Luciano Canepari & Marco Cerini (2017²)

Chinese Pronunciation & Accents

Geo-social Applications of the Natural Phonetics & Tonetics Method

6.
Chinese vowels & diphthongs

Vowels

6.1. After thorough consideration, we have chosen to include *eight* vowels in our analysis of Mandarin phonology: /i, y, ɯ, u; ɛ, ɤ, ɔ; a/ (with 21 taxophones, as we will soon see). In ƒig 6.1-A, the various vocoids and their taxophones –*ie* contextual allophones– are grouped by phonemes; while, in ƒig 6.1-B, we present the same taxophones arranged in 'functional' sets: /V#, Vn, VN/. Let us start by systematically considering each phoneme and its most typical realizations in neutral pronunciation.

6.1.1. First, we find the close (high) front phoneme *i* /i/ (but notice that we use *ï* for /ɯ/, as we will see below, cf § 6.3 & following): /(#j)i, (#j)in, (#j)iN/ [(#j)i, (#j)ɪn, (#j)ɪN] *(C/y)i*, *(C/y)in*, *(C/y)ing*: [¯ji] /ˊji/ *yī*, [ˌn̩i̩i̩] /ˌni/ *nǐ*, [ˈjɪn] /ˈjin/ *yín*, [¯tɕɪn] /ˇtsin/ *jīn*, [\jɻɴ] /\jiN/ *yìng*, [¯tɕɪN] /ˇtsiN/ *jīng*.

Too many Chinese authors, in teaching beginners the fundamentals of Chinese pronunciation, avoid *IPA*, which they use in an extremely poor way, unsatisfactorily derived from *pīnyīn*. On the contrary, a good phonic transcription is the most useful way to actually teach and learn Chinese pronunciation (as for any other language), because it is not a purely decorative embellishment, 'just to make it look better'. Quite the opposite, in fact.

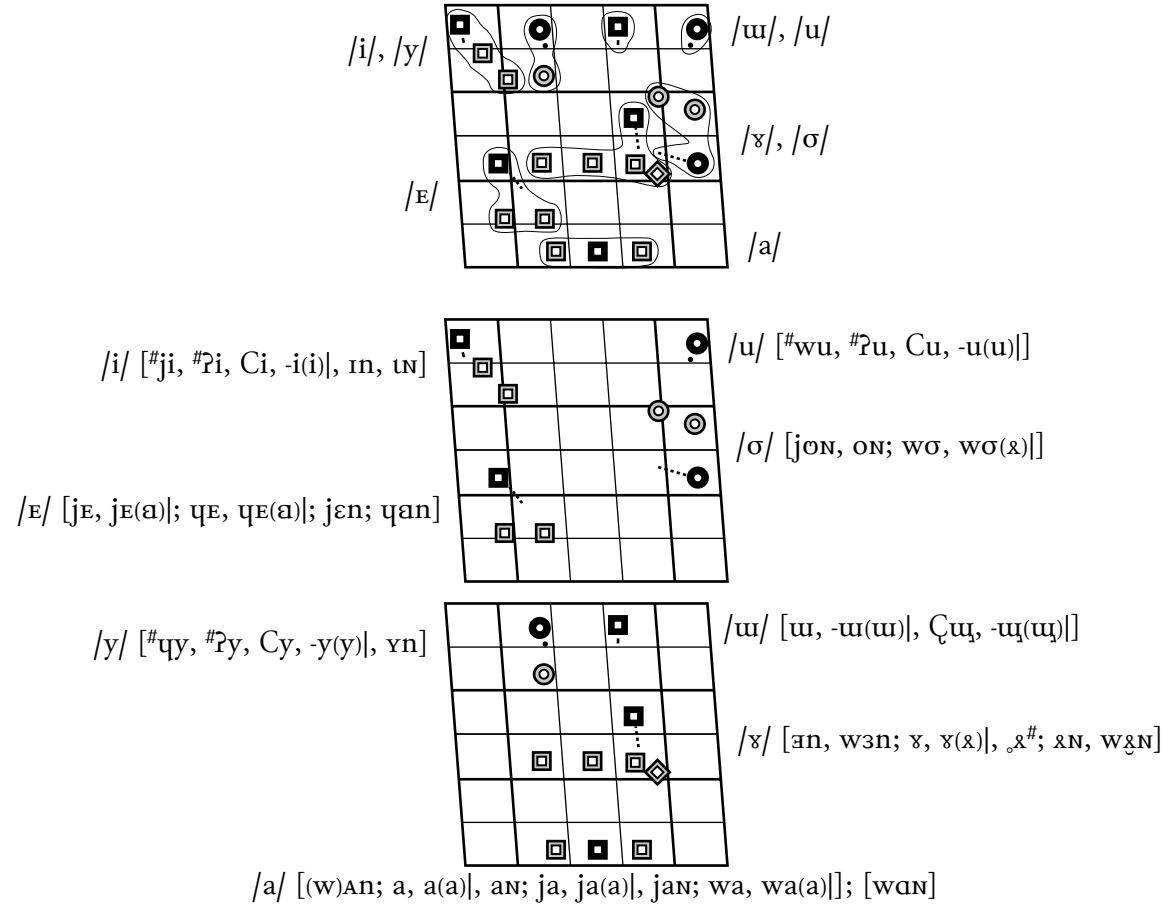
6.1.2. Secondly, there is the close (high) front-central rounded phoneme *ü* /y/, with a fairly limited distribution: /(#ɥ)y, (#ɥ)yn/ [(#ɥ)y, (#ɥ)yn].

The official orthography adopts *ü* only in *nü* and *lü*, while leaving *u* unmarked in *ju(n)*, *qu(n)*, *xu(n)*, *yu(n)*. We see no advantage in doing so and prefer to keep a friendly *ü* anywhere: *jū* /ˇtsɥ/ [¯tɕy], *qǔ* /ˌtsɧɣ/ [ˌtɕʰɣ̌], *xù* /\sɥ/ [\ɕɣ̀], *yǔ* /ˌɥɣ/ [ˌɥɣ̌], *jūn* /ˇtsyn/ [¯tɕyn], *qǘn* /ˊtsɧyn/ [ˊtɕʰyn], *xùn* /\syn/ [\ɕn̩], *yǘn* /ˊɥyn/ [ˊɥyn].

More conveniently, *pīnyīn* could have used *zü*, *cü*, *sü*, *yü*, by phonemically interpreting these syllables as we do: /tsy, tsɧy, sy/ [tɕy, tɕʰy, ɕy].

6.1.3. In addition, we find a back rounded vowel, *u* /u/ ([u] not to be confused with British English '/uː/' –or, better, /uu/– whose various taxophones are rather back-central, and often diphthongized): *gūdú* /ˇkuˊtu/ [¨ɡ̊uˊtu], *wǔrǔ* /ˌwuˌʐu/ [˙wuˌʐu̯].

ƒig 6.1-A. Mandarin fundamental vowels and taxophones, grouped by phoneme (*cf* ƒig 6.1-B).



6.1.4. Following, we have the open (low) vowel *a* /a/, with its distributions and various taxophones: /(C)a#, (w/C)an, (j)aɴ, waɴ/ [(C)a, (w/C)ɑn, (j)ɑɴ, wɑɴ]. A few examples: *āza* /¯atsa/ [¯ʔa◦dza, ···d̥za], *tà* /\tha/ [\thą], *ǎn* /˩an/ [˩ʔą̈ą̈ɳ], *pán* /ˊphan/ [ˊphɑn], *páng* /ˊphaɴ/ [ˊphɑn], *wǎn* /˩wan/ [˩wą̈ą̈ɳ], *yáng* /ˊjaɴ/ [ˊjɑn], *wǎng* /˩waɴ/ [˩wɑɑɴ].

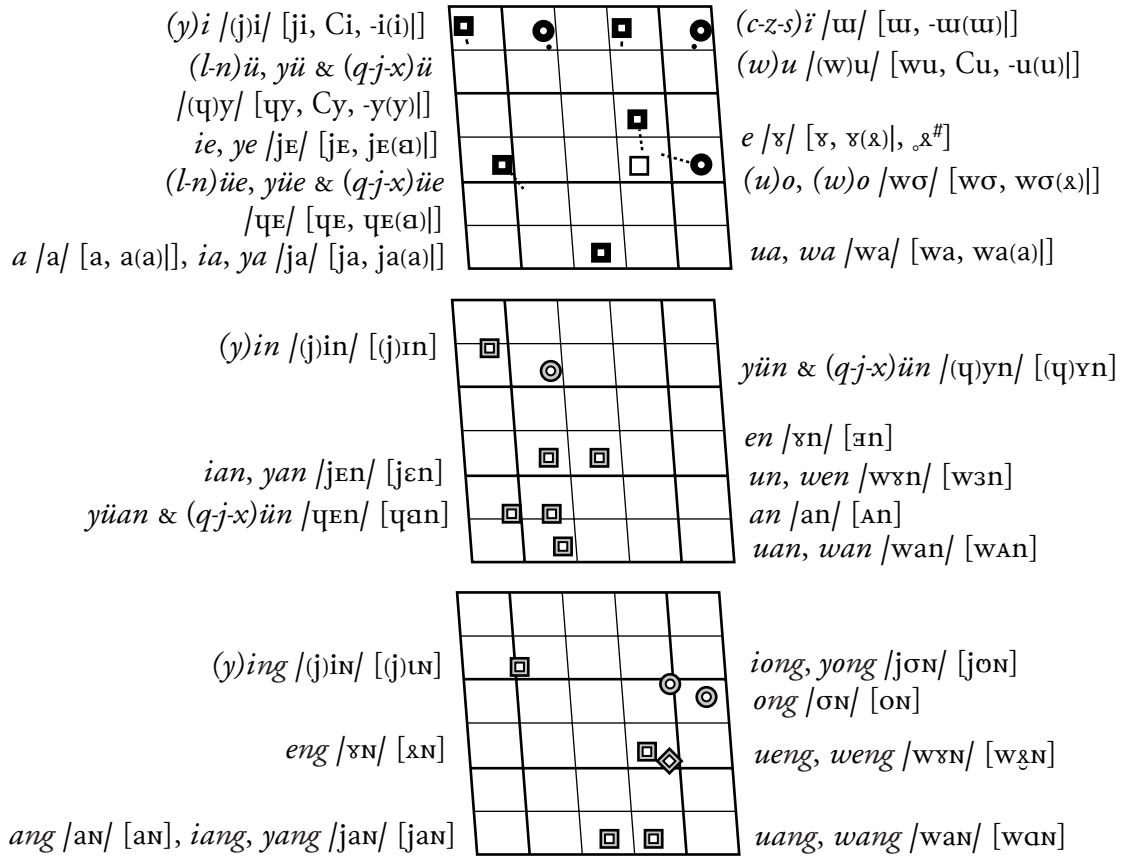
As far as /‖jan, ɥan‖ are concerned, we think they are better represented by /jɛn, ɥɛn/ (in spite of the official rendering with *-an*): *yán* /ˊjɛn/ [ˊjɛn], *diǎn* /˩tjɛn/ [˩tj̈ɛ̈ɳ], *yüàn* /\ɥɛn/ [\ɥą̈ɳ], *qüán* /ˊtshɥɛn/ [ˊtɕʰɥɑn].

6.1.5. The vocalic inventory of neutral Chinese continues with three mid vowel phonemes (lower-mid): *e* /ɛ, ɤ/, *o* /ɔ/. 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#, ɥɛ#/ [jɛ, ɥɛ].

Pīnyīn writes *ye*, when lacking a graphic initial; otherwise, *ie*: *jiě*, *qiē*, *xié*, *niè*, *liè* /˩tsjɛ, ¯tshjɛ, ˊsjɛ, \njɛ, \ljɛ/ [˩tɕj̈ɛ̈, ¯tɕʰjɛ, ˊɕjɛ, \ɲjɛ̈, \ljɛ̈], and so on. Any combination is permitted, except *fie* and *zhie*, *chie*, *shie*, *rie* (since the Chinese 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 of the Peking Opera –*Jīngjù* /¯tsiɴ\tsy/ [¨d̥ʑɪɴ\tɕʏ̌]– often chant their lines using a form of stage pronunciation which still retains archaic syllables, like *zie* and *siang*.)

ƒig 6.1-B. Mandarin fundamental vowels and taxophones, grouped by syllables: /V#, Vn, VN/.



For the latter sequence, we have *jue*, *que*, *xue* and *yue*, but *lüe* and *nüe* (no other combinations occur, even those which might be articulated without problems, like *müe*). But, more friendly, as with with *ü(n)*, 'our own' *pīnyīn* will always show *üe*, instead: *jüě*, *qüē*, *xüé*, *nüè*, *lüè* /ˌtsɥɛ, ˉtsʰɥɛ, ˊsɥɛ, \nɥɛ, \lɥɛ/ [ˌtɕɥɛ̞̍, ˉtɕʰɥɛ, ˊɕɥɛ, \ɳɥɛ̞̍, \lɥɛ̞̍].

6.1.6. The back-central vowel *e* /ɤ/ plays an important role, in giving Mandarin its distinctive sound, and shows the following distributions: *e*, *er*, *en*, *wen*/*Cun*, *eng*, *weng* /(#ʔ)ɤ#, (#ʔ)ɤʐ; (#ʔ)ɤn, wɤn; (#ʔ)ɤɴ, wɤɴ/ [(#ʔ)ɤ(ɤ|), (#ʔ)ɤɻ̩; (#ʔ)ən, wɜn; (#ʔ)ʌɴ, wɻ̩ɴ]. A few examples: *kēkè* /ˉkhɤ\khɤ/ [ˈˈkhɤ\kʰɻ̩̍(ɻ̩|)], *Zhōu Ēnlái* /ˉʈʂoʊˉʔɤnˊlaɛ/ [ˉʈ̟ʂoʊˈˈʔənˊlʌɜ, -ˈˈʔəlˊlʌɜ], *érqiě* /ˊɤʐˌtɕʰjɛ̞̍/ [ˈˈʔɤɻ̩ˌtɕʰjɛ̞̍], *gēnběn* /ˉkɤnˌpɤn/ [ˈˈg̊əmˌp̰ə̤n̰], *Sūn Wén* /ˉswɤnˊwɤn/ [ˈˈswɜnˊwɜn], *fēngshēng* /ˉfɤɴˉʂɤɴ/ [ˈˈfʌɴˉʂʌɴ], *fùwēng* /\fuˉwɤɴ/ [\fu̘ˉwɻ̩ɴ].

6.1.7. Finally, *o* /(#ʔ)o/ [(#ʔ)o(ɤ|)], which can occur in the interjections *o!*, *yo!* /ʔo, jo/ [o(ɤ|), jo(ɤ|)], or can be preceded by /w/ in the sequences *mo*, *bo*, *po*, *fo*, *wo*/*Cuo* /mwo, pwo, phwo, (C)wo/ [m̂o(ɤ|), p̂o(ɤ|), p̂ʰo(ɤ|), f̂o(ɤ|), (C)wo(ɤ|)]. A few examples: *mómò* /ˊmwo\mwo/ [ˈm̂o\m̂o̞̍(ɻ̩|)], *luòtuo* /\lwoˌthwo/ [ˈl̥o̞ˊp̂ʰo(ɤ|)] /ˊpwoˊphwo/ *bópó*, [\lwo̞̍.two̞̍(ɻ̩|)], *zuò cuò* /\tswo\tsʰwo/ [ˈd̥zwo\tsʰwo̞̍(ɻ̩|)], *hángyō* /ˊxlaɴˉjo/ [ˈɤɑɴˉjo], *ò!* /\ʔo/ [\ʔo̞̍(ɻ̩|)].

As far as *(i)ong* is concerned, a traditional approach favors '/(j)uɴ/', and we con-

sider it as perfectly legitimate (along with '/ɥuɴ/', for mediatic taxophones). Nonetheless, we have observed that [oɴ] and [jɵɴ] are by far the most typical neutral realizations, and the phonemic status bestowed on /ʊ/ suggests us to switch to the more realistic /(j)ʊɴ/ (because of its taxophones). Two examples will suffice: *gōng-gòng* /˥kʊɴ\kʊɴ/ [˙˙g̊oɴ\k̩o̩ɴ̩], *qióngjiǒng* /˦tsɦjʊɴˏtsjʊɴ/ [˙tɕhjɵɴˏtɕjo̩o̩ɴ̩].

Phonemic conciseness *vs* realism: the interphonemic choice

6.2. We are perfectly aware that from an intraphonemic point of view, *ie* just within the (Mandarin) Chinese language, [ɤ, ɘ, ɜ, ʌ, ʌ̰] and [ᴇ] could –or should– be considered as taxophones of the same phoneme. Besides, in contact with labials (*mo*, *bo*, *po*, *fo*, *ou*, *wo*/*Cuo*), [ʊ] could be considered as a variant of /ɤ/ too, as many Chinese phonologists do (*offIPA* '/wɤ/' or '/wə/'). 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 booklet on Mandarin Pronunciation for Italians (2010) as a practical *passe-partout* for the various occurrences of /ɤ/: [ɜ] (while keeping [ᴇ, ʊ] apart, as /ᴇ, ʊ/, for the reasons given above).

We are confident in our interphonemic choice to list /ᴇ, ʊ/ 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.

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

Accordingly, our /jᴇ, ɥᴇ, ᴇi; wʊ, (j)ʊu, (j)ʊɴ/ should appear as '/jə, ɥə, əi; wə, (j)əu, (j)wəɴ/' (or '/(j)uɴ/', for the last one, as previously noted). In other words, our /ᴇ, ʊ/ would be treated as taxophon(em)es of the ubiquitous '/ə/', when in contact with /i, j/ or /u, w/. Inevitably, /aᴇ, aʊ/ 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 (unlike *e!*). In our analysis, we do not have to think twice: /ᴇ, ʊ/ (usually: *ề!*, *ò!* /\ᴇ, \ʊ/ [\ᴇ̩, \ʊ̩], but Chinese dictionaries permit any toneme for both, with great paraphonic liberty, including possible additional timbres, [\e̩, \ɛ̩; \o̩, \ɔ̩]).

The point is that interjections and onomatopoeia are written down in Chinese characters too, as any other lexemes, which may lead to think that they can be lexicalized and phonemicized accordingly. This is debatable, since 'unorthodox'

phones appear sporadically in any languages, precisely in interjections and onomatopoeia, but are not generally included in their phonemic inventory.

6.2.2. A glaring example: in Italian, laughter is represented by *ah ah ah!* (or something like that); the basic pronunciation should be something like [ˈˈ(h)aː haˈhaˑ], but we can actually hear anything ranging between [h, ɦ, ɧ, ɧ, x, ɬ; ɑ, ʌ, a, ɐ, ɜ, ɑ, ʌ, ɤ], just to list a few. None of these belong to a 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 [aiˈfaˑi], while [haiˈfaˑi] 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 (as a stylistic phoneme, for paraphonic purposes), without actually being a fully functional element of Italian phonology (as in Canepari [2004, particularly Ch 2-6], or, more concisely [2007², Ch 3: *Italian*]).

6.2.3. Should we want to stick to the '/i, y; ə, a; ɯ, u/' scheme rigidly, there would be little choice but retaining /(ɛ, ɔ)/ as *incidental phonemes –ie paraphonemes–* in Chinese, like /(v)/ (which, more exactly, is rather a *xenophoneme*). To make it even more captious, one may go as far as treating [ɛ, ɔ] as *incidental monophthongizations* of '/əi, ai/' and '/əu, au/' 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 practical and easy to use.

We prefer, instead, to regard '/ə/' as an *archiphoneme* and to split it into three different *semi-phonemes* (so to say): /ɛ, ɤ, ɔ/. It has to be underlined that these are not the same as *incidental phonemes*: they are true phonemes, but simply not 'crucial' ones. In fact, native speakers may perceive no difference between [ɔʊ] and [ɤʊ], for instance, or [ʌɜ, ʌə] and [ʌɪ], as well as between [uɴ, ʊɴ] and [oɴ, ɔɴ]; 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 official symbols at the phonemic level; and, according to the most adamant traditionalists, not merely official, 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, ɛ, ɛ, ɜ, ɤ], &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 /ɛ/ [ɛ], for which official /ɛ/ would be better than /e/ (but it is finally, and happily, gaining popularity among scholars).l

6.2.4. In some analyses of Chinese, we find sober compromises like '/e, ɤ, o/' or, better, '/ɛ, ɤ, ɔ/', but also masterpieces of schizophrenia, such as '/e, ɤ, ɔ/' or '/ɛ,

ɤ, o/' – and even '/ɛ, ə, ɤ, o/', mixing phones and phonemes! This is highly confusing (not only) 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 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 /ˈɛ/ is [ɛ] as well, in those languages; neutral German even distinguishes between '/ˈeː, ˈɛː, ˈɛ/' [eː, ɛː, ɛ]). 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 between symbols.

6.2.5. 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, Spanish, and Japanese) /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, ɛ, ɐ], '/e/' makes it necessary to revive the unrealistic '/jan, ɥan/', to explain *yan*, *yüan* [jɛn, ɥɑn].

Multilingual applicability is the other advantage of the *interphonemic* approach, in that it helps polyglots (or those who aim at being such), 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, ɛ] 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* /ˈσː/ is [ˈσː] in modern Standard British, but [ˈo·o] in Mediatic British, and [ˈɔː] in most American accents. To an untrained ear accustomed to hear and speak only American English, [ˈo·o] may easily suggest *owe* /ˈσω/ [ˈσ·ω] at first, but once the listeners are able to detect [o] in isolation, they would promptly recognise a mono-timbric diphthong like [ˈo·o], and distinguish it from [ˈσ·ω] (also guided by the meaning of actual words).

On the phonemic status of *-ï*

6.3. 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. But, in our view, both belong to the same phoneme.

We strongly support '/ɯ/' as the phonemicization of what *pīnyīn* transcribes as *i* after *z*, *c*, *s* and *zh*, *ch*, *sh*, *r*. As explained below –cf Ch 9: *Chinese Consonants*– we adopt *ï* instead, in order to mark, even at the graphemic level, and in parallel to /ɯ/, that we are dealing with another fundamental component of Mandarin phonology, *not* with some mysterious taxophones of /i/.

The first author of this book is firmly determined to state that, unfortunately, there are 'phonologists' who posit such senseless and unreasonable 'rules', which follow the absurd Soviet 'phonemic' principle by which a consonant can determine the change in quality of a following vowel, instead of the opposite. This happens when the obvious two Russian phonemes /i, ɨ/ are schizophrenically mystified as a unique and unnatural '/i/' entity. On the contrary, it is more natural and clearly simple that it is a vowel that can change the quality of a preceding consonant.

Thus, instead of '/sa, si, ʂi, Ci/', for [sa, sɯ, ʂɰ, Ci] (pīnyīn: *sa*, *xi*, *shi*, *Ci*), we posit /sa/ and /Ci/, but /sɯ, ʂɯ/, by a clear and natural assimilation process – and, accordingly, we write: *sa*, *sï*, *shï*, *Ci*. Our choice falls upon /ɯ/, 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).

6.3.1. Many –if not all– Sinologists favor two symbols not enclosed in the *offIPA* inventory: ɿ, ʅ; the former to represent *ï* in *zï*, *cï*, *sï*, the latter, *ï* in *zhï*, *chï*, *shï*, *rï*. Though called 'vowels', both are often presented as 'syllabic constrictives' (although they are approximants; and note that we prefer 'intense' to 'syllabic', which is less ambiguous), 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. But, the prevailing theory is that they are simply some queer taxophones of /i/).

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

6.3.2. 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 [ɹ̩]), 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, [ɯ], *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*, [ɹ̩] (while [ɹ̩], seen above, is a plain intense approximant), may result *auditorily* quite similar to [ɯ], but the *articulatory* nature of this phone is primarily vocalic.

To sceptical &/or traditionalist readers, we suggest a simple experiment, provided they can pronounce Chinese at least decently: articulate a long, slow and emphatic *sī̈*, paying attention to any transition from *s* to *ī̈*. If this is done correctly, the result should resemble something like ⟦¯sːz̥zźδ̩ɗ̩ɯː⟧. (In this time of majestic technological achievements, many instrumental, 'acoustic', phoneticians underestimate the importance of kinaesthesia and proprioceptive abilities. To those interested in 'redeeming' themselves, we may suggest Catford, 1988).

