerous taxophonic variants that may occur. Where possible, a realistic symbol has been employed, instead of a generic one (as the *offIPA* practice recommends): thus, */ɛ, ʒ, ʌ/*, instead of something like ‘*/e, □, x/*’.

Vowels and consonants

We chose to include *eight* phonemes in the vocalic system of Mandarin Chinese, as that makes it easier and more natural to manage the numerous taxophones required in neutral pronunciation. Diphthongs have been analysed with full realism, rejecting traditional renditions like ‘*/ei, ai, au, ou/*’ even at the phonemic level.

We are particularly proud of § 2, where the *érhuà* phenomenon is treated in great depth; once again, avoiding superficiality and abstractness. Instead, a set of *laterally contracted vocoids* has been introduced, finally explaining why *érhuà* does not sound exactly like the ‘American *r*’, as many phoneticians still seem to think.

Our treatment of Mandarin contoids accepts no compromise, and clearly indicates –for the very first time, as far as we know– three stress-related consonantal degrees for stops, stop-strictives and constrictives. This three-level system can imply a partial or total voicing of the contoid; and, for /Ch/, a difference in ‘aspiration’ strength. Other linguists have had a similar intuition, too, but sadly never implemented it in a truly systematic way, as we have done in this work.

Tonemes, tones and intonation

Instead of focusing only on consonants and vowels, we devised an effective notational system, to indicate tonemes and intonemes as well. At the phonotonic level, our transcription is even capable of indicating stress, stress-related taxotones and important intonation details *at the same time*, without jeopardising readability and typographical elegance. It may *look* complex, of course, but that is simply because what it represents *is* complex. However, it is not complicated, nor visually chaotic, since each symbol of our inventory univocally corresponds to a single element that is indispensable for a neutral pronunciation, and has been carefully designed not to interfere with other symbols.

Suggested pronunciation

Most people learning Chinese are not phoneticians, and do not actually need –or want– to acquire a genuinely neutral pronunciation. The average student would be quite happy with a *reasonably* simplified scheme, but not as generic as that generally presented in Chinese handbooks. For this specific purpose, we have provided a *suggested pronunciation* which, while renouncing some phonetic subtleties, yet features no peculiarities that an average Chinese native speaker would consider regional, or even typical of a foreign accent.

As a final statement, we hope that this work can provide a useful contribution for a better understanding of the pronunciation of Mandarin Chinese, whose importance as a global language will consistently increase in the near future.

Bibliography

- BOUQUIAUX, L. & CLOAREC-HEISS, F. & THOMAS, J.M.C. (1976) *Initiation à la phonétique*. Paris: PUF/ORSTOM (a vinyl record to be used alone or in reference to THOMAS & BOUQUIAUX & CLOAREC-HEISS; expanded IPA).
- CANEPARI, L. (1983) *Phonetic Notation · La notazione fonetica*. Venezia: Cafoscarina (with 2 enclosed audiocassettes; almost ^{can}IPA).
- CANEPARI, L. (2000) *Dizionario di pronuncia italiana* ['A Dictionary of Italian Pronunciation']. Bologna: Zanichelli (^{can}IPA).
- CANEPARI, L. (2004) *Manuale di pronuncia italiana* ['A Handbook of Italian Pronunciation']. Bologna: Zanichelli (with 2 enclosed audiocassettes; ^{can}IPA).
- CANEPARI, L. (2006) *Avviamento alla fonetica* ['Breakthrough in Phonetics']. Torino: Einaudi (^{can}IPA).
- CANEPARI, L. (2007) *Natural Phonetics & Tonetics. Articulatory, auditory, functional*. München: Lincom (^{can}IPA).
- CANEPARI, L. (2007²) *A Handbook of Pronunciation: English, Italian, French, German, Spanish, Portuguese, Russian, Arabic, Hindi, Chinese, Japanese, Esperanto*. München: Lincom (^{can}IPA).
- CANEPARI, L. (2010) *The Pronunciation of English around the World*. München: Lincom (^{can}IPA).
- CANEPARI, L. (2010²) *English Pronunciation*S. Roma: Aracne (^{can}IPA).
- CANEPARI, L. & CERINI, M. (2009) *Pronuncia cinese per italiani* ['Chinese Pronunciation for Italians'] (^{can}IPA).
- CANEPARI, L. & CERINI, M. (2010) 'Mandarin: the *r* grammeme and the so-called *érhuà* phenomenon' (downloadable from the ^{can}IPA *Natural Phonetics* website: <http://venus.unive.it/canipa/>; ^{can}IPA).
- CAO, W. (2000) *Hanyu fayin yu jiuyin* ['Chinese Pronunciation and Orthoepy']. Peking: Peking University Press (simplified IPA).
- CAO, W. (2002) *Hanyu yuyin jiaocheng* ['A Course in Chinese Phonetics']. Peking: Beijing Yuyan Daxue Chubanshe (almost IPA).
- CAO, W. (2008) *Chinese Pronunciation Practice*. Peking: Peking Univ. Press (IPA).
- CATFORD, J.C. (1977) *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press (IPA).