This being said and recognized, 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') ⟦z̥, z, z̤̥, ź, δ, ɗ⟧ and to keep only [ɯ] as the real distinctive element of *sï* [sɯ], which we therefore phonemicize as /sɯ/ (with /ts(ɧ)ɯ/, too), clearly discarding 'ɿ' and similar questionable expedients.

6.3.3. A slightly different criterion must be applied to /Çɯ/, *ie* when /ɯ/ follows the post-alveolar initials: *zhï*, *chï*, *shï*, *rï*. In this case, an automatic assimilation affects the entire articulation of /ɯ/, and cannot be regarded as a merely transitional event: even an untrained ear can detect a continuous 'buzz', incomplete –or even absent– in /ɯ/; auditorily, the latter has a very different timbre too, even if the (medium-)dorsum generally remains in the same 'box' as [ɯ].

Much debated is the true nature of these /Çɯ/ sequences. As we explain hereby, neutral pronunciation actually admits more than just one realization: it is surprizing 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, some Chinese authors often refer to how 'Peking speakers' would pronounce something, as if the *pǔtōnghuà* spoken in Peking were to be assumed as a model for neutral pronunciation (rather than a kind of mediatic Mandarin accent, when not too broad, as in everyday life). 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 even by foreigners.

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 'ʅ' is still employed describe it as something of the same nature as *ɚ*, 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. Thus, 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 to regard it as 'a less fricative [ʐ̻]'. Provided 'less fricative' does not stand for a *semi-constrictive* phone (^{can}[ɻ̞]), we can safely infer that an *intense post-alveolar approximant* is actually meant (^{can}[ɻ̟]), in accordance with some officially sanctioned descriptions.

6.3.4. 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: [ʐ̺].

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 [ʂ], and they are willing to 'speak proper': in doing so, they literally *retroflect* their apex, instead of raising it gently behind the alveolar ridge, 'just to be on the safer side'. The resulting phones [ʈʂ, 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 [ɻ̟] into a sequence like [ɯɻ] (or worse, [ɯʐ], even elongated in [ɯʐː], which might be misunderstood as /ʁʐ/). 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 [ɻ], these speakers would actually use any one of their most familiar *r* phones, namely [ɹ, ɾ, ɻ, ɽ] &c (with their intense counterparts, [ɹ̟, ɾ̟, ɻ̟, ɽ̟] &c, for /ɻ̟/). These contoids act quite well as surrogates, but produce a fairly recognisable accent: those who aim at a more convincing pronunciation, then, should choose [ɻ̟]. (It might be 'illuminating' to read Ꞵ 26: The 'whole truth' on English *r*, in Canepari's *English Pronunciation & Accents*).

6.3.5. 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 [ɯˌ], which –for the sake of conciseness– we shall describe here as a *laterally contracted* [ɯ]. 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 Ꞵ 7, 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 favor [ɯˌ] is both of an *articulatory* and *euphonic* nature. By comparing the orograms of [ɯˌ] and [ɻ] (*fig* 7.1 & *fig* 9.5), it is quite clear that this last phone requires little –if any– movement of the apex, yet the auditory result is not very different from that of [ɻ̟]. 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(ɧ), ʦ(ɧ), s, l, n/ &c).

Subjective and questionable as it can be, euphony certainly does play a role, even if [ɯ̹] were not selected on purpose, nor were 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 [ɯ̹] suggests no far-fetched 'stage pronunciation' at all, like –for example– [r, ɾ] 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 [ɯ̹] instead of [ɻ̟] (or should alternate between them).

There is a certain timbric difference between [ɯ] and [ɯ̹], though mostly as a side effect of the co-articulation phenomena occurring in the latter phone. We have lateral contraction, in the first place, primarily responsible for the characteristic 'buzz' typical of any [ɣ] (laterally contracted vocoids). Thus, its post--alveolarity, which plays just a marginal role). Apart from this, [ɯ] and [ɯ̹] share almost the same dorsum position.

6.3.6. Nonetheless, we have noted a mediatic more advanced taxophone ([ɨ̟], *cf* *fig* 13.2), not very common, but adopted even by some professional voices, as well: a few examples of it can be heard at the beginning of the Chinese dubbing of *Hiroshima mon amour* (Alain Resnais, 1959), which, indeed, may sound somewhat awkward to most native listeners. But, this phone is more a mediatic one (*cf* *fig* 13.2).

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. Kinaesthesia reveals that it is possible to articulate a fully functional [ɣ] as front as in the area of [ɨ̟], but not fronter than that, otherwise it would be very hard both to contract the sides of the tongue and, at the same time, raise the apex.

Diphthongs

6.4. Modern Mandarin phonology only has four diphthongs (*fig* 6.2): /ɛi, aɛ, aɔ, ɔu/ [ɛɪ, aɜ, ɑɒ, ɔʊ]. The former two may be preceded by /w/, the latter two by /j/, with slightly different vocalic traits: /wɛi, waɛ, jaɔ, jɔu/ [wəɪ, waɜ, jaɒ, jɔʊ]. 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/' (or '/uej, uaj, iaw, iow/').

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 rather reasonable.

When *wai* and *yao* follow another initial, their spelling changes into *uai* and *iao*,

making grapho-syllabification easier in some cases: *chēnguài* ‘to blame, to rebuke’ is unmistakably /˥ʈʂʅɤn\kwaᴇ/, while *chéngwài* ‘outside of a town/city’ leads automatically to /˦ʈʂʅɤɴ\waᴇ/; otherwise, *pīnyīn* would abound in apostrophes or hyphens, like older romanization devices: **chēn'guài* vs **chéng'uài*, or **chēn'gwài* vs **chéng'wài*.

The choice to use *ng*, instead of plain *ng*, in our Chinese examples helps, too (and *sh*, *ch*, *zh*, as well – although we do not do the same for English, or for Chinese words used without their tonemes).

6.4.1. Nonetheless, we find absurd inconsistencies: instead of changing *wei*, *you* into *uei*, *iou*, the official orthography prescribes *ui*, *iu*, with no real advantages for the learner (unless someone prefers to use a mediatic accent, cf Ch 13). However, in all honesty, this is odd even at the typographical level.

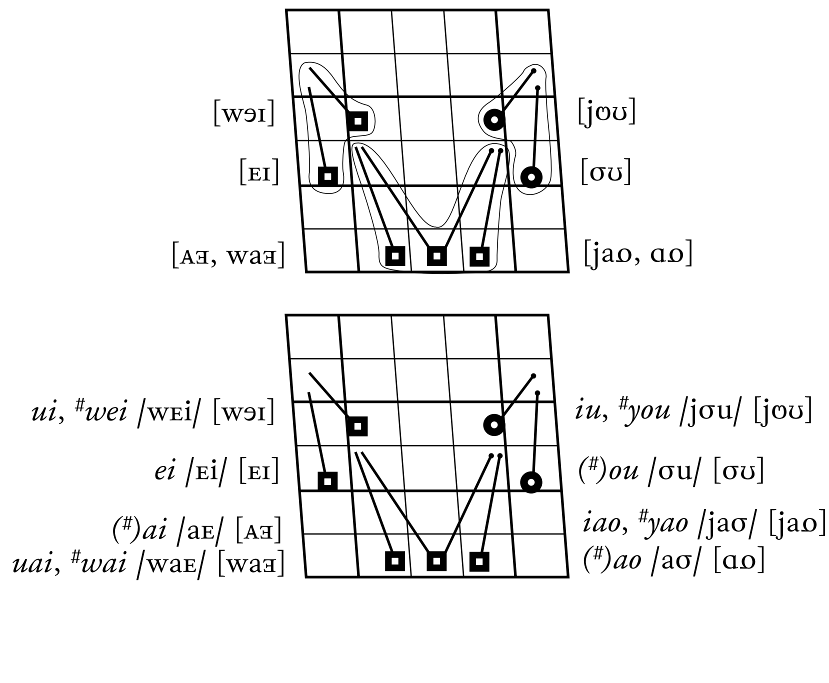
Lacking their real *nuclei*, we are forced to put the relevant diacritics on what we must consider their syllabic *codae*, instead, ie *uī*, *iú*, *uǐ*, *iù* – rather than the more friendly and more correct (as more neutral) ‘*uēi*, *ióu*, *uěi*, *iòu*’!

Besides, *ui* and *iu* easily lead foreign students to think that /wᴇi, jɤu/ 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 strongly *pīnyīn* influences their ‘understanding’ of Mandarin phonology.

To prove their point, they would refer to mediatic variants, where –nonetheless– we do not hear anything like ‘[wi, ju]’, but rather, in mediatic pronunciation, [wɩi, jɷu] with tonemes 1 & 2, but [wɜe, jɵo] with tonemes 3 & 4 – fig 13.1). As a matter of fact, we have: *neutral* pronunciation [˭wɘɪ, ˦jɷʊ], and [˩wɘ̣ɪ̣, \jɵ̣ʊ̣], but *mediatic* pronunciation [˭wɩi, ˦jɷu], and [˩wɜ̣ẹ, \jọọ].

A few neutral examples: *fēi* /˥fᴇi/ [˭fᴇɪ], *duī* /˥twᴇi/ [˭ṭwɘɪ], *ǎi* /˩aᴇ/ [˩ʔᴀ̣ᴀ̣ɜ̣], *wài* /\waᴇ/ [\wạɜ̣], *ǎo* /˩ao/ [˩ʔɑ̣ɑ̣ọ], *hǎo* /˩ʞaɤ/ [˩ʞɑ̣ɑ̣ọ], *yáo* /˦jaɤ/ [˦jaɵ], *Ōu* /˥ɤu/ [˭ʔɤʊ], *mǒu* /˩mɤu/ [˩mɤ̣ɤ̣ʊ̣], *diū* /˥tjɤu/ [˭ṭjɷʊ].

fig 6.2. Mandarin diphthongs.



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 considering.

Vowel reduction & devoicing

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

6.5.1. fig 6.3 shows the most important reductions which *may* occur in fast &/or casual speech, and in connection with the so-called *qīngshēng*, /˳$/, 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 a tune (for the first variants), up to modifications produced within a sentence, in a pre-tune (for the other variants). For /Vn, Vɴ/, 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 [Vɴ] merge into [Ṽ].

Anyway, in order not to make our transcriptions more complex than necessary, we will not employ such reductions ordinarily. On the one hand, they do not occur automatically, nor universally; on the other hand, 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).

In a more colloquial version of fast speech, even the consonants may be more attenuated, in unstressed syllables, /˳$/. In fact, we can also find: /p, t, k/ [b, d, g] → [ß, ꞡ, ɣ], /ts(j), ʈʂ/ [dz, dʑ; dʐ] → [ᵈz, ᵈʑ; ᵈʐ]; /f; s, ɕ, ʂ/ [v̥; z̥, ʑ̊, ʐ̊] → [v, z, ʑ, ʐ] (while /h/ [ḫ], generally, does not change).

6.5.2. As can be seen, [ɪ] applies to /i(n/ɴ), y(n)/: *dìdi* [\ṭi̹..d̥i̹; -dɪ̹], *guānxi* [¯ḳwᴀn̹-..ɕi; -ʑ̊ɪ̹], *xiàqü* [\ɕja̹..tɕhy̓; -tɕi̹, -tɕɪ̹]; and [ɯ], to /ɯ, u/ (with [ɰ] for /Çɯ/): *yìsï* [\ji̹-..z̥ɯ̹; -z̥ɯ̹], *rènshï* [\ʐə̹ɳ̹..ʂɰ̹; -ʐ̊ɰ̹], *shūfu* [¯ʂu..fu; -v̥ɯ].

In addition, we have [ɘ] in /jᴇ, ɥᴇ/, [ə] in /(w)ɤn/, and [əɩ] in /(w)ᴇi/: *jiějie* [˓tɕ̇jᴇ̹ᴇ̹-¨d̥ʑ̷ᴊᴇ; ˙dʑ̷ᴊᴇ, -˙dʑ̷ᴊɘ], *àiren* [\ʔᴀ̹ɜ̹..ʐ̹ɜn̹; -ʐ̹ɜn̹, -ʐ̹ən̹], *mèimei* [\mᴇ̹ɪ̹..mᴇ̹ɪ̹, -mə̹ɪ̹]; furthermore, [ɤ], for /ɤ, (w)ɤɴ, (w)ʊ/, and [(j)ɤɷ], for /(j)ʊu/: *gēge* [¯k̇ɤ··g̊ɤ; --gɤ], *dǐngzhe* [_ṭɪ̹ɴ̹¨d̥ʐɤ; -˙dʐɤ],

fig 6.3. Vowel reduction in /◦$/ (in fast speech, not in a tune; whereas, in a tune, the distinct timbers of fig 6.1-2 are used). Note that even the approximants are reduced, becoming semi--approximants: [j, ɥ, w] → [ʝ, ɥ̞, ɯ̞].

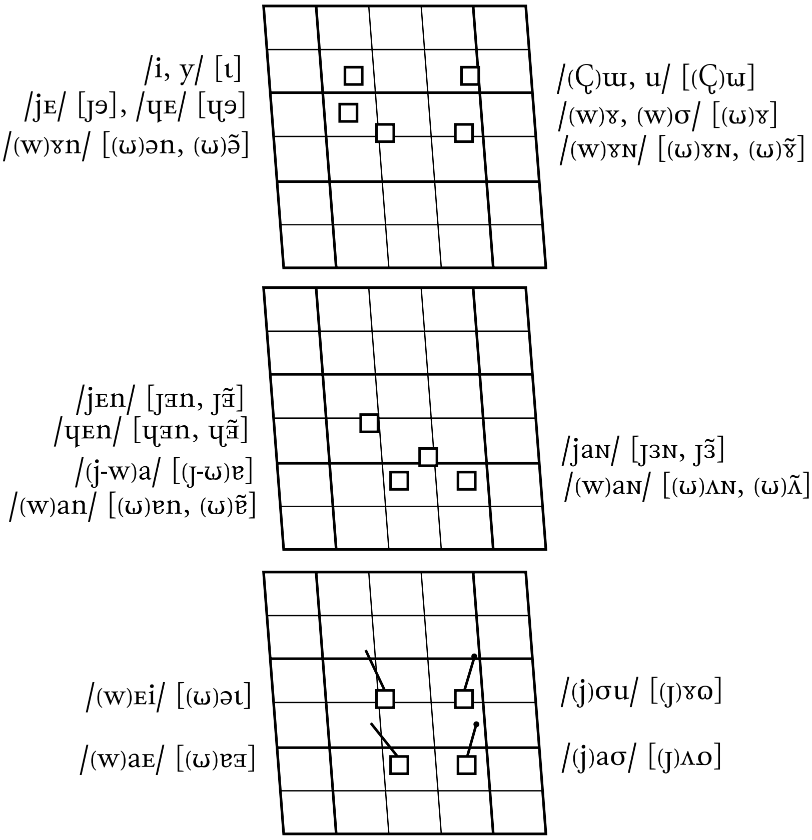
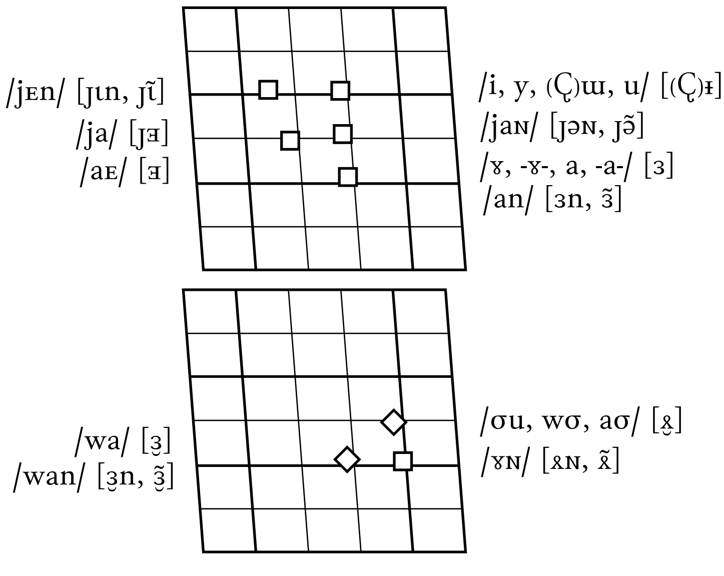


fig 6.4. Further reductions in syllables with 'zero tones' (in even quicker speech, not in a tune; whereas, in a tune, the realizations of fig 6.3 are possible, too). Let us also notice [wV] → [V̫].



xiānsheng [¯ɕjɛŋ··ʂɤɴ; ··ʐ̊ɤɴ], *jièmo* [\tɕjɛ̤..m̤ɔ; ·m̤ɤ], *nuǎnhuo* [_nwʌ̤ɲ̤¨ʜwɔ; ·ŋ̊ɦʊɤ], *yòngtou* [\jɔ̤ŋ̤..thɔ̤ʊ̤; ·tɤ̤ɵ̤], *péngyou* [ʼphɤɴ··jɔʊ; ··jɤɵ].

Lastly, we have: [ɐ; ɐɜ] in /(j-w)a, an; (w)aᴇ/: *bàba* [\pa̤..ba̤; ·bɐ̤], *māma* [¯ma··ma; ··mɐ], *xǐhuan* [_ɕi̤i̤¨ʜwan; ·ɦʊɐn], *tàitai* [\thʌ̤ɜ̤..thʌ̤ɜ̤; ·tɐ̤ɜ̤]; and [ʌ; ʌɒ] in /(j)aɴ; (j)aɒ/: *wǎnshang* [ˌwʌ̤ʌ̤ŋ̤¨ʂaɴ; ·ʐ̊ʌɴ], *zhīdao* [¯ʈʂɻ̩··dɑɒ; ··dʌɒ].

6.5.3. In even quicker speech, and not in a tune, further reductions are possible (fig 6.4): [ɪ] for /i, y, ɯ, u/: *dìdi* [\ti̤..di̤; ·dɪ, ·dɪ̤], *guānxi* [¯kwʌɲ··ɕi; ··ʑ̊ɪ, ··ʑ̊ɪ], *xiàqü* [\ɕja̤..tɕhẙ; ·tɕi̥, ·tɕɪ, ·tɕɪ̥], *yìsï* [\ji̤..zɯ̤; ·zɯ̥, ·zɪ̥], *rènshï* [\ʐɜ̤ŋ̤..ʂɯ̤; ·ʐ̊ɯ̥, ·ʐ̊ɪ̥], *shūfu* [¯ʂu··fu; ··vɯ, ··vɪ̥]; and [ɤ̥] for /ɔu, wɔ, aɔ/: *nuǎnhuo* [_nwʌ̤ɲ̤¨ʜwɔ; ·ŋ̊ɦʊɤ, ·ŋ̊ɦɤ̥], *yòngtou* [\jɔ̤ŋ̤..thɔ̤ʊ̤; ·tɤ̤ɵ̤, ·tɤ̥], *péngyou* [ʼphɤɴ··jɔʊ; ··jɤɵ, ··jɤ̥], *zhīdao* [¯ʈʂɻ̩··dɑɒ; ··dʌɒ, ··dɤ̥].

Besides, [ɜ] in /aᴇ, ja/ [ɜ, jɜ]: *tàitai* [\thʌ̤ɜ̤..thʌ̤ɜ̤; ·tɐ̤ɜ̤, ·tɜ̤], *rénjia* [ʼʐɜŋ··dʑja; ··dʑjɐ, ··dʑjɜ]; and [ɜ] for /ɤ/ and for any other /a/: *gēge* [¯kɤ··g̊ɤ; ··gɜ], *dǐngzhe* [_ti̤ɲ̤¨dʐ̊ɤ; ·dʐɜ], *àiren* [\ʔʌ̤ɜ̤..ʐɜ̤n; ·ʐɜ̤n], *xiānsheng* [¯ɕjɛŋ··ʂɤɴ; ··ʐ̊ɜɴ], *bàba* [\pa̤..ba̤; ·bɐ̤, ·bɜ̤], *māma* [¯ma··ma; ·mɐ, ·mɜ], *xǐhuan* [_ɕi̤¨ʜwan; ·ɦʊɐn, ·ɦʊɜn, ·ɦɜn], *wǎnshang* [_wʌ̤ŋ̤¨ʂaɴ; ·ʐ̊ɐɴ, ·ʐ̊ɜɴ].

6.5.3. In *fast speech*, the high vowels, /i, y, ɯ, u/ [i, y, ɯ, u] (including /Çɯ/, which are simply [Çɻ̩]), occurring in 'zero-toneme' syllables realized as non-high in pitch, *ie* [.$, ·$], can often be devoiced when preceded by fully voiceles contoids, though in varying degrees and –certainly– not obligatorily.

First, the *partial devoicing of close vowels* after obstruents, /p(h), t(h), k(h); ts(h), ʈʂ(h); f, s, ʂ, h/, *ie* not after (re)sonants, and especially with the fourth toneme, particularly its 'wide' taxotone [\] (see below): /i, y, ɯ, u/ [i̥, ɪ̥, ɩ̥; ẙ, ʏ̥; ɯ̥; u̥, ʊ̥]. Be noted that partially devoiced vocoids do not feature any 'creaky voice' (*ie* laryngealization, [V̰]), as their fully voiced counterparts would, when associated with low taxotones.

A few examples: *pìrú* [\phi̥ʼʐu], *bùfǎ* [\bu̥ˌfa̤a̤], *tìshēn* [\thi̥¯ʂɜn], *dùkǒu* [\du̥ˌkhɔ̤ɔ̤ʊ̤], *jùtuán* [\dʑẙʼthwʌn], *kùchǎ* [\khu̥ˌʈʂha̤a̤], *gùkè* [·g̊u̥\khɤ̤], *cìjī* [\tshɯ̥¯tɕi], *zìmǔ* [\dzɯ̥ˌmṳṳ], *chùfàn* [·ʈʂhu̥\fʌ̤n̤], *jìngzuò* [·dʑɪ̥ɴ\ʈswɔ̤], *fùbù* /\fu\pu/ [·fu̥\pṳ], *xìnyǎng* [\ɕɪ̥ŋ̊ˌja̤a̤ŋ̤], *xùqǔ* [\ɕẙˌtɕhẙẙ], *shùfú* [\ʂu̥ʼfu], *hùzhào* [·ʜu̥\ʈʂa̤ɒ̤].

Further examples: *qǐsī* [.tɕi̥¯sɯ, .tɕɩ̥-, .tɕɪ̥-], *tǔdòu* [.tu̥\tɔ̤ʊ̤, .tɯ̥-, .tɪ̥-], *qǐdòng* [.tɕi̥\tɔ̤ŋ̤, .tɕɩ̥-, .tɕɪ̥-], *tǔdì* [.tu̥\ti̤, .tɯ̥-, .tɪ̥-], *cǐdì* [.tsɯ̥\ti̤, .tsɯ̥-, .tsɪ̥-], *dòufu* [\tɔ̤ʊ̤fu̥, -ɯ̥, -ɪ̥], *xǔduō* [.ɕ̊ẙ¯twɔ, .ɕ̊ɩ̥-, .ɕ̊ɪ̥-], *shǔjià* [.ʂu̥\tɕja̤, .ʂ̊ɩ̥-, .ʂ̊ɪ̥-].

More rarely, even non-close vowels can be devoiced: *kěnéng* [.kɤʼnɤɴ, .kɤ̥-], *yùchǎng* [\ɥẙ.ʈʂɐ̤ŋ̤, -ɐ̥ŋ̤], *hěn hǎo* [_ʜɜ̤ŋ̊·ɦʌɒ, -ɦʌ̥ɒ̥].

6.5.4. In *broader* and even *faster speech*, thus moving pronunciation further away from a real neutral accent, this devoicing can also be complete, [V̥], instead of partial, [V̥]. However, the vocoids do not drop, but are still there. In fact, there is no syllable reduction, although less clearly prominent, with no *[C̩], nor *[Cː, C], instead of [CV], but [CV̥] (in addition to [CV̥]).

A few examples with devoiced elements on their final syllable: *gàosu* [\ka̤ɒ̤.sɯ̥, -ɯ̥, -ɪ̥], *bēizï* [¯pᴇɪ·tsɯ̥, -ɯ̥, -ɪ̥], *cóucou* [ʼtshɔʊ·tsɔʊ, -ɯ̥, -ɪ̥], *dōngxi* [¯tɔɴ·ɕi̥, -ɩ̥, -ɪ̥], *yǎnjing* [_jɛ̤ɛ̤ŋ̊·tɕɪɴ, -ɪ̥ɴ], *chūqü* [¯ʈʂhu·tɕẙ, -ɩ̥, -ɪ̥], *zhīshï* [¯ʈʂɻ̩·ʂɯ̥, -ɯ̥, -ɪ̥], *bāngzhu* [¯paɴ·ʈʂu̥, -ɯ̥, -ɪ̥], *qīngchu* [¯tɕhɪɴ·tʂu̥, -ɯ̥, -ɪ̥].

In addition: *lìqi* [\lḭtɕḭ, -ɿ̥, -ɨ̥], *dōngxi* [¯ṭoɴ·ɕ̥ḭ, -ɿ̥, -ɨ̥], *dàifu* [\ṭa̰ɜ̰f̰ṵ, -ɯ̥, -ɨ̥], *suànshù* [\swa̰ŋ̰ʂ̰ṵ, -ɯ̥, -ɨ̥], *chūqü* [¯tʂu·tɕḭ, -ɿ̥, -ɨ̥], *wèntí* [\wɜ̰ŋ̰ṭḭ, -ɿ̥, -ɨ̥], *xīnkŭ* [¯ɕiŋ·kṵ, -ɯ̥, -ɨ̥], *shìzï* [\ɕḭṭsɯ̥, -ɯ̥, -ɨ̥], *sāncï* [¯san·tsɯ̥, -ɯ̥, -ɨ̥], *yŏu yìsï* [˩jo̰ʊ̰\jḭʂɯ̥, -ɯ̥, -ɨ̥], *lìzhì* [\lḭtʂɯ̥, -ɯ̥, -ɨ̥], *dànshì* [\ṭa̰ʂɯ̥, -ɯ̥, -ɨ̥], *yīqǐ* [¯ji·tɕḭ, -ɿ̥, -ɨ̥], *lìshǐ* [\liʂɯ̥, -ɯ̥, -ɨ̥], *yìsï* [\jḭʂɯ̥, -ɯ̥, -ɨ̥], *shàngcì* [\ʂa̰ŋ̰ṭsɯ̥, -ɯ̥, -ɨ̥], *zhēngqǚ* [¯tʂʌɴ·tɕẙ, -ɿ̥, -ɨ̥].

6.5.5. The combination of personal pronouns (1) with frequent grammeme-like words (2) generally produces reduced forms typical of *very fast speech*.

(1) *wŏ* [˩wo̰ɤ̰] (I), *nǐ* [˩ɲḭḭ] ('you', sing.), *tā* [¯tħa] ('he, she, it'), *-men* [·mɜn] (plural marker): *wŏmen* ('we') [˩wo̰ɤ̰·mɜn], *nǐmen* ('you' pl.) [˩ɲḭḭ·mɜn], *tāmen* ('they') [¯tħa·mɜn];

(2) *yĕ* [˩jḛḛ] ('also, too'), *yào* [\ja̰ɤ̰] ('will, want'), *jiù* [\tɕjo̰ʊ̰] ('then'), *yŏu* [˩jo̰o̰ʊ̰] ('to have'), *yī* [¯ji] ('one, a'), *hái* [˩˥ɣlaɜ] ('still, something more'), *hùi* [\ɣlwḛɪ̰] ('will, can'), *de* [·dɤ] (particle), *shì* [·ʐɯ̰] ('to be').

wŏ yĕ [˩wo̰ɤ̰] + [˩jḛḛ] → [·ɷɘɘ, ·ɷɘ]
wŏ yào [˩wo̰ɤ̰] + [\ja̰ɤ̰] → [·ɷɤ̥ɤ̥, ·ɷɤ̥]
wŏ jiù [˩wo̰ɤ̰] + [\tɕjo̰ʊ̰] → [·ɷɤ̥ɹɯ, ·ɤ̥ɹɯ]
wŏ yŏu [˩wo̰ɤ̰] + [˩jo̰o̰ʊ̰] → [·ɷɤ̥ɯ, ·ɷɤ̥]
wŏ yī [˩wo̰ɤ̰] + [¯ji] → [·ɷɤ̥ɩ, ·ɤ̥ɩ]
wŏ hái [˩wo̰ɤ̰] + [˩˥ɣlaɜ] → [·ɷɤ̥ɐɜ, ·ɷɐɜ]
wŏ hùi [˩wo̰ɤ̰] + [\ɣlwḛɪ̰] → [·ɷɤ̥ɩ]
wŏ de [˩wo̰ɤ̰] + [·dɤ] → [·ɷɤ̥ɯ, ·ɷɯ]
wŏ shì [˩wo̰ɤ̰] + [·ʐɯ̰] → [·ɷɤ̥ɰ, ·ɷɰ];

nǐ yĕ [˩ɲḭḭ] + [˩jḛḛ] → [·ɲɩə]
nǐ yào [˩ɲḭḭ] + [\ja̰ɤ̰] → [·ɲɩɐ, ·ɲɩɐɤ̥]
nǐ jiù [˩ɲḭḭ] + [\tɕjo̰ʊ̰] → [·ɲɩɤ̥]
nǐ yŏu [˩ɲḭḭ] + [˩jo̰o̰ʊ̰] → [·ɲɩɯ]
nǐ yī [˩ɲḭḭ] + [¯ji] → [·ɲiɩ, ·ɲi]
nǐ hái [˩ɲḭḭ] + [˩˥ɣlaɜ] → [·ɲɩɐɜ]
nǐ hùi [˩ɲḭḭ] + [\ɣlwḛɪ̰] → [·ɲɩɷəɩ, ·ɲɩəɩ, ·ɲɩɩ]
nǐ de [˩ɲḭḭ] + [·dɤ] → [·ɲɩdə, ·ɲɩɡ̥ə, ·ɲɩɤ, ·ɲɩə]
nǐ shì [˩ɲḭḭ] + [·ʐɯ̰] → [·ɲiɰ];

tā yĕ [¯tħa] + [˩jḛḛ] → [·tɐɩə, ·tɐɩ]
tā yào [¯tħa] + [\ja̰ɤ̰] → [·tɐ·ɟɤ̥ɯ, ·tɐɩɤ̥]
tā jiù [¯tħa] + [\tɕjo̰ʊ̰] → [·tɐ·ɹɯ, ·tɐɩɤ̥]
tā yŏu [¯tħa] + [˩jo̰o̰ʊ̰] → [·tɐ·ɟɯ, ·tɐɩɯ]
tā yī [¯tħa] + [¯ji] → [·tɐɩ]
tā hái [¯tħa] + [˩˥ɣlaɜ] → [·tɐɜ]
tā hùi [¯tħa] + [\ɣlwḛɪ̰] → [·tɐ·ɷɩ, ·tɐɩ]
tā de [¯tħa] + [·dɤ] → [·tɐɯ]
tā shì [¯tħa] + [·ʐɯ̰] → [·tɐɰ];

wǒ hái ('and I') [ˏwǫǫ] + [ˊɦʌɜ] → [·ɷɐɩ]
wǒmen [ˏwǫǫ] + [·mɜn] → [·ɷɤ̞m̩, ·ɷɤ̞m, ·ɷm̩, ·ɤ̞m]
nà wǒmen ('then we') [\nạ] + [ˏwǫǫ·mɜn] → [·nɞm, ·nɤm]
nǐmen [ˏɲịị] + [·mɜn] → [·ɲɩm̩, ·ɲɩm]
tāmen [ˉtʰa] + [·mɜn] → [·tɐm̩, ·tɐm].

6.5.6. Let us also notice, again most typical of *very fast speech*:
kěyǐ ('can, be able to') [ˏkʰɤ̞ɤ̞] + [ˏjịị] → [·kɷɤ̞ɩ]
bú huì ('cannot be') [ˊp̣u] + [\ɦwəɪ̣] → [·bɷɤ̞ɩ]
jiù yào ('going to') [\tɕjọʊ̣] + [\jạọ] → [·ʥɹɤ̞ɤ̞]
bǎ tā ('constrain') [ˏp̣ạạ] + [ˉtʰa] → [··p̣a]
suǒyǐ ('therefore') [ˏswǫǫ] + [ˏjịị] → [·sɷɤ̞ɩ]
zhè yàng ('in this way') [\ʈʂɤ̞] + [\jạŋ̣] → [\tɕjạŋ̣]
dàjiā ('everybody') [\ṭạ] + [ˉtɕja] → [·dɐɩə, ·dɐɤ]
bǐjiào('compared with') [ˏp̣ịị] + [\tɕjạọ] → [·bɹɐɤ̞]
yǒu shíhour ('sometimes') [ˏjǫǫʊ̣] + [ˊʂɻ̩·ɦɤ̞ɤ̞] → [·ɹɤ̞ɤ, ·ɹɤ̞ə]
bù yí yàng ('differently') [\p̣ụ] + [ˉji] + [\jạŋ̣] → [\p̣jạŋ̣].
dōu yǒu ('completely') [ˉṭɞʊ] + [ˏjọʊ̣] → [·ṭjọʊ̣]
zhǐyào ('if') [ˏʈʂɻ̩ɻ̩] + [\jạọ] → [··ʥ̊aɒ]
dāngrán ('of course') [ˉṭaɴ] + [ˊʐʌn] → [··ḓ̊ãẽ]
ránhòu ('and then') [ˊʐʌn] + [\ɦọʊ̣] → [·ʐɞʊ]
zhè-zhǒng ('this kind of') [\ʈʂɤ̞] + [ˏtɕọŋ̣] → [··ʥ̊ʐɞɴ]

Possible changes due to emphasis

6.6. On the contrary, *fig* 6.5 shows the opposite results which can occur when, instead of reductions, due to 'zero tonemes', speakers resort to emphasis on certain words or syllables, for particular communicative purposes.

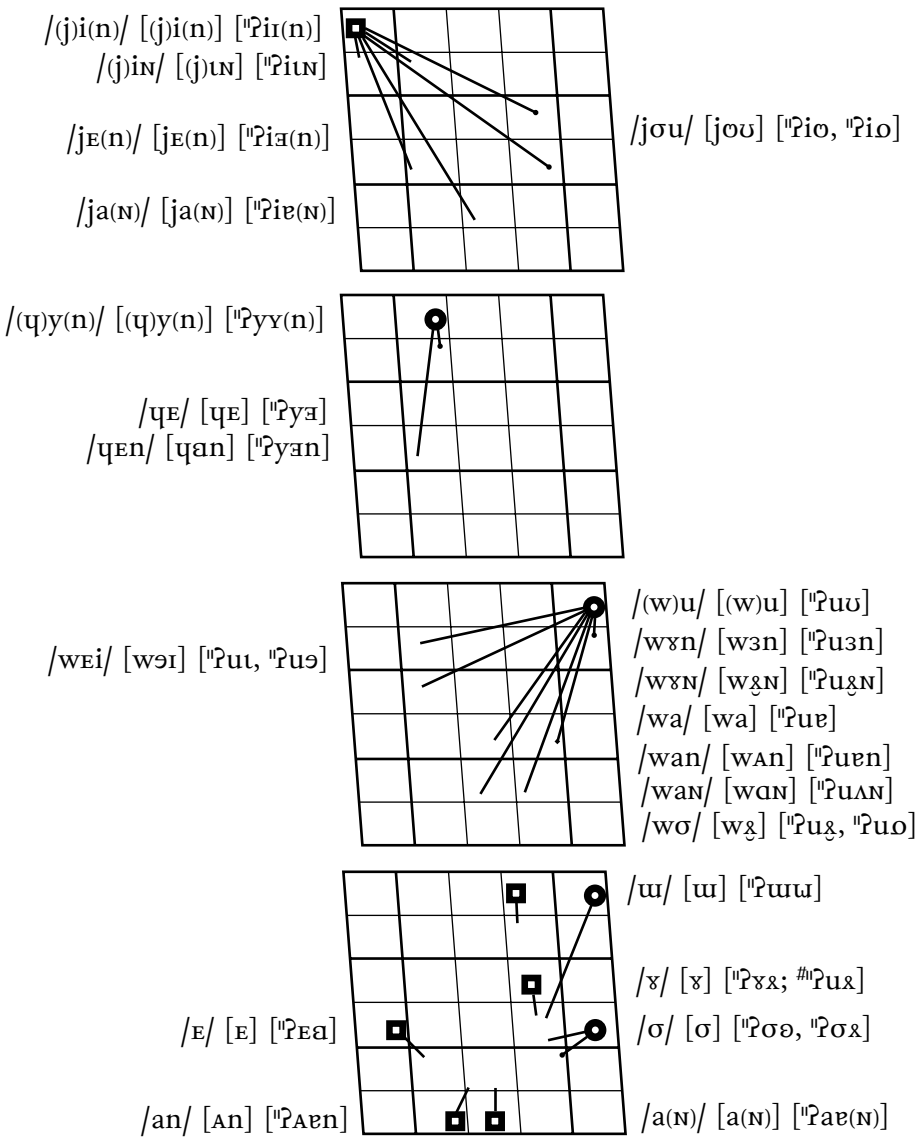
The vocograms given in *fig* 6.5 might even explain, but certainly not justify, the unique and excessive use in Chinese books of /VV, VVV/ sequences, instead of normal /JV, JVV/ ones (*cf* Ch 13).

Of course, perhaps their authors may mean something like that, but it is certainly not the best way to 'explain such facts'.

As a matter of fact, they should –at least– reverse their 'norms', by giving something like our own sequences, [ˈJV, ˈJVV], and only afterwards also adding [ˈˈVV, ˈˈVVV], including paraphonic peculiarities, which do not exclude at all the more convenient possibility [ˈˈJV, ˈˈJVV], which is actually used, as well.

The first author goes as far as to think that, probably, their problem originates from an insufficient analysis of the phonic structure they try to handle, or from excessive, unfounded, confidence: *zhuāng* [ˉʈʂwɑɴ, ⁼ʈʂwɑ(˙)ɴ, ⁼ʈʂuʌɴ] (also with [⁼ʈʂ-]).

fig 6.5. Frequent emphatic (or hyperprecise, or very slow) variants, which can be heard instead of the more normal ones (always with emphatic stress and paraphonics). Also notice the presence of [ʔ], and even /ˌV/ [ˌV̥V̥] [ˌV̥ʔV̥].



6.7. In addition, *fig* 6.6.1-2 show that /jV, ɥV, wV/ sequences, too often still called and classified as 'rising diphthongs', are nothing but normal sequences formed by approximant consonants and vowels (or diphthongs).

Of course, similar tracings could be done in Ꮆ 7 and others, as well. It is very important to discover the real difference between such sequences and true diphtongs (and triphthongs, too).

fig 6.6.1. /jV, ɥV, wV/ sequences, which are clearly different from true diphthongs.

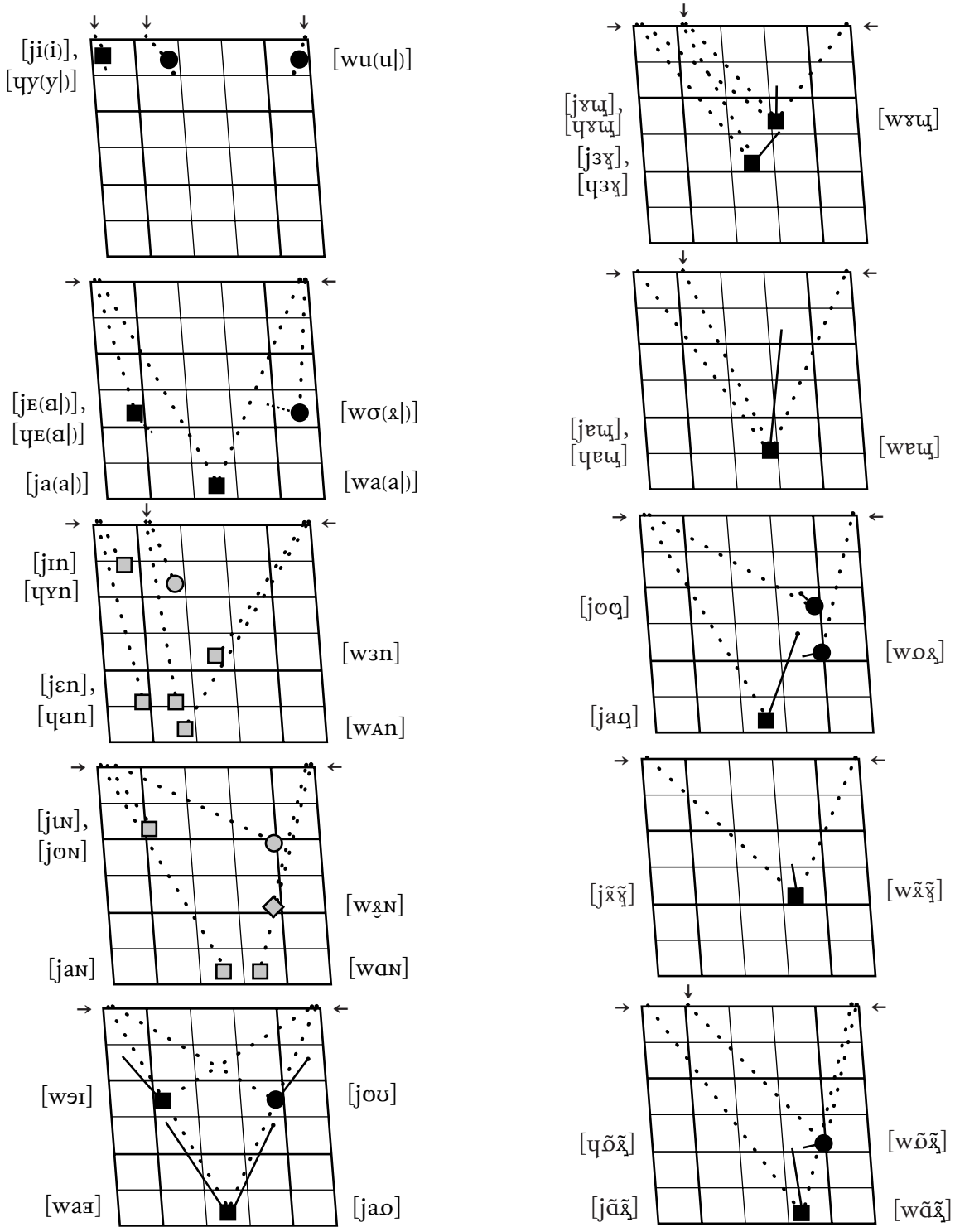
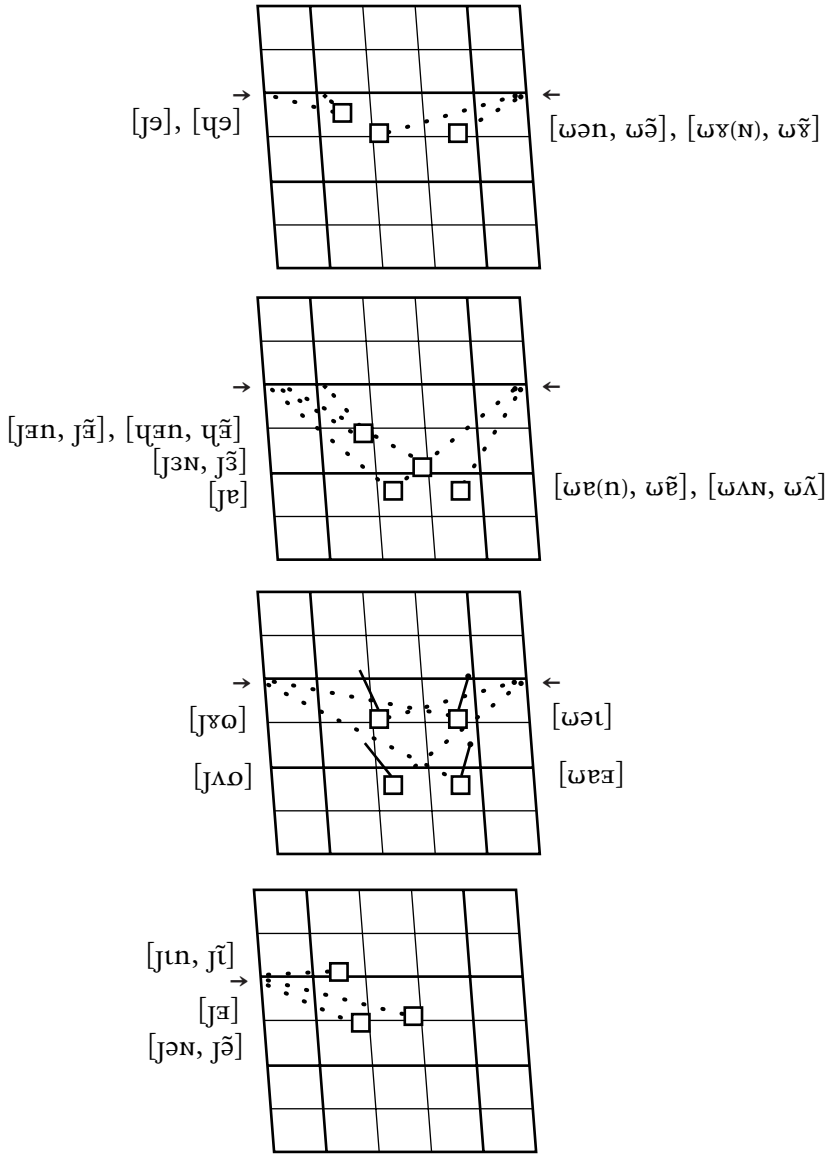


fig 6.6.2. /jV, ɥV, wV/ sequences, when occurring in weak syllables.



Some curios

Variants of zhè *(this),* nà (*that*), nǎ *(which one?)*

6.8. In Chinese, *zhè*, *nà*, and *nǎ* are three very frequent monosyllables, which we usually classify as 'pronouns', while Chinese linguists adopt the probably better definition of 'substitutes': depending on their position within a sentence and the words that might be affixed to them, substitutes can act as English demonstrative and interrogative adjectives/pronouns, locative and modal adverbs, and so on.

When followed by the numeral *yī* 'one', *zhè yī*, *nà yī*, and *nǎ yī* can be fused together into the colloquial variants *zhèi*, *nèi*, and *nǎi/něi*. By logic, they should not be used when another numeral follows, but the use and abuse of the language shows a certain trend to recognize these contractions as simple alternatives, though

–in fact– they are different lexemes. For *nà* only, the *Contemporary Chinese Dictionary* also provide the variant *nè*.

More exactly, *nǎ* is actually pronounced as such when used alone; but when it is followed by a classifier or a numeral plus a classifier, the pronunciations *nǎi*/*něi* are admitted: *nǎi*/*něi ge*, *nǎi*/*něi huìr*, *nǎi-*/*něixiē*, *nǎi-*/*něiyàng(r)*. When used alone or immediately followed by a verb, *nà* is pronounced *nà*/*nè*; when followed by a classifier or a numeral plus a classifier, the pronunciation *nè(i)* is admitted: *nè(i) chéngzï*, *nè(i) ge*, *nè(i)-huìr*, *nè(i)xiē*, *nè(i)yàng(r)*.

The lexeme *nàme* 'in that way' has the colloquial variant *nème*. Finally, *zhè* keeps this pronunciation when used alone or followed by a substantive; when followed by a classifier or a numeral plus a classifier, the pronunciation *zhèi* is admitted: *zhè(i) chéngzï*, *zhè(i)-ge*, *zhè(i)-huìr*, *zhè(i)xiē*, *zhè(i)yàng(r)*.

The lexeme *zhème* 'in this way' also has the colloquial variant *zème*. However, when speaking or reading aloud in a solemn ‰r official manner, it is advisable for native and foreign speakers alike to stick to the basic forms.

Variants of the modal grammeme ***a***

6.9. The grammeme *a* can only occur at the end of a sentence. Usually, it conveys familiarity, surprise, or enthusiasm; but it can also act as a full pause, possibly with a suspense flare, before passing to the following sentence. Being a weakly-stressed syllable with no specific meaning, *a* will be regularly absorbed by a *le* that precedes, resulting in *la*.

Furthermore, there are some euphonic variants that occur after certain syllabic finals. Such variants are even reflected in *pīnyīn*, though not always systematically (in fact, broader transcriptions usually present *a* only): after *i(e)*, *-ai*, *-ei*, *-ü(e)*, and *-uo*, *ya* is used; after *-ao*, *-(o)u*, there is *wa*; after *-n*, there is *na* (which is written with the same Chinese character representing the substitute *nǎ*); while after *-ng*, *pīnyīn* tends to prefer a simple *a* for what actually is /◦na/ (with a 'zero toneme'), as there is no specific Chinese character, though a rarely-seen *nga* is admitted.