- CATFORD, J.C. (1988) *A Practical Introduction to Phonetics*. Oxford: Clarendon Press (IPA).
- CHAO, Y-R. (1948) *Mandarin Primer*. Cambridge (MASS.): Harvard University Press (non-IPA).
- CHENG, C-C. (1973) *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton (almost IPA).
- Contemporary Chinese Dictionary, The (Chinese-English Edition, 2002)* Peking: Commercial Press.
- DEFRANCIS, J. (2003) *ABC Chinese-English Comprehensive Dictionary*. Honolulu: University of Hawai‘i Press.
- DELATTRE, P. et alii (1951) *vwajel sētetik a dō fōrmāt e vwajel kardinal*, in ‘Le Maître Phonétique’ (IPA).
- DOW, F. D-M. (1972) *An Introduction to the Pronunciation of Chinese*. Edinburgh: Edinburgh University Press (IPA).
- DOW, F. D-M. (1972) *An Outline of Mandarin Phonetics*. Canberra: Australian National University Press (IPA).
- Duden Aussprachewörterbuch* (2005⁶, 1962¹) Mannheim: Bibliographisches Institut & F. A. Brockhaus AG (IPA).
- HAUDRICOURT, A.G. & THOMAS, J.M.C. (1976) *La notation des langues. Phonétique et phonologie*. Paris: Institut Géographique National (with 2 enclosed vinyl records; IPA).
- HE, P. (2006) *Chinese Pronunciation Course*. Peking: Peking University Press (with 2 CD’s; IPA, but not systematically).
- HUANG, B-R. & LIAO, X-D. (2002) *Xiandai Hanyu* [‘Modern Chinese’]. Peking: Gaodeng Jiaoyu Chubanshe.
- HYMAN, L.M. (1975) *Phonology: theory and Analysis*. New York: Holt, Rinehart & Winston (non-IPA).
- JONES, D. (1956) *Cardinal Vowels*. London: Linguaphone Institute (2 [78 rpm] records with booklet; now partially downloadable by surfing on the Internet; IPA).
- JONES, D. (1967³) *The Phoneme: its Nature and Use*. Cambridge: Heffer (IPA).
- KRATOCHVÍL, P. (1968) *The Chinese Language Today*. London: Hutchinson (almost IPA).
- LADEFOGED, P. & MADDIESON, I. (1996) *The Sounds of the World’s Languages*. Oxford: Blackwell (almost IPA).

- LAVER, J. (1980) *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press (with a non-enclosed audiocassette; IPA).
- LAVER, J. (1994) *Principles of Phonetics*. Cambridge: C. Univ. Press (IPA).
- LI, M. & SHI, P-W. (1986, repr. 2004) *Hanyu putonghua yuyin bianzheng* ['The Correct Distinction of the Common Chinese Language Phones']. Peking: Beijing Yuyan Daxue Chubanshe (almost IPA).
- LIN, Y-H. (2007) *The Sounds of Chinese*. Cambridge: Cambridge University Press (IPA, but almost exclusively as a phonemic counterpart of *pīnyīn*).
- LU, Y-ZH. (2001) *Qingsheng he erhua* ['Zero Toneme and *érhuà*']. Peking: Commercial Press.
- MIONI, A. M. (1973) *Fonematica contrastiva*. Bologna: Patròn (IPA).
- MIONI, A. M. (2001) *Elementi di fonetica*. Padova: Unipress (IPA).
- NORMAN, J. (1988) *Chinese*. Cambridge: Cambridge University Press (IPA).
- O'CONNOR, J.D. (1973) *Phonetics*. Harmondsworth: Penguin (IPA).
- Practical Chinese Reader* (1985) Peking: The Commercial Press (almost IPA).
- QIN, Z. (1980) *On Chinese Phonetics*. Peking: The Commercial Press (almost IPA).
- SCURFIELD, E. (1991) *Teach Yourself Chinese*. Sevenoaks: Hodder & Stoughton (with an audiocassette).
- SHAN PAO LI, J. (1975) *Introduction to the Chinese Language*. Taipei: Ke Ji Publishing Co (IPA).
- SMALLEY, W.A. (1964²) *Manual of Articulatory Phonetics*. Terrytown (NY): Practical Anthropology (with 33 non-enclosed [18 cm, 19 cm/s] reels, lasting 32 hours; non-IPA).
- SPESHNEV, N.A. (1980) *Фонетика китайского языка*. Ленинград: Издательство Ленинградского Университета (hybrid IPA).
- THOMAS, J.M.C. & BOUQUIAUX, L. & CLOAREC-HEISS, F. (1976) *Initiation à la phonétique*. Paris: PUF (completed by the vinyl record of Bouquiaux & Cloarec-Heiss & Thomas; expanded IPA).
- TRUBECKOJ, N.S. (1939) *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht – (1969) *Principles of Phonology*. Berkeley: University of California Press (non-IPA).
- ZHU, CH. (1997) *Waiguo xuesheng Hanyu yuyin xuexi duici* ['Counter-measures in the Study of Chinese Phonetics by Foreign Students']. Peking: Yuwen Chubanshe (scattered use of IPA).