The simple rendition *a* is maintained for [ɥa, ɥa], too, which occur after *-ï*. One may want to resort to something like *ra*, which nevertheless should not be articulated as a proper [ʐa]: foreign speakers should better aim at a sort of an unstressed diphthong, like *-ï+a* [◦ɯa; ◦ɥa, ◦ʲʐa], which –by the way– are very close to native articulations.

Stress variants in numerals and numeral syntagms

6.10. Foreign speakers can safely pronounce all numerals by applying standard stress patterns, as explained in the relevant chapter. However, once enough fluency and proficiency in basic pronunciation are secured, they may want to benefit from the few tips that we list in this paragraph for an even smoother pronunciation of numerals.

Within a numeral syntagm, that is to say, a string composed of a numeral and a classifier, the latter is weakly stressed, pretty much like the generic classifier *gè*

(which is always pronounced with a 'zero toneme' and is even written as such: *ge*): *sān-ge lǎoshī̈* [ˉsʌŋ··g̊ɤ̤.lɑ̜o̜ˉʂɰ̢, -·gɤ-, -·gɤ̞-] 'three teachers', *sì̈-zhāng zhǐ̈* [\̩sɰ̢.dʐ̥a̜ɴ̜˩ʈ͡ʂɰ̢ɰ̢, -.dʐa̜ɴ̜-] 'four sheets of paper', *wǔ-běn shū* [˩wu..b̥ɜ̜ɳ̜ˉʂu, -.bɜ̜ɳ̜-] 'five books', *liù-bǎ yǐzï* [\ljo̜ʊ̜·b̥a_jḭ·dzɯ] 'six chairs'; but *jiǔ-suì de háizï* [..dʑ̥jo̜ʊ̜\swə̜ɪ̜..d̥ɤ̜˩ʞʌɜ·dzɯ, -.dɤ̜-, -.dɤ̞̜-] 'nine-year-old boy'.

The numeral *shí̈* 'ten', used in isolation, is pronounced fully stressed: [˩ʂɰ̢]. In two--syllable numerals, *shí̈* features a semi-weak stress if it is the first syllable: *shí̈'èr* [·ʂɰ̢-\(ʔ)ɤ̜ɤ̜̆] 'twelve'; otherwise, it is fully stressed: *èrshí̈* [\̩(ʔ)ɤ̜ɤ̜̆˩ʂɰ̢]. When *shí̈* occurs between two numerals, or between a numeral and a classifier, it usually has a weak stress, and it can even pass to a 'zero toneme' in fast connected speech: *èrshí̈-zhāng* [\̩(ʔ)ɤ̜ɤ̜̆·ʂɰ̢ˉʈ͡ʂaɴ, -··ʂɰ̢-, -···ʂɰ̢-, -..ʂɰ̢-, -·ʂɰ̢-, -.ʂɰ̢̥-] 'twenty sheets', *sānshí̈sān* [ˉsʌɳ·ʂɰ̢ˉsʌn, -··ʂɰ̢, -·ʂɰ̢̥-] 'thirty-three'.

Foreign words and names

6.11. In this section, a concise account is provided of how foreign words, including proper names and a small number of scientific lexemes, are rendered in Chinese. Forcing a word to fit in another language where its structure sounds completely unnatural can be unpleasant, but in Chinese, it is often disappointing and even irritating, as the Chinese not only insist on adapting the original phones to those more familiar to them (which is the right thing to do: all languages work like that).

They also stubbornly refuse to go anywhere beyond the meager native taxophonic inventory, those 400-something syllables that are described at the beginning of this book. The reason is quite simple: as native speakers expect everything Chinese to be written down in Chinese characters, one obviously has to stick to their original pronunciation, even in those cases where one could *orally* dare to articulate some unorthodox, yet possible, combination.

For example, nothing prevents Sinophones from producing syllables ending in *-m* (which, by the way, was very common in the Mandarin spoken just a few centuries ago). In fact, the common plural suffix *-men* is usually reduced to its sole initial *-m* in fast connected speech: *wǒm(en)*, *nǐm(en)*, *tām(en)*.

However, no Chinese character permits to represent such combinations in writing, anymore. Therefore, the English name *Tom* /ˈṭhɒm/ becomes *Tāngmǔ* or just *Tāng*, also because there seems to be no real rules. Very frequently, a character is chosen for its good meaning or beautiful shape rather than for its fidelity to the original sound.

The English surname *Nixon*, for example, has been variously rendered as *Níkèsōng*, *Nígésōng*, and *Níkèsūn*, because when the then US President, Richard Nixon and China commenced the talks that eventually led to the establishment of diplomatic relations between the United States and the People's Republic of China, all Chinese-language media in the world, both in China and abroad, started writing or broadcasting about that historical turning point at the very same time, with little coordination, and different sources literally 'made up their own Nixon'.

For decades now, the People's Republic of China's authorities have issued offi-

cial lists of foreign words 'transcribed' in Chinese characters. Unfortunately, these lists contain neither rules nor norms, but mere combinations of characters that have to be used and memorized as they are.

In other words, even when one knows how to pronounce a foreign word or name in Chinese, it is most often impossible to deduce which characters are actually used, and only a small portion of them is systematically employed to express certain sounds.

6.12. Furthermore, the People's Republic of China and the Republic of China (Táiwān) not always follow the same criterion. While in mainland China a decent approximation to the original pronunciation is the principle now universally applied, in Táiwān it is quite common that even prominent international figures are given some sort of a 'Chinese name'.

For example, the surname of the former US Secretary of Defence, Donald E. Rumsfeld is *Lāmǔsīfēi'ěrdé* in the RPC, which some 'more refined' speakers have the courtesy of pronouncing somehow like [ˈɻʌmsfɤɻtʂ̩]. In Táiwān, the official version of *Rumsfeld* is the more Chinese-sounding *Lún Sīfěi*, which we transliterate as such because the Chinese themselves perceive it as a sequence of a monosyllabic surname, *Lún* and a dissyllabic given name, *Sīfěi*.

It is not infrequent that even certain official renditions are at odds with embarrassingly-sounding homophones, but the Chinese have not been consistent about this matter: Marco Polo's surname is *Bōluó* and it sounds just like *bōluó* 'pineapple'. Similarly, *Luómǎ* 'Rome' cannot be distinguished from *luó-mǎ* 'mules and horses', exept in writing.

On the other hand, the former Italian President, Carlo Azeglio Ciampi was assigned a very suggestive *Qiánpí*, lit. 'money skin', which sounds quite apt for someone who used to be the Chief Governor of the Bank of Italy. At least, he was spared something even odder, like a 'toneless' *Qiangbi*, given the phonetic proximity to *qiángbí* 'wall, lock tightly', or *qiǎngbī* 'compel by force', and especially… *qiāngbì* 'to execute by firing squad'!

6.13. Finally, there is a category of foreign words that are doomed to suffer the ultimate affront: not just a mere phonetic transposition to Chinese, but total –even cultural– assimilation based on the common (Sino-)graphic heritage. Since the Chinese characters are still used in Japan and Korea, and were used for a long time in Vietnam as well, the Chinese assume that if a term can be written in Chinese characters, it can consequently be pronounced as if it were Chinese.

Therefore, a Japanese gentleman named Fujimoto Ichirō will be obliged to introduce himself in Chinese as *Téngběn Yīláng*, pretty much as if we forced the former F1 champion, Michael Schumacher to rename himself as *Miguel Zapatero* in Spain and *Michel Cordonnier* in France! Nor will poor Mr Fujimoto be allowed to resort to other Chinese characters that would better represent the pronunciation of his name and surname: his passport bears those four Chinese characters, and their only possible pronunciation in Chinese is the one indicated above.

In any case, it is true that the Japanese and Koreans themselves will usually prefer that meaning is preserved over sound. At least, the Koreans have a chance to take a good revenge over their neighbours, as they systematically Koreanize all Chinese words, including proper names, by means of their phonemic script *hangul*.

In Japan, on the other hand, a peculiar and interculturally respectful system is applied to Chinese words, especially those that have been introduced more recently: the original Chinese characters are preserved, and a generally passable rendition of the original pronunciation is expressed by means of the *furigana* (small superscript square-shaped syllabic *katakana* symbols).

In this fashion, *Máo Zédōng* can be legitimately pronounced *Mō Takutō*, but a more realistic *Mao Tsoton* will be suggested by the *furigana*.

6.14. In our view, a fair solution to the deplorable treatment that foreign words receive in Chinese would be that of keeping the original orthography wherever possible: thus, Chinese characters for Korean and Japanese loans, accompanied by their relevant romanization; the Latin alphabet, which forms the basis of *pīnyīn* itself and is –therefore– officially recognized by the Chinese government, could be usefully employed for all the languages that already use it or for which a standard romanization scheme has been devised.

In this manner, those who are more gifted and more enlightened would have a chance to improve their pronunciation. For the lamentably numerous ones who are unredeemable, a transposition based on Chinese characters might well be added between parenthesis, provided it is not used as the official version.

However, we are so confident in the potentiality of the International Phonetic Alphabet (even in its current, incomplete official form) that we do not hesitate to boldly propose its extensive use in all texts whose readers are supposed to have a certain knowledge of foreign languages: first and foremost, conversation handbooks, as well as linguistics, phonetics, and philology handbooks, but also books on the history and geography of other countries.

This is typically applicable to university students of foreign languages, politics, and economy; in short, to whomever will sooner or later come in contact with foreigners.

Those who read such books must know that the famed eighteenth-century Italian politician, Camillo Benso, Count of Cavour [khɯˈvɷ˕ɐ, -ˈvʊː] ([kaˈvuːr] in Italian) cannot possibly be referred to as *Jiāfùʼěr* [ˉtɕja˦fu̹˩(ʔ)ɣ̹ɣ̹˨] , as it is today, outside of a strictly Chinese-speaking environment. Nobody expects a perfect-pitch Italian or French pronunciation, which would even sound affected and pretentious, but a totally viable rephonemicization like /kaˈwuʐ, kʰ-, -ˈf-/ would do nicely, without much suffering… or inflicting unneeded sufferance on others.

7.
The *-r* grammeme and the so-called *érhuà* phenomenon

7.1. As a result of an evolution process which lasted for centuries, the northern Chinese dialects (Pekingese and Standard Chinese itself among them), developed a series of monosyllabic unstressed suffixes, still in use today, like *zï* and *li*… 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 contoid, [ɻ̩]), 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.

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 *stød*, when describing Danish pronunciation), or to resort to something like '*r*-coloring' 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 with a phonic signifier and various realizations, as we will see right away.

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

bāng /⁻paɴ/ [⁻ṗaɴ] 'to help' ᴠs *bāngr* /⁻paɴɻ̩/ [⁻ṗɑ̃ɻ̩̃] 'gang';
gān /⁻kan/ [⁻ḳʌn] 'to be dry' ᴠs *gānr* /⁻kaɻ̩/ [⁻ḳɐɻ̩] 'dried food';
huà /\ɧwa/ [\ɬwạ] 'to depict' ᴠs *huàr* /\ɧwaɻ̩/ [\ɬwɐ̯ɻ̩̯] 'a painting';
huó /ˏɧwσ/ [ˏɬwσ(ɤ|)] 'to live' ᴠs *huór* /ˏɧwσɻ̩/ [ˏɬwɔɤ̯] 'activity, manual labor';
shù /\ʂu/ [\ʂụ] 'to stand vertically' ᴠs *shùr* /\ʂwɤɻ̩/ [\ʂwɤ̯ɻ̩̯] 'vertical brush stroke in Chinese calligraphy'.

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

tóu /ˏtɧσu/ [ˏtɬσʊ] 'head' ᴠs *tóur* /ˏtɧσuɻ̩/ [ˏtɬσɑ] 'chief';
xìn /\sin/ [\ɕi̯ɲ̟] 'letter' ᴠs *xìnr* /\sjɤɻ̩/ [\ɕjɤ̯ɻ̩̯] 'information, news';
bàntiān /\pan⁻tɧjɛn/ [\b̥ạɲ̟⁻tɬjɛn] 'a long time' ᴠs *bàntiānr* /\pan⁻tɧjaɻ̩/ [\b̥ạɲ̟⁻tɬjɐɻ̩] 'half a day';

māoyǎnr /ˉmaʊˏjaʐ/ [¨mɑɒˏjɐ̹ɐ̹ɣ̹] 'peephole' *vs* *māoryǎn* /ˉmaʊʐˏjᴇn/ [¨mɑɒ̹ˏjɛ̹ɛ̹ɳ̹] 'opal'.

7.1.3. Finally, an *r*-suffixed word may easily suggest a diminutive, reductive *&/or* affectionate acceptation:

bīnggùnr /ˉpiɴ\kwɤʐ/ [¨b̥ɪɴ\k̹wɣ̹ɣ̹] 'ice lolly';
nánhái r /ˊnanˊɦaʐ/ [˙nᴀɴˊɦɐɣ̹] 'baby boy';
ménfèngr /ˊmɤn\fɤɴʐ/ [˙mɜɱ\fɣ̹ɣ̹̃] 'a tight crack between a door and its frame, or between two doors';
xiǎo Zhāngr /ˏsjaʊˉʈʂaɴʐ/ [..ɕja̹ɒ̹ˉʈ̹ʂɑ̃ɣ̹̃] 'young Zhāng';
lǎo māor /ˏlaʊˉmaʊʐ/ [.lɑ̹ɒ̹ˉmɑɒ̹] 'old dear kitten';
àiqíngr /\ʔaᴇˊʨɦjɤɴʐ/ [\(ʔ)ᴀɜ̹ˊʨɦjɣ̃ɣ̹̃] 'unimportant love affair';
méi shìr /ˊmᴇi\ʂɤʐ/ [˙mᴇɪ\ʂɣ̹ɣ̹] 'no problem';
xiǎoshuōr /ˏsjaʊˉʂwʊʐ/ [..ɕja̹ɒ̹ˉʂwɒɣ̹] 'novelette'.

7.2. With all its taxophones, *érhuà* provides one of the most prolific fields of research on neutral (and mediatic) Chinese pronunciation: if a simple addition of [ʐ] to any syllabic rhyme –*V*, *VV*, *Vn* and *Vng*– might be sufficient for an 'international' kind of pronunciation to sound understandable and convincing, phonetic reality is quite more complex.

Only certain vocoidal phones are able to coexist with /ʐ/, which is primarily a *post-alveolar approximant*, [ʐ]. 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, good audio files have revealed such a wide range of taxophonic oscillations, that it is no longer possible to passively accept what the 'phono-sinological' tradition has nearly passed off as a sacred dogma.

7.2.1. For decades, nonsensical and frankly embarrassing transcriptions like '[Vr]' 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* '/Vr/' or '[Vɹ]', but still inadequate, both phonemically and phonetically. Such inadequacy derives both from the conceptual and typographical limits of *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 [ʐ] contoid 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), [Vʐ, Ṽʐ̃], in truly neutral and spontaneous pronunciation –neither emphatic nor mannered– /Vʐ/ sequences are more often realized as [Vɣ̹, Ṽɣ̹̃], *ie* a simple vocoid followed by a *provelar* (: back-central) *vocoid with lateral contraction*.

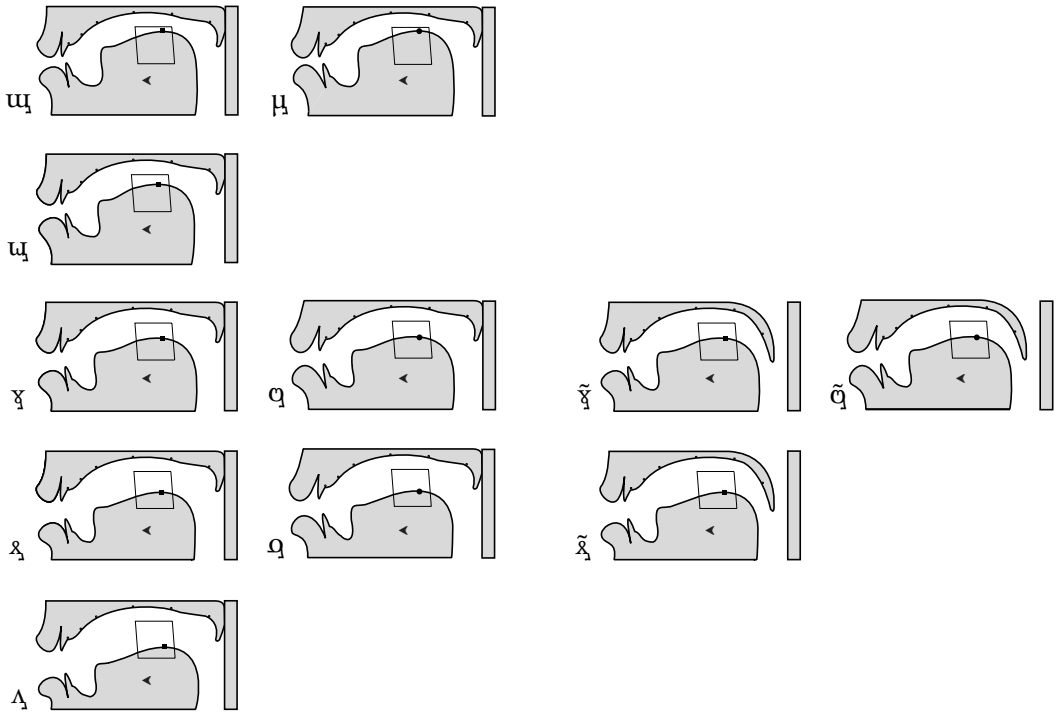
7.3. As shown in fig 7.1, the apex is *slightly* raised behind the alveoli, certainly not 'curled', and that is why we disagree with the so-called 'retroflex vowel [ɚ]', which is frequently encountered in traditional descriptions of *érhuà*. A certain degree of postalveolarity in any [ɣ] is indisputable, but the essential feature to be underlined is the *lateral contraction*, which inevitably causes the apex to move back and raise.

Concisely, any [ɣ] 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 voicing and potential nasalization; and of *consonantal traits* fairly similar to those of a *slightly rounded and postalveolarized prevelar semi-approximant with lateral contraction*, [ɹ̠̊˔].

The corresponding *approximant* is nothing but the famous 'General American /ɹ/' [ɹ]. (while Southern British /ɹ/ is, rather, a *slightly rounded postalveolar approximant*, [ɻ], *pace* what most –even native– phoneticians still seem to think, or to accept).

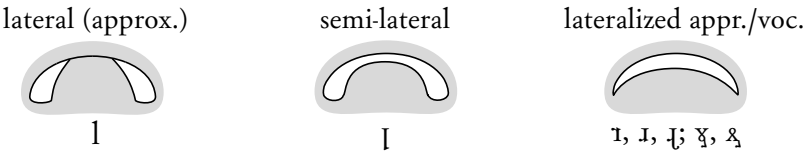
The very symbol '[ɹ]' suggested us the presser-foot-like diacritic in '[ɣ]'. Yet, there are some differences worth noticing: *Pŭtōnghuà* [ɣ] are generally *pro*-velar, while [ɹ, ɹ̠] are *pre*-velar; the latter two are also *intrinsically slightly rounded.* Instead, the potential roundedness of any [ɣ] depends on that of their 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*, [ɻ̟]. Therefore: [V] + [ɻ̟] → [ɣ].

fig 7.1. Vocograms of the laterally contracted vocoids (cf fig 5.3-4 for normal vocoids, and fig 13.2.1 for a Pekingese variant).



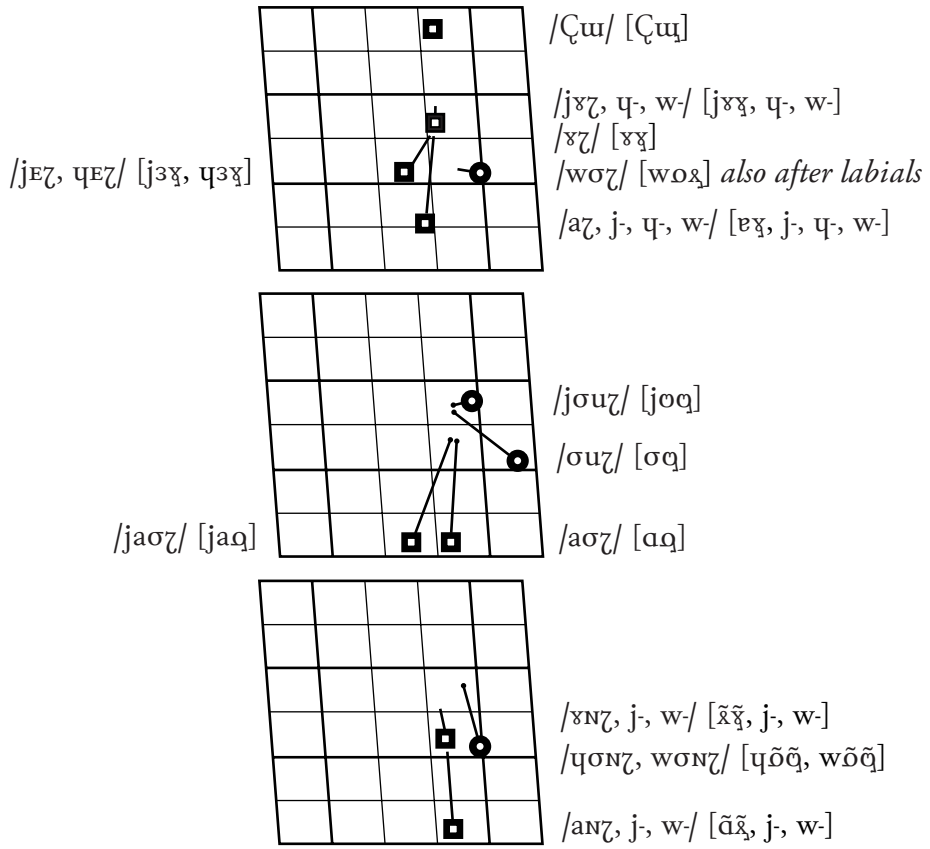
7.4. Before analyzing how these peculiar vocoids behave in *érhuà*, it is important to note that, even in /ɕɯ/, we normally find a lateralized articulation, [ɯɻ], rather than an intense approximant like [ʑ], as various *offIPA* notations –*eg* '[ɻ, ɻ̟]'– would suggest. Many sinologists use '[ɭ]', a non-standard symbol whose *offIPA* equivalent should be '[ɻ̩]', described as a 'syllabic retroflex fricative' or even 'retroflex vowel' (though written with a contoid glyph! By the way, let us note that *canIPA* [ɭ] is an apical-postalveolar flap, instead). Undoubtedly, [ʑ] belongs to neutral pronunciation as [ɯɻ] does, but it is associated with a tenser articulation and a certain emphatic component (if not with a marked Peking accent).

fig 7.2. Frontal orograms showing a different perspective.



7.5. Starting with /(J)Vʐ/, /(J)a, ɤ, wo/ and /ʐ/ interact without phonemic adjustments: /aʐ, jaʐ, waʐ; ɤʐ; woʐ/. Chinese taxophonics does not allow /‖iʐ, yʐ‖/ as such, and adjusts them to /jɤʐ; ɥɤʐ/. Exceptionally, /‖iʐ, yʐ‖/ present oscillations that might suggest quite peculiar phonemicizations: [iɻ, yɻ] /iɤʐ; yɤʐ/; but all the sound files that we have examined clearly show that these oscillations are always related to inten-

fig 7.3. Vocograms of the laterally contracted vocoids for the *suggested pronunciation*.



tionally slow &/or careful speech. In addition, //Cɯʐ, Çɯʐ// change into /Cɤʐ, Çɤʐ/, in order to be better distinguished from [Cɯ] (or [Cʑ̩]) and especially from [Çɯ̢] (or [Çʐ̩]). Herein, we list the most frequent taxophones, including those with ↕[ʐ] (where '↕' suggests a less spontaneous pronunciation, often used by native speakers with a regional background and an intentional effort to approach neutral pronunciation.

7.5.1. 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 neutral realization shown on the list, and it needs not to: for example, [ɐɰ̢] is the most typical for /aʐ/, yet we 'suggest' [ɐɣ̢], because it sounds as much as natural to natives, and it employs what we consider to be a very versatile taxophone, [ɣ̢]:

ar; *yar*, *iar*; *war*, *uar* /aʐ, j-, w-/ [ɐɰ̢, j-, w-] — ˢ[ɐɣ̢, j-, w-];
er & *ïr* /ɤʐ/ (//ɯʐ//) [ʌɣ̢] — ˢ[ɤɣ̢];
yir, *ir*; *yür*, *ür* /jɤʐ, ɥɤʐ/ [jɤɰ̢, ɥ-] — ˢ[jɤɣ̢, ɥ-];
wor, *uor*; *(m-b-p-f-)or* /wɔʐ/ [wɒʌ̢; Çɒʌ̢] — ˢ[wɒʌ̢].

7.5.2. For *yer*, *ier*; *yüer*, *üer* //jᴇʐ, ɥᴇʐ//, the most typical realizations are [jɜɣ̢; ɥɜɣ̢], which we still prefer to assign to /-ᴇʐ/ (in a more useful interphonemic analysis, according to the principle of Natural Phonetics, rather than to (mediatic) /-ɤʐ/, in a supposed 'more economical' –but more abstract and less useful– intraphonemic analysis). So, we recommend the taxophonic sequences shown, [jɜɣ̢; ɥɜɣ̢] (or, at most, ↕[jɜʐ, ɥɜʐ], which risk to sound too precise, because of the use of the contoid [-ʐ], instead of the more typical laterally contracted vocoids, *cfr fig* 7.1):

yer, *ier* & *yüer*, *üer* /jᴇʐ, ɥᴇʐ/ [jɜɣ̢, ɥ-].

7.5.3. Similarly, //uʐ// provides two choices, as in mediatic pronunciation (Ch 13): /wɤʐ/, which we consider as neutral; and /uʐ/, certainly not indispensable, but useful to improve one's understanding of spoken Chinese:

(w)ur /wɤʐ/ [wɤɰ̢] — ˢ[wɤɣ̢].

7.5.4. The second elements of the diphthongs /(w)ᴇi, (w)aᴇ/ are incompatible with the place of articulation of /ʐ/; they are therefore reduced to a 'zero' element, leaving /(w)ɤʐ, (w)aʐ/; in theory, this causes homophony with *er* and *ar*, though the very few cases where this happens are substantially negligible.

7.5.5. Regarding /jaɔʐ, aɔʐ/, both vocoids 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 back-central quality of [ɒʌ̢]. A similar behavior distinguishes /jɔuʐ/ from /ɔuʐ/, as clearly shown (*cfr fig* 7.3), in spite of the high number of mediatic variants

currently occurring for this diphthong:

aor /aσɀ/ [ɑɑ̯] — ˢ[ɑɑ̯];
yaor, *iaor* /jaσɀ/ [jaɑ̯] — ˢ[jaɑ̯];
our /σuɀ/ [σʊ̯] ;
your, *iur* /jσuɀ/ [jʊʊ̯].

7.6. In *-nr*, /n/ disappears, too. Consequently, *gēr* 'song' and *gēnr* 'root' are pronounced alike, /ˉkɤɀ/ [ˉk̟ɤɤ̯]; as well as *bàr* 'handle' and *bànr* 'obstacle, petal', /\paɀ/ [\p̀p̟ɤ̀ɤ̯]. It is important to note that *pīnyīn* is certainly not the actual phonic structure of Chinese (as many are led 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– phonemic level, more like our possible // // symbolization. For instance, we might have the sequences //inɀ, ynɀ, wɤnɀ//, for *inr*, *ünr*, *enr*; but /jɤɀ, ɥɤɀ, wɤɀ/ are certainly better.

As for //jɛnɀ, ɥɛnɀ//, it must be remembered that /jɛn, ɥɛn/ are simply more realistic phonemicizations for [jɛn, ɥɑn] than '/jan, ɥan/'; but, when /n/ disappears –by neutralization– the syllabic nucleus resumes its theoretical 'original' vowel quality: /jaɀ, ɥaɀ/ (//janɀ, ɥanɀ//), with the same values as described above for /aɀ/, eg *yüánqüānr* /ˈɥɛnˉʦʰɥaɀ/ [ˑɥɑɲˉʧʱɥɛɤ̯].

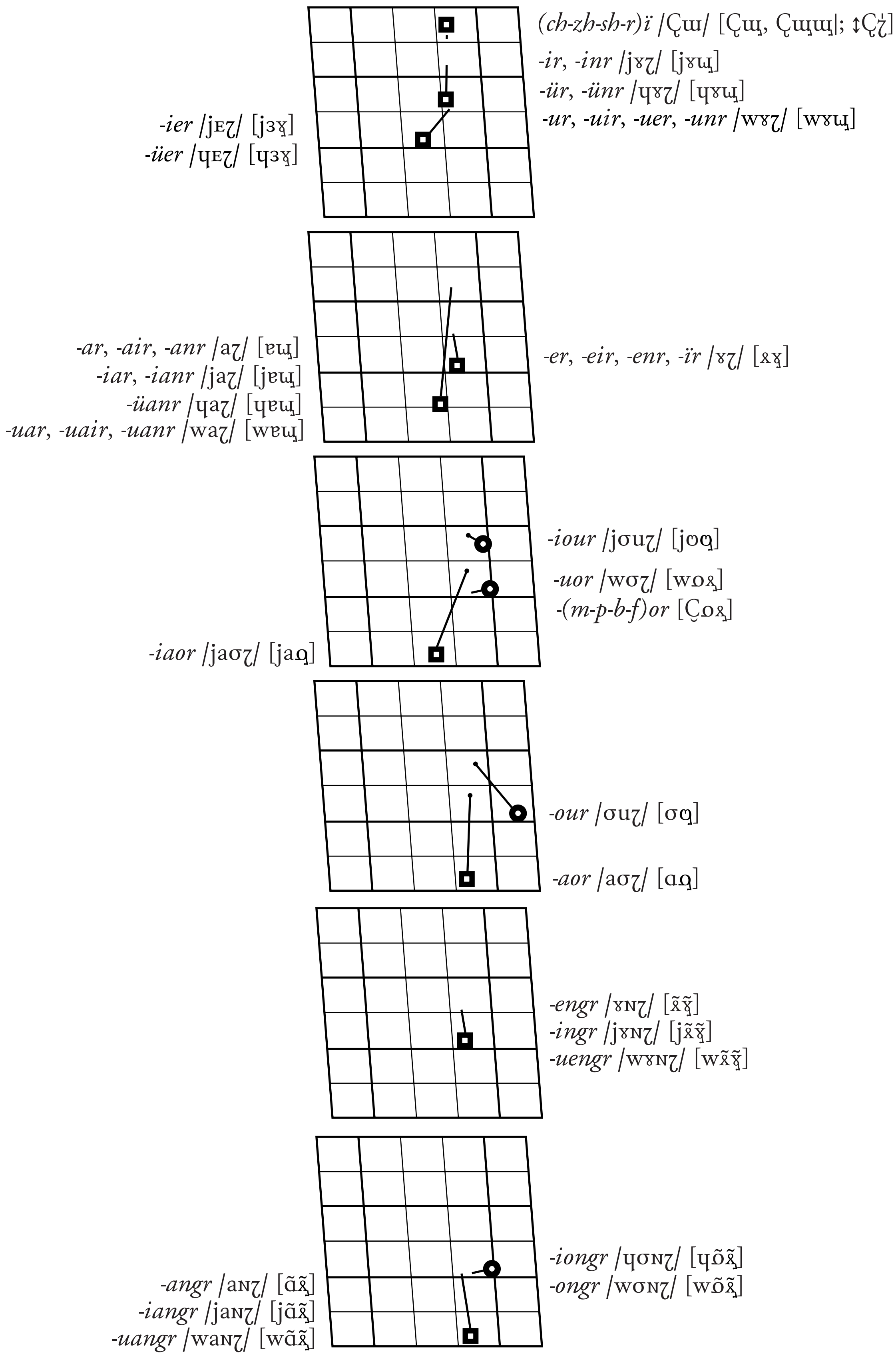
7.6.1. Also *-ng* /ɴ/ is dropped before /ɀ/, but it does not vanish totally; instead, it nasalizes both the preceding vowel and the following /ɀ/. The neutralizations and readjustments explained for ordinary /V(n)ɀ/ largely apply to /Vɴɀ/, as well:

yingr, *ingr* /jɤɴɀ/ [j̃ɤ̃ɤ̯];
angr; *yang*, *iangr*; *wangr*, *uangr* /aɴɀ, j-, w-/ [ɑ̃ɤ̃, j-, w-];
engr /ɤɴɀ/ [ɤ̃ɤ̯];
wengr /wɤɴɀ/ [wɤ̃ɤ̯];
ongr /σɴɀ/ [σ̃ɤ̃];
yongr, *iongr* /ɥσɴɀ/ [ɥʊ̃ɤ̃] — ˢ[ɥʊ̃ʊ̯].

7.7. 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' mediatic pronunciation (cf § 13). 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), also the mediatic taxophones, very often, may come by themselves (to be used mainly as a comunicative understanding task).

A. /aɀ/: [ɐɤ̯] *ar*, *air*, *anr*; [jɐɤ̯] *iar*, *ianr*; [ɥɐɤ̯] *üanr*; [wɐɤ̯] *uar*, *uair*, *uanr*:
/aɀ/ (//a(i)ɀ, anɀ//): *huà-huàr* [ˑʞwa\ʞwɐ̀ɤ̯];
/jaɀ/ (//jaɀ, janɀ//): *chuāngliánr* [ˈˈʈʂʰwɑɴˈljɐɤ̯];
/ɥaɀ/ (//ɥanɀ//): *yüánqüānr* [ˑɥɑɲˉʧʱɥɐɤ̯];

fig 7.4. Vocograms of the laterally contracted vocoids with no simplification.



/waʅ/ (//wa(i)ʅ, wanʅ//): *hǎowánr* [.ʁɑ̞ʊ̞˧wɐɣ˞].

B. /ɤɻ/: [ɤɻ] *er*, *eir*, *enr*, *ïr*; [jɤɻ] *ir*, *inr*; [ɥɤɻ] *ür*, *ünr*; [wɤɻ] *ur*, *uir*, *unr*:

/ɤɻ/ (/‖ɛiɻ, ɤ(n)ɻ, ɯɻ‖/): *bèikér* [\ˌbˌɐˌˈkʰɤɻ];
/jɤɻ/ (/‖i(n)ɻ‖/): *shǐ̈jìnr* [ˌˌʂɯ\tɕjɤɻ];
/ɥɤɻ/ (/‖y(n)ɻ‖/): *qūqür* [¯tɕʰy·tɕɥɤɻ];
/wɤɻ/ (/‖uɻ, wɛiɻ, wɤnɻ‖/): *kǒuwèir* [ˌkhɤˌ\wɤɻ].

Again, *yer*, *ier* & *yüer*, *üer* might be given as mediatic /jɤɻ, ɥɤɻ/ too, yet we prefer to interpret them precisely as /jɛɻ, ɥɛɻ/, thus:

/jɛɻ/: *shǐ̈jiēr* [ˈʂɯ¯tɕjɜɻ];
/ɥɛɻ/: *mùjüér* [\mɯˈtɕɥɜɻ].

C. /ɔuɻ, jɔuɻ/: [ɔɳ] *our*, [jɔɳ] *iur*; /wɔɻ/: [wɔɻ, ɰɔɻ] *(u)or*; /(j)aɔɻ/: [ɑɑ] *aor*, [jaɑ] *iaor*:

/ɔuɻ/: *tǔdòur* [ˌthɯ\tˌɔɳ];
/jɔuɻ/: *zhīliúr* [ˈdʑɯˈljɔɳ];
/wɔɻ/: *shānpōr* [ˈʂʌm¯pʰʊ̟ɔɻ];
/aɔɻ/: *nǎimáor* [ˌnˌɐˌˈmɑɑ];
/jaɔɻ/: *màimiáor* [\mˌɐˌˈmjaɑ].

D. /ɑnɻ/: [ɑ̃ɻ] *angr*; [jɑ̃ɻ] *iangr*; [wɑ̃ɻ] *uangr*:

/ɑnɻ/: *piānchǎngr* [ˈphjɛnˌtʂʰɑ̃ˌɑ̃ɻ];
/jɑnɻ/: *dòujiāngr* [\dˌɔˌ¯tɕjɑ̃ɻ];
/wɑnɻ/: *tiānchuāngr* [ˈthjɛn¯tʂʰwɑ̃ɻ].

E. /ɤnɻ/: [ɤ̃ɻ] *engr*; [jɤ̃ɻ] *ingr*; [wɤ̃ɻ] *uengr*:

/ɤnɻ/: *bǎndèngr* [ˌbˌɐn\tˌɤ̃ɻ];
/jɤnɻ/: *qǐ míngr* [ˌtɕhiˌˈmjɤ̃ɻ];
/wɤnɻ/: *xiǎo wèngr* [ˌɕjɐˌɔ\wɤ̃ɻ].

F. /ɔnɻ/: [ɔ̃ɳ] *ongr*; [ɥɔ̃ɻ] *iongr*:

/ɔnɻ/: *tōukòngr* [ˈthɔʊ\kʰɔ̃ˌɻ];
/jɔnɻ, ɥ-/: *kūqióngr* [ˈkhuˈtɕʰɥɔ̃ɻ].

7.8. Compared to the huge number of the Chinese-speaking population worldwide, just a tiny minority can use *érhuà* natively and properly, being it a peculiarity of Peking and of the Peking area: many accents, including some of the 'northern' group, do not contemplate the *érhuà* phenomenon at all (unless attempting to speak 'well'), or lack the phones necessary to realize it correctly. In other dialects, something comparable to *érhuà* has crystallized in a more conservative phase than that of Peking *érhuà*, or 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*

a 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 and tones, as [ˈ$˳$], but also as [˳$ˈ$]! Other speakers, instead, would articulate only one, but possibly with minor &/or non-standard neutralizations (all not to be imitated), such as:

shùr /\ʂwɤɽ/ *[\ʂu̹.ɣ̽] 'number; vertical stroke in Chinese calligraphy';
xífur /˧si˳fwɤɽ/ *[˙ɕi¯fu·ɣ] 'wife';
xiǎo qiúr /˩sjaσ˧ʦɧjσuɽ/ *[..ɕja̹o̹˧ʨɦjʊʊ·ɣ] 'small ball';
cuòr /\ʦɧwσɽ/ *[\ʦɦwo̹.ɣ̽] 'mistake'.

7.8.1. 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 often 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.

7.8.2. 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 *-zï* 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.

7.8.3. In an 'international' kind of pronunciation, slightly –but reasonably– simplified, when compared to the 'normalized' neutral one, just three sequences might suffice, namely [aɽ, ɜɽ, σɽ], to which any other /V(V)ɽ, Vnɽ/ shall adhere and conform.

For /Vɴɽ/, [ãɽ̃, ɜ̃ɽ̃, σ̃ɽ̃] are fair enough to make oneself understood; but attention should be paid not to omit this nasalization, given that [ãɽ̃, ɜ̃ɽ̃, σ̃ɽ̃] are *not* to be merged with [aɽ, ɜɽ, σɽ].

Let us examine a few examples: *bānr* [¯paɽ] 'job; scar' *vs* *bāngr* [¯pãɽ̃] 'band, gang'; *bèir* [\pɜɽ] 'totally; generation' *vs* *bèngr* [\pɜ̃ɽ̃] '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 7.1. For *(y)ingr*, we recommend [jɜ̃ʐ̃]; while, for *yongr*, *iongr*, and *ongr*, we recommend [jõʐ̃] and [õʐ̃]. It goes without saying that an approximation to native vocoids produces a more authentic result and is, of course, highly recommendable (and certainly appreciated).

A. Related to ^{i}[aʐ], we have: ^{i}[aʐ] *ar*, *air*, *anr*; ^{i}[jaʐ] *iar*, *ianr*; ^{i}[ɥaʐ] *üanr*; ^{i}[waʐ] *uar*, *uair*, *uanr*:

^{i}[jaʐ]: *chuāngliánr* ^{i}[¨ʈʂhwaŋ˩lјaʐ], ^{s}[¨ʈʂhwɑɴ˩ljɐʐ̞];
^{i}[ɥaʐ]: *yüánqüānr* ^{i}[˙ɥɛɲ¯tɕhɥaʐ], ^{s}[˙ɥaɲ¯tɕɦɥɐʐ̞];
^{i}[waʐ]: *hǎowánr* ^{i}[..haʊ˩waʐ], ^{s}[..ʁɑ̹o̹˩wɐʐ̞].

B. Related to ^{i}[ɜʐ], we have: ^{i}[ɜʐ] *er*, *eir*, *enr*, *ïr*; ^{i}[jɜʐ] *ir*, *ier*, *inr*; ^{i}[ɥɜʐ] *ür*, *üer*, *ünr*; ^{i}[wɜʐ] *ur*, *uir*, *unr*:

^{i}[jɜʐ]: *shǐjìnr* ^{i}[..ʂʐ̩\tɕjɜʐ], ^{s}[..ʂɰ̩\tɕ̬jɤ̞ʐ̞];
^{i}[ɥɜʐ]: *qūqür* ^{i}[¯tɕhy·tɕhɥɜʐ], ^{s}[¯tɕɦy·tɕɥʐ̞];
^{i}[wɜʐ]: *kǒuwèir* ^{i}[..khoʊ\wɜʐ], ^{s}[..kho̹ʊ̹\wɤ̹ʐ̞].

C. Related to ^{i}[ʊʐ], we have: ^{i}[ʊʐ] *our*; ^{i}[jʊʐ] *iur*; ^{i}[wʊʐ] *uor*; ^{i}[aʊʐ] *aor*; ^{i}[jaʊʐ] *iaor*:

^{i}[ʊʐ]: *tǔdòur* ^{i}[..thu\tʊʐ], ^{s}[..thu̹\ṭo̹ɑ̹];
^{i}[jʊʐ]: *zhīliúr* ^{i}[¨ʈʂʐ̩˩ljʊʐ], ^{s}[¨ɖ̥ʐɰ̩˩ljoɑ];
^{i}[aʊʐ]: *nǎimáor* ^{i}[..naɛ˩maʊʐ], ^{s}[..nʌ̹ɐ̹˩mɑɒ];
^{i}[jaʊʐ]: *màimiáor* ^{i}[\maɛ˩mjaʊʐ], ^{s}[\mʌ̹ɐ̹˩mjaɒ];
^{i}[wʊʐ]: *shānpōr* ^{i}[¨ʂam¯phwʊʐ], ^{s}[¨ʂʌm¯p̂ɦoʁ̞].

D. Related to ^{i}[ãʐ̃], we have: ^{i}[ãʐ̃] *angr*; ^{i}[jãʐ̃] *iangr*; ^{i}[wãʐ̃] *uangr*:

^{i}[ãʐ̃]: *piānchǎngr* ^{i}[¨phjɛɳ˩ʈʂhããʐ̃], ^{s}[¨phjɛɳ˩ʈʂɦɑ̹̃ɑ̹̃ʐ̞̃];
^{i}[jãʐ̃]: *dòujiāngr* ^{i}[\toʊ¯tɕjãʐ̃], ^{s}[\d̥o̹ʊ̹¯tɕ̬jɑ̃ʐ̞̃];
^{i}[wãʐ̃]: *tiānchuāngr* ^{i}[¨thjɛɳ¯ʈʂhwãʐ̃], ^{s}[¨thjɛɳ¯ʈʂɦwɑ̃ʐ̞̃].

E. Related to ^{i}[ɜ̃ʐ̃], we have: ^{i}[ɜ̃ʐ̃] *engr*; ^{i}[jɜ̃ʐ̃] *ingr*; *uengr*:

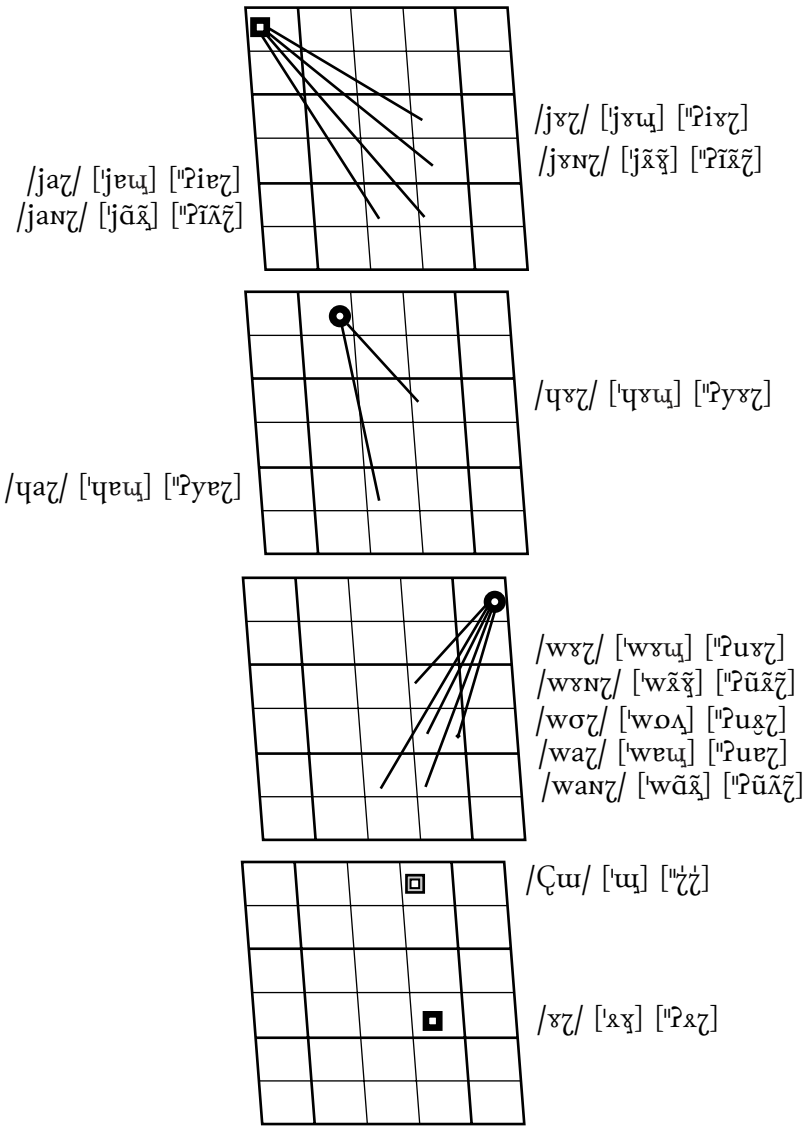
^{i}[ɜ̃ʐ̃]: *bǎndèngr* ^{i}[..pan\tɜ̃ʐ̃], ^{s}[..b̥ʌ̹n̹\ṭɤ̹̃ʐ̞̃];
^{i}[jɜ̃ʐ̃]: *qǐmíngr* ^{i}[..tɕhi˩mjɜ̃ʐ̃], ^{s}[..tɕhi̹˩mjɤ̃ʐ̞̃];
^{i}[wɜ̃ʐ̃]: *xiǎowèngr* ^{i}[..ɕjaʊ\wɜ̃ʐ̃], ^{s}[..ɕja̹o̹\wɤ̹̃ʐ̞̃].

F. Related to ^{i}[õʐ̃], we have: ^{i}[õʐ̃] *ongr*; ^{i}[jõʐ̃] & [ɥõʐ̃] *iongr*:

^{i}[õʐ̃]: *tōukòngr* ^{i}[¨thoʊ\khõʐ̃], ^{s}[¨thoʊ\kɦõ̹ɑ̹̃];
^{i}[jõʐ̃]: *kūqióngr* ^{i}[¨khu˩tɕhɥõʐ̃], ^{s}[¨khu˩tɕɦɥõɑ̃].

7.9. Let us end with fig 7.5, which shows some frequent emphatic (or hyperprecise, or very slow) variants, that can be heard instead of the more normal ones (always with emphatic stress and paraphonics). Also notice the presence of [ʔ], and even /˩V/ [˩V̹V̹] [˩V̹ʔV̹].

fig 7.5. Possible emphatic variants for some /-ɀ/ structures.



10. Chinese tonemes & stress

10.1. 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, thus, 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* and *taxotones*– with as much care and perseverance as possible.

Whether Old Chinese was a tone language since the beginning is still under debate, but it appears plausible that tonemes developed at a relatively early stage of its long history, though not necessarily in its more important stage yet. In fact, 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, with no need to add such a complex mechanism as that of tonemes.

10.1.1. 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 restricted syllabic inventories enriched by the use of tonemes. Otherwise, the lexemic monosyllabism so typical of those languages would simply be impossible.

Although Japanese has its *akusento*, of a clear tonal nature, that language has no tonemes, in the strict sense of the term, *ie* not syllable-tonemes, but just word--tonemes. No surprise, thus, that polysyllabic words abound in purely indigenous Japanese vocabulary (note that [˼] or [˃], at the end of the transcription of an isolated Japanese word, indicates that a possible following grammeme, like *ga* [◦ŋɐ], tonetically is [.ŋɐ] or [·ŋɐ], respectively).

So, we have: *atarashii* [.ɐ-tɐ·ɺɐ·.ɕii˼] 'new', *kokoro* [.kʊ̥-kʊ.ɺʊ˼] 'heart' – *vs* *shin* [˴ɕiŋ̍˼] 'heart; new', the Sinic pronunciation of both the corresponding Chinese characters, as in *shinbun* [˴ɕim̩·bɯ̆ŋ̍˃] 'newspaper', and *shinri* [˴ɕin̩.li˼] 'mentality, psychology; truth' (*cf* Canepari, 2007², Ch 12: *Japanese*, or Canepari & Cerini, *Japanese Pronunciation & Accents* [forth.]). In Chinese, *shinbun* and *shìnri* are *xīnwén* [¨ɕɪnˊwɜn] and *xīnlǐ* [¨ɕɪn˼li̦i̦, ¨ɕɪl̰-, ¨ɕɪl-], respectively.

10.1.2. As far as Chinese 'dialects' –actually different Chinese *languages*– are concerned, some of them count as many as eight tonemes; others have six or five full tonemes. *Pǔtōnghuà* and the Peking 'dialect' have only four marked tonemes, or *shēngdiào* [¨ʂɤɴ\t̪jа̹o̹], collectively called *sì-shēng* [\sɹ̩¯ʂɤɴ] 'four voices'. They are usually 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 (but, since these names do not necessarily reflect modern pitch contours, it is better to avoid them):

– *yīnpíngshēng* [¯jɪm·phɪɴ¯ʂɤɴ, -¨phɪɴ-, -˙pɪɴ-], 'lower level', nowadays rather a steadily *high* and *level* toneme, in neutral Chinese;
– *yángpíngshēng* [ˀjaɴ·phɪɴ¯ʂɤɴ, -¨phɪɴ-, -˙pɪɴ-], 'higher level', nowadays a *mid-to-high rising* toneme (in its basic form);
– *shǎngshēng* [..ʂa̹ɴ¯ʂɤɴ], 'ascending', nowadays nearly always a *low* and slightly *descending* toneme (only seldom with an added rising coda, as its basic form);
– *qùshēng* [\tɕhy̌¯ʂɤɴ], 'departing', apparently coherent with the modern steep, abruptly *high-to-low* falling contour (*cf* Huang & Liao, 2002:78-90; Li & Shi, 1986: Ch 9).

Numerous expedients have been devised to indicate tonemes in Latinized transcriptions; the one 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* /¯ʦhin/ [¯tɕhɪn], and *qiān* /¯ʦhjɛn/ [¯tɕhjɛn];
– T-2, *lóu* /ˀlɔu/ [ˀlɔʊ], but *liú* /ˀljɔu/ [ˀljɵʊ] (instead of a better form like '*lióu*');
– T-3, *hěn* /ˏxɤn/ [ˏxə̹ə̹n̹], but *gǔn* /ˏkwɤn/ [ˏkwə̹ə̹n̹] (instead of '*guěn*');
– T-4, *hèi* /\xɛi/ [\xɛ̹ɪ̹], but *huì* /\xwɛi/ [\xwə̹ɪ̹] (instead of '*huèi*').

10.1.3. 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. Their actual pitches –again– are determined by the tonal environment; but, unlike in destressed syllables, they have no actual relation to any of the marked tonemes. Therefore, another tonal group must be considered: a true *unmarked toneme*. The Chinese call it *qīngshēng* [¨tɕhɪɴ¯ʂɤɴ] '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. (A full treatment of *qīngshēng* can be found in Lu, 2001; also, *cf* He, 2006, vol. a, Ch 6; & Li & Shi, 1986, Ch 11.)

Dealing with Chinese tonemes –and all their relevant taxotones– is decidedly

hard, not just for foreigners whose mother tongues do not use pitch for lexical purposes, but also for many Chinese whose native tongues have different tonal behaviors. 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, in Hénán Province, the locals pronounce T-1 in a very similar way to T-3, while T-2 and T-4 of *pǔtōnghuà* are practically exchanged: thus, they have the 'right' tones, but these are not assigned to the right tonemes, and a sentence like *tā xüé Hànyǔ* 'he/she is learning Mandarin' becomes a puzzling **tǎ xüè Hányū*.

However, controlling pitch to use the tones correctly is not an option: practicing tones 'the hard way' from the start is the only way –if not to fully 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 well worthwhile.

Taxotones & stress

10.2. 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. Unfortunatly, 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 (and boring).

Instead, the four tonemes have three 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!

10.3. For each one of the marked tonemes, there is a basic 'full' tone (*fig* 10.1), to be used in [ˈ$]: [ˉ], [ˊ], [ˬ], [˨] (and three possible taxotones [ˍ], [ˊ], [ˋ], *fig* 10.2, while [ˬ], [˨] are limited to prepausal position (or, to be more precise, when an interruption in the tonal chain intervenes, at the end of a rhythmic unit not semantically and structurally linked to the next one). (Here, too, we might perhaps refer our readers to Canepari, 2007², § 11.3.3.1 *et seq.*, with some enrichments and improvements.)

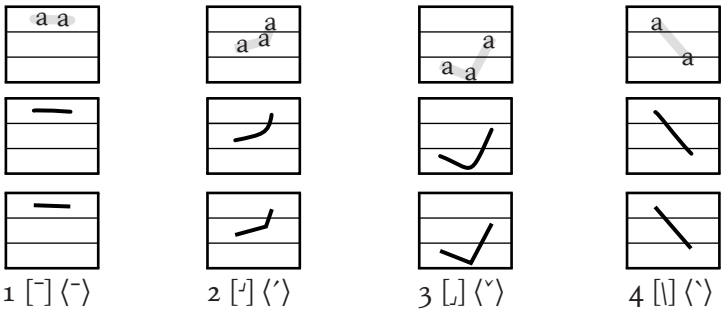
Let us observe, once and for all, that a simple vowel, in fully stressed free syllables (especially in a tune), has different degrees of length depending on the tone it has. Thus, with a full T-3, the vocoid is geminated, [ˬV̩V̩], and we always show it like that (although a finer indication might be ⟦ˬV̩·V̩⟧, *ie* a shorter second element, after a half-lengthened first one). In the same context, a diphthong has the following length pattern: [ˬV̩V̩V̩] (or ⟦ˬV̩·V̩V̩⟧).

In addition, let us note, instead, that a full T-4, generally has a short vocoid, [\V̩], as its narrower taxotone, [ˋV]. But, full T-1 or T2 usually have a half-lengthened vocoid [[¯V·, ˊV·]], although we do not show them in our transcriptions, [¯V·, ˊV·] (and the same happens to fully stressed taxotones of T-3, [_V̩] [[_V̩·]] and [ˇV·] [[ˇV·]]). Lastly, a T-0 has no lengthening at all, [·V] (on the contrary, it often can have a shorter vocoid, which we only show here, [[˙V, ·V, ̩V]]).

In *faster speech*, the degrees of length we showed above are generally less evident, so that instead of [·] the sign [[.]] might be more appropriate, since it indicates something halfway between [·] and [ ] (short).

It goes without saying that the emphatic counterparts of the four tones and their taxotones, [=], [˩], [˩̩], [_=], [ˊ], [\\], [ˋ], in comparison, have stronger stress, ["$], and greater length, which we do not show here, also because it can have different realizations).

fig 10.1. The four basic tonemes (iconic, realistic, and schematic).



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*, which produce a rather complex and composite pitch contour. Taxotones (see *fig* 10.2) are definitely not an unnecessary nuisance: without them, it would be uneasy and tiresome for both speakers and listeners.

In syllables with a secondary stress, the tones are somewhat reduced, both in duration and in intensity. Thus, we have [ˌ$]: [¨], [·], [.], [ˌ] (not considering here other taxotones modified by intonation). Since the 'zero' toneme is regularly associated with [◦$], its actual pitch contour is so feeble that it can be represented by a single dot, and only three signs are necessary: [˙], [·], [.] (see *fig* 10.8-9); but, in *slow speech* and in a tune, T-0 *may* receive more phonatory energy and substantially merge with [ˌ$]: [¨], [··], [..].

To be entirely rigourous, a notation like '[ˌ]' would suffice to express a secondary stress with medium pitch, but we prefer '[··]' to keep consistency with [¨], [..]; besides, in non-tone languages, [ˌ] may be completely unrelated to pitch. For example, in (British) English, *Mexico* is actually [ˈmɛksɪˌkɜʊ] (diaphonemically /ˈmɛksɪkʊʊ/) But what is the actual pitch of the three syllables, when the proper intonational patterns are applied? Considering the four British neutral tunes, one would hear (Canepari, 2015, ℂ 34-37):

– *conclusive*: /ˈmᴇksɪkσɷ./ [ˈmᴇksɪ˻kɜɷ̠], *ie* ⟦ˈmᴇḳsɪ̤k3ɷ⟧,
– *interrogative*: /ˈmᴇksɪkσɷ?/ [ˈmᴇksɪ˻kɜɷ·˙], *ie* ⟦ˈmᴇk·sɪ˙kɜɷ⟧,
– *suspensive*: /ˈmᴇksɪkσɷ;/ [ˈmᴇksɪ˻kɜɷ̣·], *ie* ⟦ˈmᴇḳsɪ··kɜɷ⟧,
– *continuative*: /ˈmᴇksɪkσɷ,/ [ˈmᴇksɪ˻kɜɷ·], *ie* ⟦ˈmᴇk·sɪ··kɜɷ⟧.

10.4. 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 phono-articulatory energy– and a number of less prominent ones.

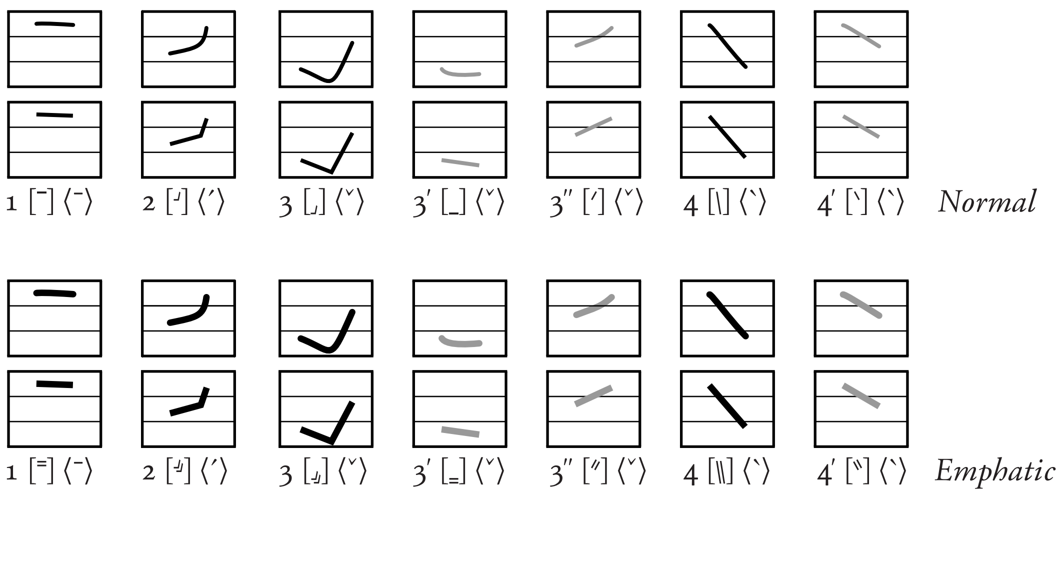
All the more so in Chinese, where even the very concept of 'word' is particularly vague; and words of more than two syllables have more than one strong stress, although, in these cases, the stronger stress is the last one.

Let us now examine some rhythmic groups containing marked tonemes only. First of all, a disyllabic sequence most commonly follows the [ˌ$ˈ$] pattern: *zhīyüán* [¨ʤ̥ɻɯ˻ʮɑn] 'employee', *liánxì* [·ˈljɛn̨\ɕi̥] 'to link', *hēibǎn* [¨ʁᴇɪ˼ṗʌ̥ʌ̥n̥] 'blackboard', *jiàoshī* [ˌ\ʥ̥jao̥ˉʂɯ̨] 'teacher'; *wǒ shuō* [..wo̜ˉʂwσ] 'I say', *huài háir* /\ɤwaᴇ˻ɤaʒ/ [ˌ\ʁwa̧ɜ̯˻ɤɐɤ] 'naughty boy'.

A useful comparison between the basic tones and three of their very important taxotones is clearly shown in *fig* 10.2. Attention should be paid to *ǎ* /˼/ [ˊ], a T-3 that closely resembles a T-2 (though not identical, as too many textbooks still describe it, instead), and occurs only before another T-3; thus, the so-called 'half T-3' is *ǎ* /˼/ [_], when preceding some other marked tonemes.

Finally, there is even a 'half T-4', *à* /\/ [ˋ], 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 [ˋ] in particular, are nothing but a 'condensed' version of their corresponding full forms, [˼] and [\].

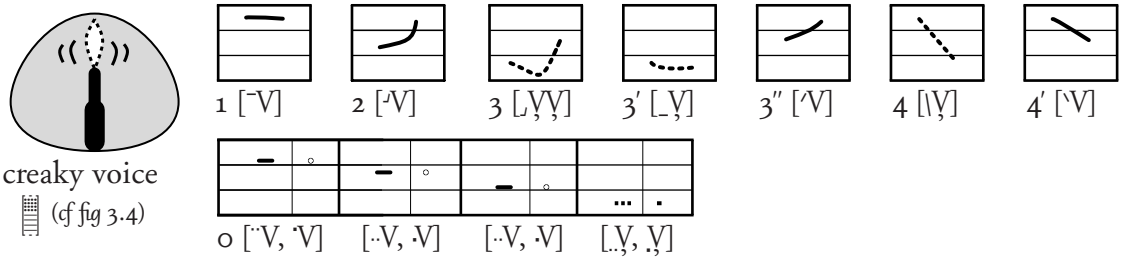
fig 10.2. Tones and taxotones.



Of all the (taxo)tones, including those in weak and reduced syllables, *fig* 10.3 shows which of them are uttered with *creaky voice*, when the low tonal band is involved. Of course, in creaky voice (or laryngealization, [m̩, a̰]), the vocal folds are less tense and less thin than usual, whereas the arytenoids remain firmly closed, so that only the part of the folds not in contact with them are free to vibrate. The folds are completely lax, and produce a series of quick glottal taps and a pitch impression which is lower than usual.

We can find creaky voice in some other languages, in conjunction with low pitch, as at the end of certain conclusive intonemes, which are falling in conjunction with the low parts of tunes. In British English, creaky voice on a conclusive tune is considered as upper-class: *Yes, indeed* [¯jɛs ɪnˈɖr̩ḭɖ..].

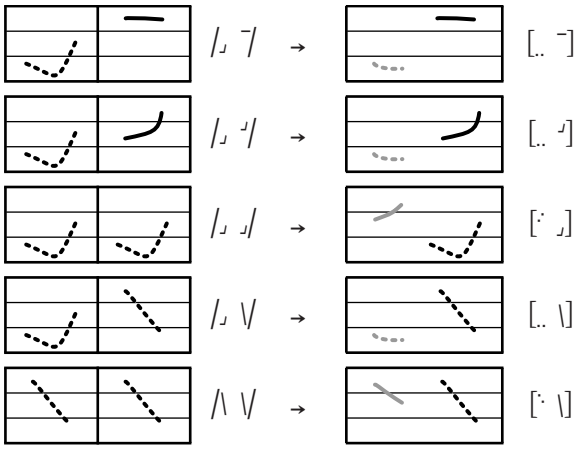
fig 10.3. Regular and creaky voice & tonetic elements.



10.5. It might be argued that the 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, these should anyway be able to recognise 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.

A better explanation of how the taxotones behave is provided by *fig* 10.4, as well as by the following examples, where secondary-stress reduced taxotones are also employed. It is clear that [˩] → [_] (& [..]) allows for a simpler and smoother transition

fig 10.4. Fundamental taxotones in disyllabic sequences.



from /˩/ to /˥/, or /˧˥/, or /˥˩/: *huǒchē* [.ʀwo̞¯ʈʂɦɤ], *děngyú* [.dɤŋ˧˥ɥy], *kǎoshì* [.khɑ̞o̞˥˩ʂɥɻ].

Similarly, a sequence like /˩$˩$/ would not be practical, if realized as such, or even with the taxotone just seen, *ie* '[.$˩$]'; instead, [ʹ] decidedly improves the enunciation: *shǒubiǎo* [˙ʂσʊ˩pjɑ̞ɑ̞o̞] (again, with a secondary stress and reduced duration: [ʹσʊ] → [˙σʊ], instead of [˩o̞o̞ʊ̞]).

In /\$\$/, there is a generally descending contour: of course, a theoretical '[\$\$]' would require an uncomfortable 'zig-zag' movement, should we actually try to make our voice 'climb up' (from [.] to [˙]) before falling down again. Instead, with [ˋ], a very straightforward [˙$\$] is obtained: *fàng jià* [˙faɴ\tɕja̞].

In short, *seven tones* and as many signs are required for [ˈ$]: [¯a], [˧˥a], [˩ɑ̞ɑ̞, _a̞, ʹa], [\a̞, ˋa]; but only *five*, for [ˌ$]: [¨a, ˙a, ˙a, \a̞, ..a̞] (because [˧˥] & [ʹ] → [˙]; [˩] & [_] → [..]); and *three*, for [◦$]: [˙a], [·a], [.a̞].

(Despite remarkable differences, in terms of tonetic notation and iconic representation of intonation, our approach has some points in common with what W. Cao presented in a simple booklet, *Hanyu yuyin jiaocheng* – A Course in Chinese Phonetics, 2002, Peking: Beijing Yuyan Daxue Chubanshe, almost *IPA*, especially § 10.4.1-4. He also used the well-known 'transitional tone', *guòdù shēngdiào*, for some occurrences of [¨] instead of [˙] – already clearly shown by Chao, 1948:111. However, it is surprisingly sad that in more recent booklets of his everything has been simplified, 'downgrading' the whole treatment of tonemes and tones to nothing more than a sketchy resumé.)

10.6. Generally speaking, trisyllables composed by marked tonemes follow the [ˈ$ˌ$ˈ$] pattern, which is substantially coherent with that of disyllables, as it can be analyzed as '[ˈ$]+[ˌ$ˈ$]'. 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 a tune: compare *Mexico* [ˈmɛksɪˌkɜɷ] *vs Mòxīgē* [\mo̞¨ɕi¯kɤ̞]. Moreover, both [\mo̞] and [¯kɤ̞] bear a primary stress, though [\mo̞] is auditorily less prominent, for intrinsic reasons, also due to the influence of rhythm. However, it is true that in *fast speech* we generally find [\mo̞¨ɕi¯kɤ̞], where [\] indicates a slightly lighter strength than [\].

Other examples: *yōumògǎn* [¯jɷʊ\mo̞˩kʌ̞ʌ̞n̞] 'sense of humour', *féngrènjī* [˧˥fɤɴ\ʐə̞n̞¯tɕi] 'sewing machine', *huǒchēzhàn* [_ɹwo̞¨ʈʂhɤ\ʈʂʌ̞n̞] 'train station', *fùmǔ shuō* [\fʊ̞..mʊ̞¯ʂwσ] '(my) parents say'. It is very interesting to notice that the primitive lexemes, if taken in isolation, would switch from [ˈ$ˌ$] to the [ˌ$ˈ$] pattern: *yōumò* [¨jɷʊ\mo̞] 'humor', *féngrèn* [˙fɤɴ\ʐə̞n̞] 'to sew', *huǒchē* [.ʀwo̞¯ʈʂɦɤ] 'train', *fùmǔ* [\fʊ̞˩mʊ̞ʊ̞] 'parents'.

10.6.1. Beside the norms explained hitherto, which occur inevitably and almost irrespective of enunciation speed (as we saw), there are other particular cases, involving [˙], where a certain degree of fluency and speed is required for the tonal adjustment to take place: [˙] → [¨].

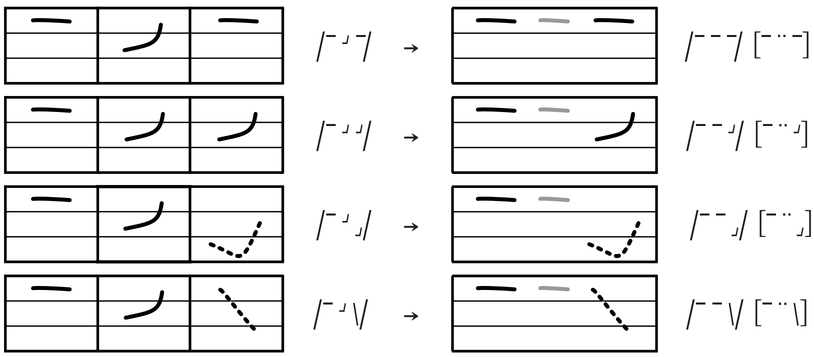
In fact, native speakers are rarely conscious of converting a T-2 or a T-3 into a T-1 (actually, the overlapping is taxotonic, not tonemic): for them, that is a perfectly natural outcome, in their speech habits. But, if asked to repeat the same sequence

at a slower pace, paying attention to its taxotones, they invariably avoid [ˑ̈] in favour of the original tones.

Foreign learners must pay attention to this last point, and may even want 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 [ˑ̈], though in the right place, might result puzzling and be judged –paradoxically!– as wrong.

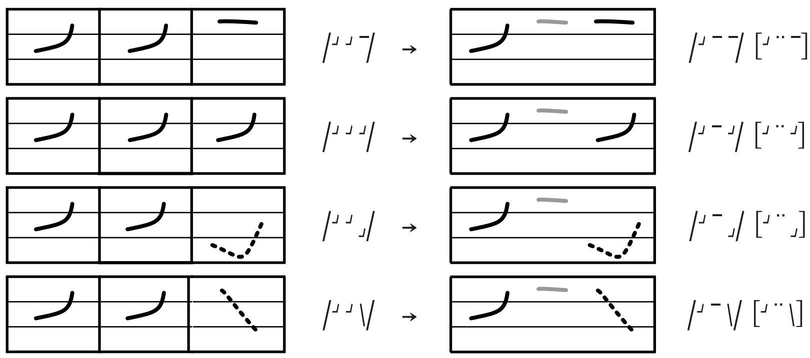
First of all, let us examine *fig* 10.5, presenting the behavior of /¯$˴$ˈ$/, *ie* /¯$˴$/ followed by any other marked toneme: [¯$˙$ˈ$] → [¯$¨$ˈ$], *eg xiānrénzhǎng* [¯ɕjɛɳ-¨ʐɜɳ˛ʈʂa̰a̰ɳ̰], *dōngnánfēng* [¯ṭoɴ¨nʌɱ¯fɤɴ], *sānniánjí* [¯saɲ̟¨ɲ̟jɛɲ̟˴ʨi], *xīhóngshì* [¯ɕi-¨ʀoɴ\ʂɰ̰ɻ̰]. As said, in slower speech, a slightly less spontaneous but 'safer' realization would resume [-˴ʐɜɳ-], [-˴nʌɱ-], [-˴ɲ̟jɛɲ̟-], [-˴ʀoɴ].

fig 10.5. Taxotones in /¯$˴$/+/ˈ$/ trisyllabic sequences.



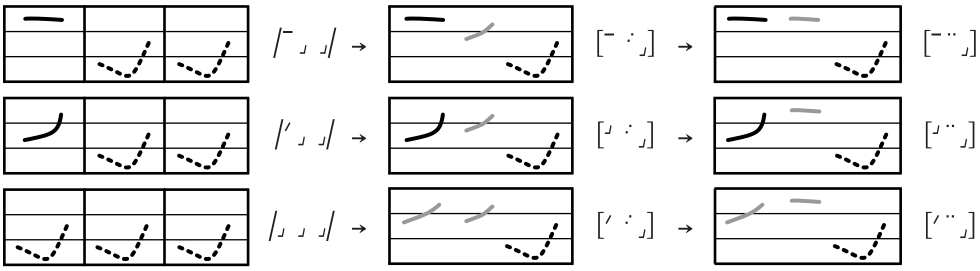
Besides, *fig* 10.6 presents the behavior of /˴$˴$ˈ$/ [˴$˙$ˈ$] → [˴$¨$ˈ$], *eg Méi Lánfāng* [˴mɛɪ¨lʌɱ¯faɴ], *wánqüán dǒng* [˴wʌɲ̟¨ʨhɥɑn˛ṭo̰o̰ɴ̰], *hái méi lái* [˴ʞʌɜ¨mɛɪ˴lʌɜ], *xǘnyáng-jiàn* [˴ɕʏɲ̟¨jaɴ\ʨjɛɲ̟].

fig 10.6. Taxotones in /˴$˴$/+/ˈ$/ trisyllabic sequences.



10.6.2. Finally, *fig* 10.7 describes what happens when a /˛$˛$/ sequence *follows* one of the first three marked tonemes: [ˈ$..$˛$] → [ˈ$˙$˛$] → [ˈ$¨$˛$] (with [ˈ$] ≠ [\$]), *eg tā hěn hǎo* [¯tʰa¨ʀɜɴ˛ʞa̰o̰o̰], *hánshǔbiǎo* [˴ʞʌɳ¨ʂu˛p̰ja̰a̰o̰], *wǒ yě yǒu* [ˊwʊ¨jᴇ˛jo̰o̰ʊ̰] (be noted: [ˊwʊ], not '[˴wʊ]', depending on rhythm and emphasis, the alternative reading [ˍwʊ̰¨jᴇ˛jo̰o̰ʊ̰] is possible as well).

fig 10.7. Taxotones in /ˈ$/+/ˌ$ˌ$/ trisyllabic sequences.



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 phono-articulatory energy on the 'outer ends', and to reduce it in the middle.

This means that, in a four-syllable sequence, the usual distribution is [ˈ$ˌ$ˌ$ˈ$], rather than [ˌ$ˈ$ˌ$ˈ$], which would not sound very smooth. However, depending on the semantic value of each component, also [ˌ$ˈ$ˌ$ˈ$] is possible, especially when a contrastive purpose on the first bisyllabic lexeme is present, which may certainly tend to its original pattern [ˌ$ˈ$].

But with five syllables, there is no reason to subvert the [ˈ$ˌ$ˈ$ˌ$ˈ$] pattern resulting from the simple juxtaposition of [ˈ$ˌ$ˈ$] + [ˌ$ˈ$], or the spontaneous adjustment deriving from [ˌ$ˈ$] + [ˈ$ˌ$ˈ$] → [ˈ$ˌ$ˈ$ˌ$ˈ$] (but, of course, lexical factors play their role too, when combining the syllables of their words).

A practical example: taken separately, *pīnyīn* 'to spell words out (phon[et]ically)' and *zı̀mŭ* 'alphabet' are both naturally [ˌ$ˈ$]: [¨phɪŋ¯jɪn], [\d̥zɰ̩˩mṳṳ]. By combining them into *pīnyīn-zı̀mŭ* 'phono-alphabetic script, romanization', the theoretical (yet possible) rendering [¨phɪŋ¯jɪn \d̥zɰ̩˩mṳṳ] normally changes into [¯pɦɪŋ-¨jɪn \d̥zɰ̩˩mṳṳ]. In English we have examples like *the seven-year itch* [ðə¯sᴇv-ənˌjɪəɻ ˈɪʧ] or *fifth-generation computers* [¯fɪfθ ʤᴇnəˈɻᴇɪʃn̩ kɯmˈphjµuɾəz̥], which combine their plain elements: *seven* [ˈsᴇv-ən], *year* [ˈjɪɐ], *itch* [ˈɪʧ] and *fifth* [ˈfɪfθ], *generation* [ˌʤᴇnəˈɻᴇɪʃn̩], *computer* [khɯmˈphjµuɾɐ] (when given in isolation).

Keeping in mind all these observations, we must admit that, in *fast speech*, the patterns [ˈ$ˌ$ˌ$ˈ$] and [ˈ$ˌ$ˈ$ˌ$ˈ$] may certainly tend to become [ˈ$˳$ˌ$ˈ$], [ˈ$˳$ˈ$ˌ$ˈ$] and also [ˈ$˳$ˌ$ˈ$], [ˈ$˳$ˌ$ˌ$ˈ$], with [ˈ] veaker than [ˈ]).

10.6.3. It has to be underlined, both from the learner's and teacher's point of view, that a rhythmic group is 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 zhŭxí* [˩mɑɒ¨d̥zɤ¯ʈoɴ ‥d̥ʐɰ̩˩ɕi] 'chairman Mao Zedong', with simple juxtaposition: [ˈ$ˌ$ˈ$] + [ˌ$ˈ$]. But, in the basic sentence *tā shì zhŭxí* '(he/she) is (the) chairperson', we have [ˈ$ˌ$+ˌ$ˈ$]: [¯tɦa\ʂɰ̩ɻ ‥d̥ʐɰ̩˩ɕi]; in fact, [¨tha\ʂɰ̩ɻ ‥d̥ʐɰ̩˩ɕi] (omitting sentence intonation symbols) would not be a mere variant –as in *pīnyīn-zı̀mŭ*– but would imply that 'he/she *is* (the) chairperson', though with less assertiveness than [¨tha\\ʂɰ̩ɻ¦ ‥d̥ʐɰ̩˩ɕi] '(he/she) *definitely is* (the) chairperson', with an

extra-strong, or emphatic, stress on *shì̈*, the copula (and a very short pause, [¦]).

On the other hand, one might say *Lǐ fùzhǔxí* [_lị \fụ..ḍ̥ʐụ˩çi] ‘(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). In fact, we have *Lǐ fùzhǔxí* [_lị· \fụ..ḍ̥ʐụ˩çi], with a continuative tune, [_lị·], although with no pause. The same is normally true of *Máo Zédōng zhǔxí* [˩mɑo¨ḍ̥zɤ¯ṭoɴ· ..ḍ̥ʐụ-˩çi], in actual speech.

10.7. A Chinese bisyllable 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*, *érzï*, *rènshï*, *luóbo*, *zhàogu*, *méigui*, &c. Such typographical convention is mandatory, and (good) teachers would reject such misspellings as **érzï̌*, **rènshḯ*, and **méiguī*.

There are also some *de-stressed* syllables, /ˈ$/ [ˈ$] → [˚$], *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.

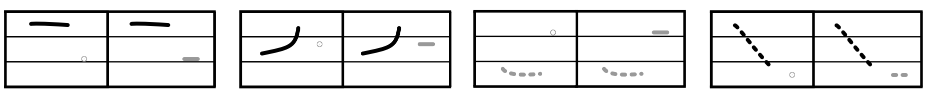
The official orthography is somewhat inconstant, on this respect: de-stressing is sometimes explicitly annotated, more often it is not (pretty much like tonal adjustments, which are given as implicitly understood by the reader/speaker). A few examples, duly transcribed, are given below.

As a general rule, the actual pitch of a ‘weak’ syllable is determined by the tone of the syllable that comes before it. In addition, in *fast connected speech*, the pitch of a following syllable may exert a certain influence as well, but always on a secondary level. As previously said, three symbols are enough: [˙ · ̣]; but, phonetically, *four* different pitches may be detected, as shown in *fig* 10.8: *jīzï* [¯ṫçi·dzɯ] ‘chicken’, *jízï* [˩ṫçi·dzɯ] ‘collected works’, *jǐzï* [_ṫçị¨dzɯ] ‘muntjac’, *jìzï* [\ṫçị.dzụ] ‘bun’.

In *slow speech* and in a tune, as said in § 10.3, the last syllable may receive a secondary stress, [ˈ$ˌ$], so that it becomes slightly more prominent than it usually is: *eg* [¯ṫçi··ḍ̥zɯ], [˩ṫçi··ḍ̥zɯ], [_ṫçị¨ḍ̥zɯ], [\ṫçị..ḍ̥zụ]. 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* 10.8, 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* [¯ṗi·p̂σ, -··p̂h̬σ], *píqi* [˩pɦi·tçi, -··tçhi]. After a T-3, which is automatically

fig 10.8. Taxotones in /ˈ$/+/◦$/ bisyllabic sequences.

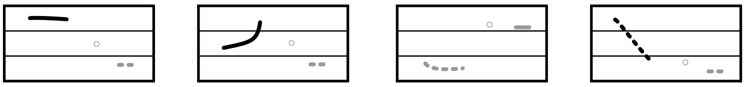


narrowed to [ˍ] and shortened in its length, we have a *lowered high* pitch. Finally, after a T-4, the 'dot' lies just in the middle of the *low* band: *bǐfang* [ˍp̣ị˙ɣ̥aɴ, ˗˙faɴ], *bìhui* [\p̣ị̣ɦwə̣ɪ̣, ˗.ʀwə̣ɪ̣].

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 &/or compressed, and many nuances may become less peculiar, especially in current speech. Even more so, then, to 'get by' with just three punctiform symbols ([˙·.], which are square, in our transcriptions, not round as the intonation dots [˙·.] – although not so clearly distinguishible).

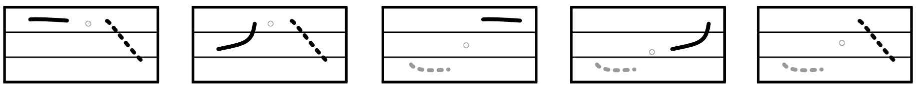
10.8. When a marked toneme is followed by *two* unmarked ones, /ˈ$◦$◦$/, the phonatory energy is redistributed as [ˈ$◦$ˌ$], akin to the general trend of reducing stress in the middle –as observed in /ˈ$ˈ$ˈ$/ [ˈ$ˌ$ˈ$]– while keeping most strength on the edges. As fig 10.9 clearly explains, we have *kāilai le* [¯kɦaɜ·lʌɜ..lɤ̥(ʀ̥|)], *nálai le* [˩na·lʌɜ..lɤ̥(ʀ̥|)], *zǒulai le* [ˍʦọʊ̣˙lʌɜ˙lɤ(ʀ|)], *sònglai le* [\sọŋ̣lʌ̣ɜ̣..lɤ̥(ʀ̥|)], with /˩laᴇ/ destressed to /◦laᴇ/ (see below).

fig 10.9. Contextual pitch of /◦$/ in the middle of some trisyllabic sequences.



Between two marked tonemes, a T-0 (fig 10.10) 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* [¯ʈʂɦɯ˙lʀ\fạṇ], *pá-bu-shàng* [˩pɦa˙bu\ʂạŋ̣], *wǒde shū* [ˍwọ·dʀ¯ʂu], *liǎngge rén* [ˍljạŋ̣·gʀ˩ʐɜn], *kàn-de-jiàn* [\kɦạṇ.dʀ\ʨjẹṇ] (& [ˋkɦan·dʀ-]).

fig 10.10. Contextual pitch of /◦$/ in the middle of the most frequent trisyllabic sequences.



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

Resultative complements are most often destressed. Taking *kàn* [\kɦạṇ] 'to see', and *qǐ* [ˏʨɦịị] 'to start; to lift', as main verbs, we can add an enclitic verb to them, like *jiàn* [\ʨjẹṇ] 'to see', or *lái* [˩lʌɜ] 'to come'. So, we obtain *kànjian* [\kɦạṇ̥.ʥjẹṇ, ˗..ʥjẹṇ] 'to see' and *qǐlai* [ˍʨɦị˙lʌɜ, ˗˙lʌɜ] 'to rise up, to commence' (and other meanings). But, if we insert *dé* [˩ṭɤ] 'to reach' or *bù* [\p̣ụ] 'not' between the verb and its resultative complement, a potential construction is obtained, where *dé* and *bù* lose

their stress, [◦dɤ̞, ◦bu], while the other verb resumes the original toneme: *kàn-de--jiàn* [\kħʌ̹n̹̊.dɤ̞\tɕ̇jɛ̹n̹], *qǐ-bu-lái* [_tɕħi̹·bu˧lʌɜ].

Let us also compare *shōu* + *qǐlái* + *le* → *shōuqilai le* [¯ʂσʊ·tɕi·lʌɜ··lɤ(ɤ̞|), -··lʌɜ·lɤ̞] 'to have collected scattered things'. If *le* comes after the main verb, the sequence *shōu* + *le* + *qǐlái* gives *shōule qǐlái* [¯ʂσʊ·lɤ̞ ..tɕħi̹˧lʌɜ, _tɕħi̹˥lʌɜ]. A monosyllabic pronoun like *tā* 'he, him; she, her; it' might be inserted too: if fully stressed, we would have *shōuqi tā lái le* [¯ʂσʊ·tɕi¯tħa ˧lʌɜ·lɤ̞], or /-◦laᴇ-/ [¯ʂσʊ·tɕi ¯tħa·lʌɜ··lɤ̞] (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 /¯ʂσu◦tshi◦tha ˧laᴇ◦lɤ/ [¯ʂσʊ·tɕi··tha ˧lʌɜ·lɤ̞].

fig 10.11. Contextual pitch of /◦$/ in the middle of all possible trisyllabic sequences.

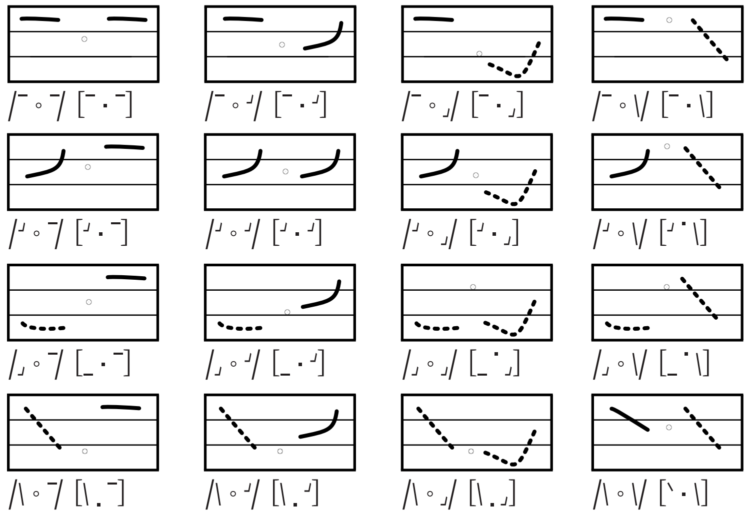
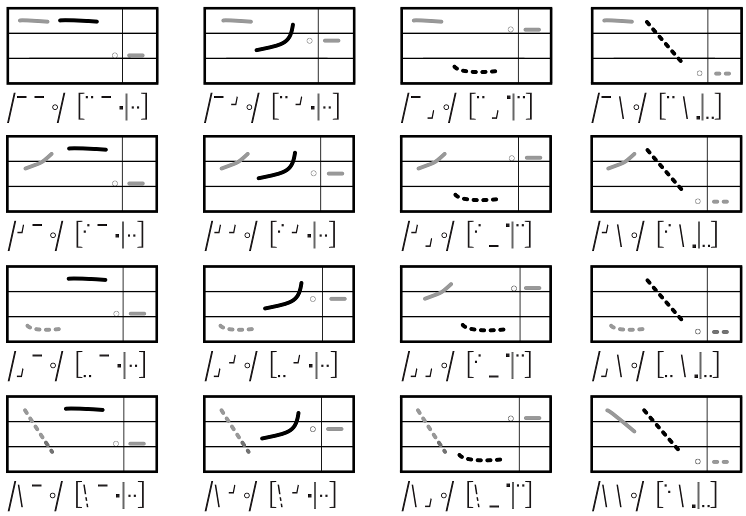


fig 10.12. Contextual pitch of /◦$/ at the end all possible trisyllabic sequences.



10.9. Let us add that the copula word *shì* is normally reduced to /◦ʂɯ/, although always written with its tone mark. It is followed by a complete inventory of three syllable structures, with a T-0 in the middle.

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* /\pu/ changes into /ˈpu/, producing a smoother combination: *wǒ bù qù* 'I don't go' /ˏwɑ\pu\tshy/ → /ˏwɑˈpu\tshy/ [ˍwǫˑb̥u\tɕɦy̌] (& [ˍwǫ·bu\tɕɦy̌], with loss of stress). For a moderately stronger assertion, like 'I do *not* go', its spontaneous pronunciation would rather be [‥wǫˈṗu·tɕy].

However, if the speaker wants to emphasize that he/she has '*absolutely no* intention to go', *bù* may retains its toneme, possibly with two distinct rhythmic units and a potential pause, *eg* /ˏwɑ\\pu. \tshy./ [ˍwǫ\\ṗụ˙¦ \tɕɦy̌˙] (our intonation symbols will be fully explained in Ch 11 – but let us notice again that intonation dots are round, [˙ · .], while tone dots tend to be square, [˙ · .]). 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* [ˍt̥ǫŋ̣ˈbuˏt̥ǫǫŋ̣]; *chī mei chī* [¯tʂɦɯ̣·mᴇɪ¯tʂɦɯ̣, -ˈmᴇɪ-].

The numeral *yī* 'one' keeps its original T-1 when the lexeme is used alone, or at the end of an utterance: *shíyī* [ˈʂɯ̣¯ji] 'eleven', *yī yī dé yī* [¨ji¯ji· ˈd̥ɤ¯ji˙] 'one multiplied by one gives one' (again, please, refer to Ch 11 for an explanation of our intonation symbols). Before a T-4, *yī* becomes /ˈji/: *yīgòng* [ˈji\k̥ǫŋ̣] 'altogether', *yī bàn* [¨ji\ṗᴀŋ̣] 'a half'. Before other tonemes, instead, *yī* changes into /\ji/: *yīdiǎnr* /\jiˏtjaɀ/ [\jịˏtjẹẹɀ̣] 'a bit', *yī tiān* [\jị¯tɦjᴇn] 'a day', *yī nián* [\jịˈɲjᴇn] 'a year'. For teaching purposes, more realistic spellings are found, though they are not systematically encountered in official orthography: *yígòng*, *yí bàn*, *yìdiǎnr*, *yì tiān*, *yì nián*.

Traditionally, *qī* 'seven' and *bā* 'eight' used to be pronounced as /ˈtshi, ˈpa/ before a T-4: *qīyüè* [ˈtɕhi\ɥẹ] 'July', *bā-wèi lǎoshī* [ˈb̥a\wẹɪ̣ ‥lạọ¯ʂɯ̣] 'seight 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* [¨tɕhi\ɥẹ], *bā-wèi lǎoshī* [¨b̥a\wẹɪ̣ ‥lạọ-¯ʂɯ̣] (*cf The Contemporary Chinese-English Dictionary*, p. 26 & p. 1508.).

10.10. Lastly, we must always keep in mind what is said about Chinese stress (in § 12.0.2): rules are rules, but nothing more. In fact, since –in Chinese– stress is not distinctive, its 'simple and clear' stress patterns are not necessarily 'respected', when speaking fluently.

This produces a more natural way of speaking, although foreign learners may certainly find it rather bewildering.

10.11. Of course, as it happens in some of the regional accents dealt with in Ch 16, we can also find even *five* heights of level tones: [¯ ¯ - ˍ _]. The use of [] is quite sufficient to clearly distinguish [¯] from [¯], and [_] from [ˍ], or [-] both from [¯] and [ˍ] (and from [-], as well). But, writing by hand, especially when taking quick notes, it is certainly safer to use [˙], or [˙], for [¯], and [.], or [.], for [ˍ], and [-] for [-].

13. Mediatic Mandarin/Pekingese pronunciation

The Peking *accent* vs Peking *dialect*

13.1. We have carefully considered whether to include analyses of all possible regional pronunciations in a book that specifically focuses on *neutral* Chinese pronunciation. That would be the logical continuation of this work (according to the series of books on the pronunciation & accents of various other languages – *cf* § 1.1). 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. Based on these recordings (also collected with a special questionnaire) of several informants and numerous other audio files, we are forming a rich sound database.

On the other hand, 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 almost an independent language, with its phonetic, lexical and grammatical peculiarities.

Surely, there is no sharp boundary between *accent* and *dialect* in 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 contexts and linguistic styles.

13.2. 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, the local pronunciation(s) will evolve more quickly, making it harder to define a univocal pattern for describing how a 'Peking accent' sounds like today.

Let us tell in advance that *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 – or, in the West: B. Hou) 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 – ie Q. Guo), 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.

The old-fashioned pronunciation, consecrated by *xiàngshengr* crosstalk masters as *Hóu Bǎolín* and *Guō Qǐrú*, 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. Their 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 the 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 is 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 in audio files attached to contemporary Chinese courses. Also the records of the first *Chinese Course* by the Linguaphone Institute offer abundant examples of Peking-accented pronunciation.

Therefore, and contrary to popular belief, the 'traditional' pronunciation reported above has little to do with the 'lazy mutter' most often –but unjustly– associated with speakers from the *hútòngr* alleys.

The *hútòngr* are ancient narrow streets, or alleys, most commonly associated with the oldest and most traditional parts of Peking. In that city, *hútòngr* 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 *hútòngr*, and joining one *hútòngr* to another. The word *hútòngr* is also used to refer to such neighbourhoods.

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

Hútòngr pronunciation is infamously known for its anarchical use of *érhuà* (see Ch 7), and the frequent reduction –if not 'annihilation'!– especially 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'. But, 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.

13.3. Upon suggestions from our informants, we managed to determine three samples of what can be honestly targeted as a 'modern' Peking accent (again: not *dialect*). First of all, Mr *Jiāng Wén* (1963- – *ie* W. Jiang), a well known movie star and director; in addition, Ms *Yáng Lán* (1968- – *ie* L. Yang), a successful entrepreneur and talk-show hostess; finally, Mr *Zhāng Bīn* (1969- – *ie* B. Zhang), 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; thus, 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- – *ie* K. Jang), a master of *xiàngshengr*, who, on the contrary, habitually speaks using the real Pekingese dialect (even if it is no longer that of *Hóu*).

13.4. 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 the last sections of Ch 6.

This is not surprising, as anyone will easily notice, by comparing the language spoken daily among relatives, friends and colleagues, with that generally adopted in dubbing foreign movies and documentaries.

The typical *buzz*, accompanying Pekingese Mandarin, is rather a consequence of the

behavior of /Çɯ/, and *érhuà*, /V(ɴ)ɻ/. Depending on speakers and their elocution speed, [ɯɻ] may feature a noticeably *stronger lateral contraction* than in neutral pronunciation. Frequently, it is realized as a postalveolar intense approximant [ɻ̟], possibly 'colored' by the addition of a certain degree of lateral contraction (representable –if really needed– with '[ɻ̟ʐ]'); including two further possibilities, though less advisable: [ɯɻ], [ɯɻ́]).

The massive use of *érhuà* is perhaps the most stereotypical characteristic of (Peking-accented) mediatic Mandarin. However, so far we have not found two single speakers from Peking who would totally agree on which words should take the 'extra' /-ɻ/ (*ie* one not strictly requested in neutral pronunciation) and which ones should not.

If a clear rule existed, we might treat the behavior 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' essays (*cf* also Lu, 2001:151-189; and He, 2006:159-162).

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 /Çɯ/. Therefore (beside neutral or nearly neutral /V(ɴ)ɻ/ [V̩]), [Vɻ] is also possible – including their nasalized counterparts, of course.

But, let us start from the *consonants*, which structurally –*ie* phonemically– are the same as those of neutral Mandarin Chinese. However, we will soon see that there are many differences in their realization. The reader are warmly invited to make their own comparative transcriptions.

As for the *nasals*, we do not generally have the taxophone [m̂], except in milder accents, verging on the neutral one. In addition, as a typical mediatic feature, we find that /Vn, Vɴ/ become [Ṽn̰, Ṽɴ̰], *ie* the vocoids are nasalized, while the nasals become seminasals (*cf fig* 13.0.1).

Passing to the plain Chinese stops and stop-strictives (*ie* /C/, not the 'aspirated' ones, /Ch/), their mediatic realizations tend to be semi-stop and semi-stopstrictive, *ie* with a laxer and incomplete closure (*cf fig* 13.0.2): [p̓, b̓; t̓, d̓; k̓, g̓] and [t̓s, d̓z, t̓ɕ, d̓ʑ; t̓ʂ, d̓ʐ], in their voiceless (plain & lenis), halfvoiced, and voiced versions, [p̓, b̥; t̓, d̥; k̓, g̊], [t̓s, d̥z; t̓ɕ, d̥ʑ; t̓ʂ, d̥ʐ]. Some speakers can also have *slit* [t̓θ, dð] &c, instead of [t̓s, dz] &c.

In *very fast* and uncontrolled speech, 'unaspirated' /p, t, k/ in unstressed syllables, can often become [β, ð] (approximants), or [ð̞] (semi-approximant), [ɣ] (semi-constrictive), while /ts(j), tʂ/ can become [ˈs, ˌz̥, ˌz; ˈɕ, ˌʑ̊, ˌʑ; ˈʂ, ˌʐ̊, ˌʐ], depending on stress.

Their 'aspirated' counterparts, /Ch/, have normal (voiceless) articulations, [p, t, k; ts, tɕ, tʂ]. But, instead of the neutral /h/ [ɦ, h, (ɦ)] taxophones, they generally have weaker realizations: [ɦ̞, h̞, ∅] (*ie* a laryngeal semi-constrictive and semi-approximant, and a zero phone, *cf fig* 13.0.3). Again, no (neutral) labialized taxophone, [p̂, b̥], usually occurs in the mediatic accent. While, instead of [ˈCɦ̞, ˌCh], before /i, y; j, ɥ/, we find [ˈCʝ, ˌCʝ] (*ie* palatalized laryngeal semiconstrictive and semiapproximant, respectively).

Also the *constrictive* phonemes have weaker articulations. In fact, instead of /f/ [f, v̥], we find [ˈf, ˌv̥, ˌv] (with no neutral-like [f̬, v̊, v̬], *cf fig* 13.0.4), and also /f/ → ↓[hʊ].

Even /#h/ is veaker than in neutral pronunciation: [ˈʜ, ˌh, ˌɦ] (velar semiconstrictive, approximant, and semiapproximant), instead of [ɻ, ʁ, h] (*fig* 13.0.5).

Passing to the grooved constrictives, again, we generally find (voiceles and half-voiced) *semi*constrictive phones (including fully voiced ones, up to their corresponding *slit* approximants, which have a typically broad flavor, *cf fig* 13.0.6): /s/ [s̆, ẓ̆, ź; ʂ̆, ẓ̊, ẓ́], /ʂ/ [ʂ̆, ʐ̆, ʐ́] (and [ɹ, ɹ, ʐ]), instead of neutral [s, z̥; ɕ, ʑ̊; ʂ, ʐ̊] (let us note that the prepalatal approximant [ɹ] is different from a palatal semiapproximant [ʝ]). Some speakers can also have *slit* [θ, ð] &c, instead of [s̆, ź] &c.

The dorsal phonemic Chinese *approximants*, /j, ɥ, w/ [j, ̥j; ɥ, ̥ɥ; w, ̥ω], are also weaker (*ie* semiapproximants) in mediatic pronunciation: [ʝ, ɥ, ω]. In addition, they have some broader (still semiapproximant) variants, as well; for the first two, we have: [ʝ] (*post*palatal), [ʉ] (*prevelar* rounded, *cf fig* 13.0.7).

While, for /#w/, we find some even broader (and more typical) variants: [ʋ] (labiodental approximant), [ʋ̂] (labiodental rounded approximant), [ʋ̂] (labiodental rounded *semi*approximant). Indeed, in mediatic pronunciation, [#ω] is generally found only before the rounded vowels, /ɔ, u/ (while, /Cw/ normally has [Cω], followed by any vowel type): *wànwéiwǎng* 'World Wide Web' [ʋɑ̃ŋ̩ˈʋɩɩ˩ʋɑ̃ɑ̃ŋ̩, -ˈʋɩɩ-]

fig 13.0.1. Pekingese mediatic pronunciation: nasals & seminasals (for comparison).

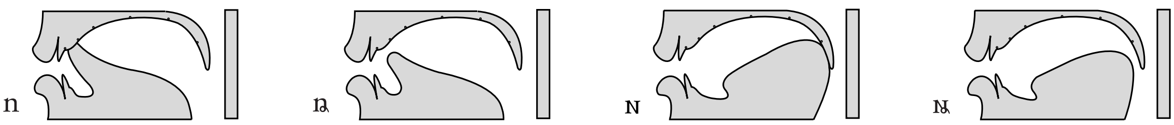


fig 13.0.2.1. Pekingese mediatic pronunciation: semi-stops (by detension) & variants.

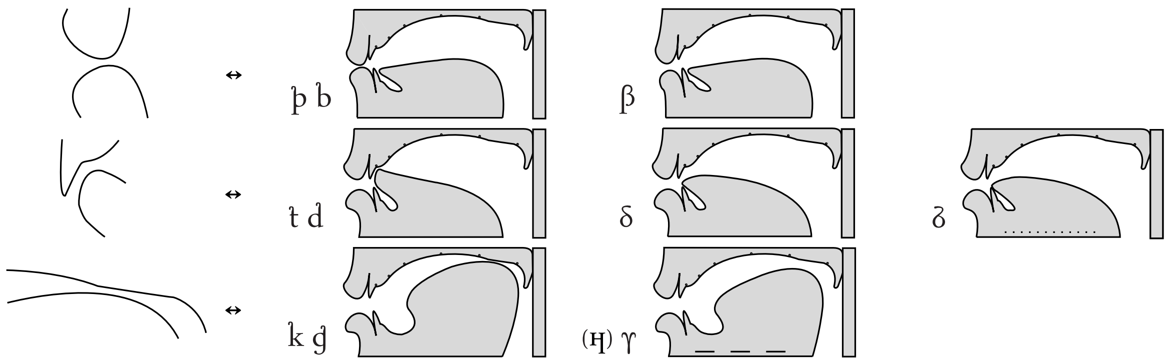
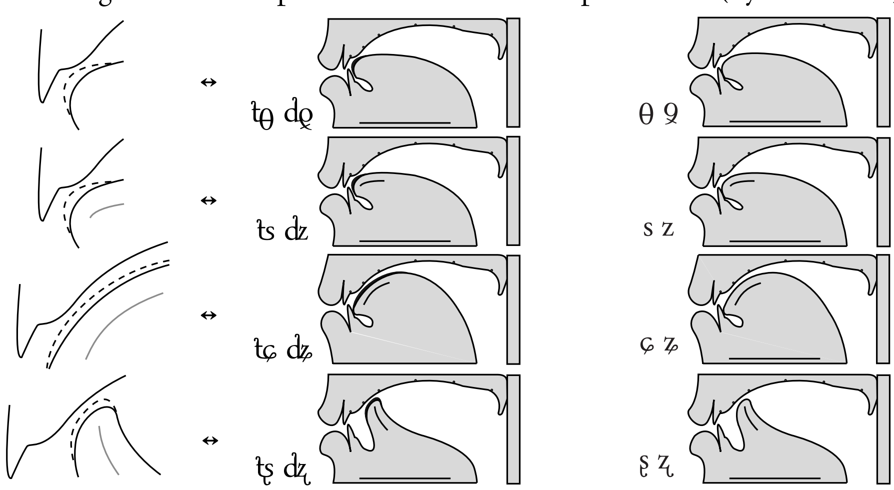


fig 13.0.2.2. Pekingese mediatic pronunciation: semi-stopstrictives (by detension) & variants.



(with further variants: see below), but *tā wèn wǒ* 'he/she asks me' [¨thɐ\ʋə̃ŋ̟ ˌωɤ̝(x̝|), -ˌωɤ̝(x̝|)], (obviously, other stress patterns are possible, too).

The labiodental variants are gradually gaining ground among youngsters, while being far less frequent, if not virtually absent, among most elders. In their early attempts

fig 13.0.3. Pekingese mediatic pronunciation: /Cɧ/. Laryngeal semiconstrictive, [ħ̞], and semiapproximant, [h̞], in comparison with neutral laryngeal constrictive, [ħ], and approximant, [h].

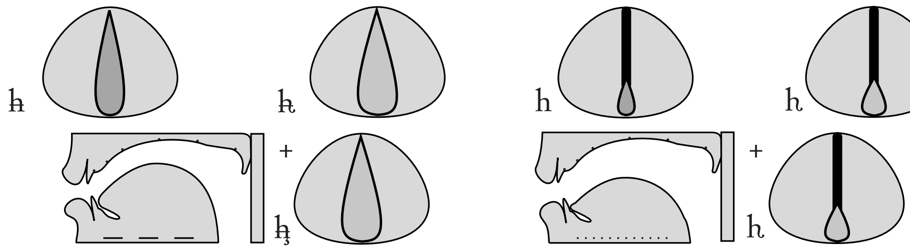


fig 13.0.4. Pekingese mediatic pronunciation: realizations of /f/ & /#ɧ/.

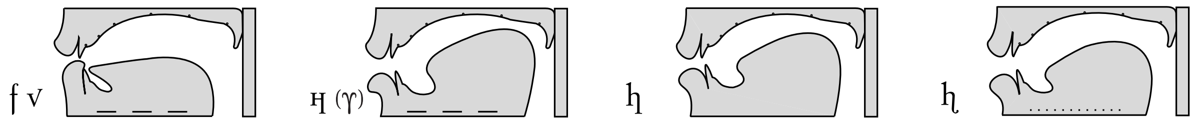


fig 13.0.5. Pekingese mediatic pronunciation: grooved semi-constrictives (& variants).

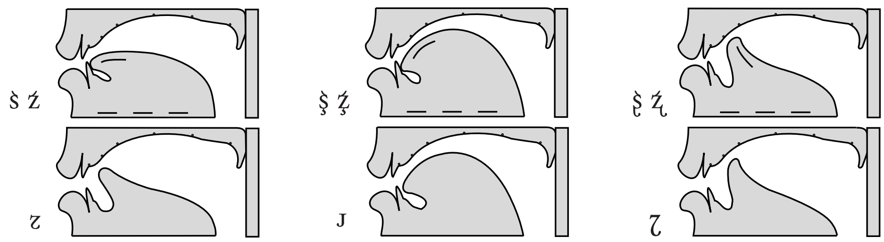


fig 13.0.6. Pekingese mediatic pronunciation: approximants (& semiapproximants).

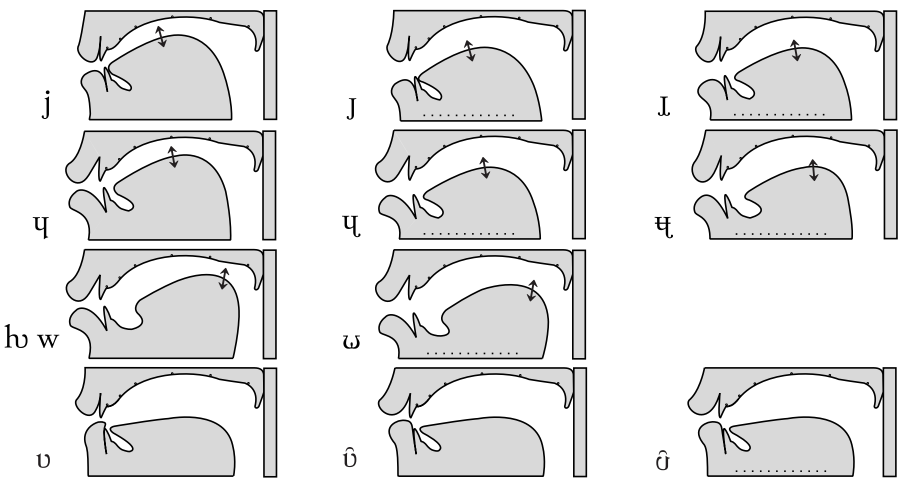


fig 13.0.7. Pekingese mediatic pronunciation: three realizations of the approximant phoneme /ʐ/.

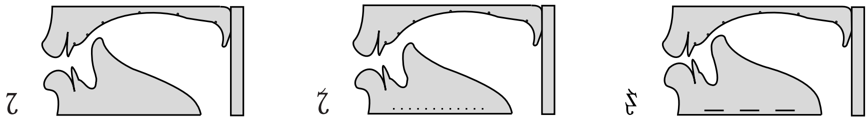
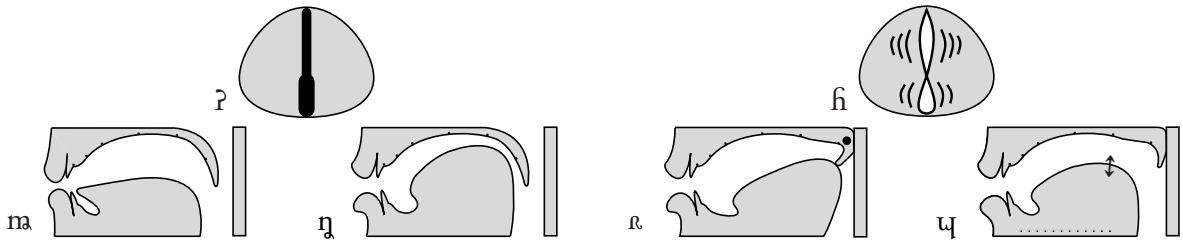


fig 13.0.8. Pekingese mediatic pronunciation: frequent contoid insertions before /#V/: [ʔ] for all vowels, including /i, y, u/ (and also in place of /j, ɥ, w/ before /i, y, u/), but [ʔ, ɦ, ɱ, ɳ, ɻ, ʯ] before non-high vocoids.



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 [ʊ] from [v] (as some Dutch speakers are, for example), /w/ [w] is the best and only choice for teaching/learning purposes.

Our last (slit) approximant phoneme, /ʐ/ [ʐ], also has two partially different realizations, which are either stronger or weaker than its corresponding neutral one (*fig* 13.0.8): in adition to [ʐ], we find [ʑ] (semiapproximant, especially in weak syllables), and [ʓ] (*slit* semiconstrictive, especially in stressed syllables). This last phone might, perhaps, explain why some '*off*IPA phoneticians' consider /ʐ/ as if it actually were '[ʒ]'.

13.5. We will, now, consider the *vowels* of Mandarin Chinese, by means of the two sets of *fig* 13.1 & *fig* 13.2, which have to be inspected very carefully. The interested readers should also prepare comparative transcritions with examples of neutral Chinese.

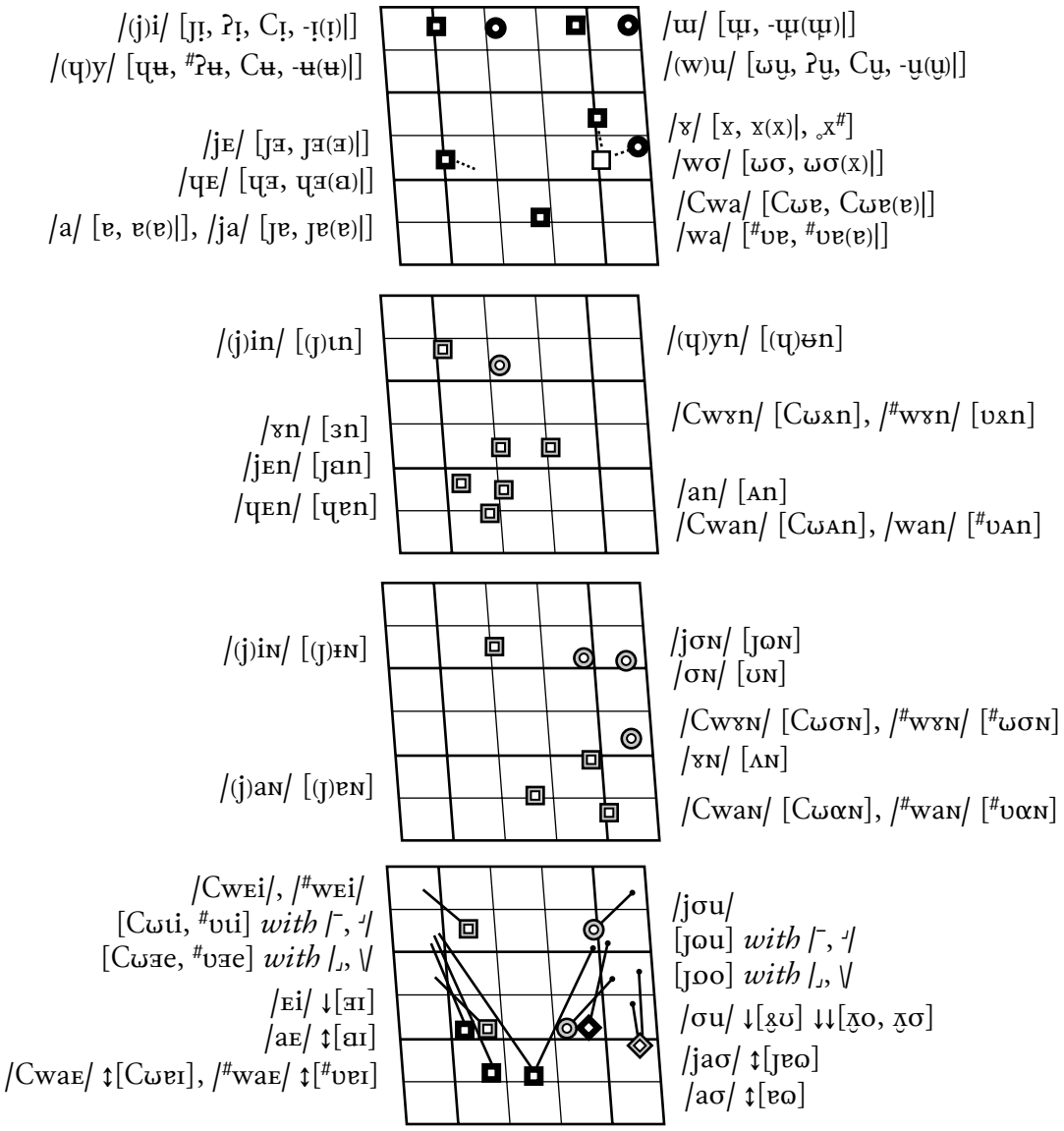
Among other things, let us notice a popular trend among young Peking speakers, which involves articulating /ɐi, ʊu/ as narrower diphthongs: [ɐe, ʊo] (*fig* 13.1.1-2). In *faster speech* and in protunes, even monotimbric realizations are possible, [ɐɐ, ʊʊ] (particularly in /◦$/, weak syllables). These variants occur in addition to [↓ɐɪ] and [↓ʁ̞ʊ, ↓↓ɤ̞ʊ], as well.

Vowel reductions are shown in *fig* 13.1.3, while mediatic *érhuà* peculiarities are given in *fig* 13.2.1-7.

13.6. What can be said with respect to vowel and consonant reduced forms, in ordinary pronunciation, largely applies to mediatic speech too, of course, and even more strongly (as well as the typical paraphonic setting with a *raised larynx*, ⟨∴⟩, and not only for women. Other Chinese languages, instead, use a normal –or modal– laryngeal setting, or an opposite one, with a *lowered larynx* – thus another marked one, ⟨∵⟩).

In addition to that, non-native speakers should be able to recognize two less common types of reduction, affecting vowels and consonants especially in *fast speech* and in a *protune* (ie not in a *tune*), for they can make Chinese less easy to understand. They are not at all indispensable for acquiring a good pronunciation. However,

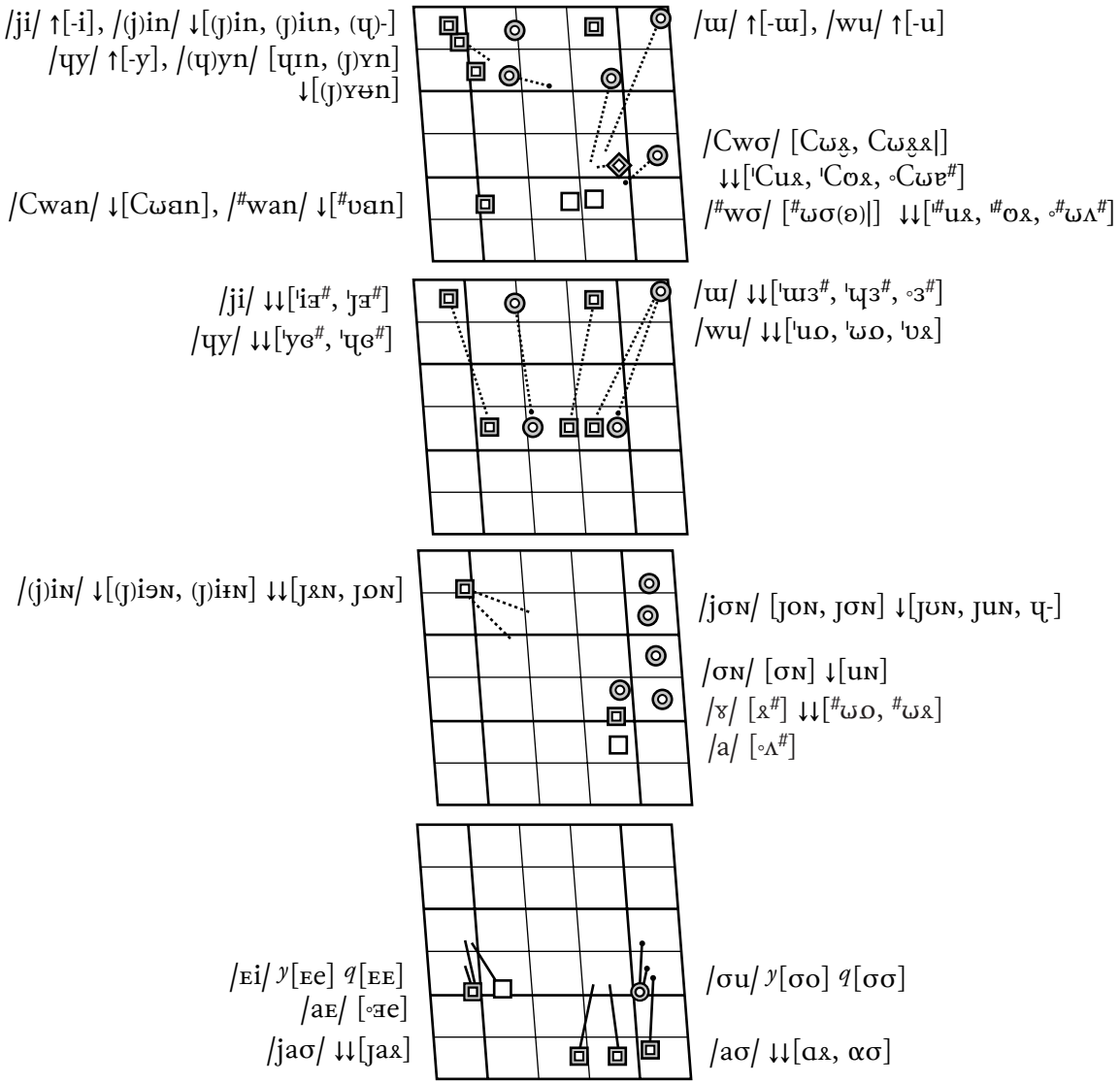
fig 13.1.1. Pekingese mediatic pronunciation: vowels & diphthongs (*cf fig* 6.1-B).



should foreign learners want to apply them in an appropriate manner, naturalness would certainly be improved. And, of course, those foreigners who readily recognize them certainly improve their own understanding of current Chinese.

13.7. In addition to the typical paraphonic *postdorsal vowel* setting (*fig* 12), a *reduced vowel* setting, ⟨V̬⟩ (*fig* 13.3.1), is typical of mediatic Mandarin pronunciation (as it is of the Peking *dialect* itself, even in a stronger form, and combined with ⟨V̭⟩). This setting reduces the space in the mouth available for the production of both stressed and unstressed syllables, which means both vocoids and contoids, of course. Thus, all phones become less defined than usual, producing a kind of muffled speech, certainly more difficult to be understood, even by native speakers, especially those from different areas. All this is added to the various attenuations already seen for neutral Chinese pronunciation (*cf* § 6.5[.1-3]), causing further reductions.

fig 13.1.2. Pekingese mediatic pronunciation: some variants, including /ɛi, σu/ for ʸ (= younger speakers) & ᑫ (= quicker speech, especially in protunes, or weak syllables – *cf fig* 6.1-B).



Thus, marked mediatic accents may use the (normal) neutral vowels and diphthongs, shown in the vocograms given in *fig* 6.1-4, but they are modified by the paraphonic setting of *fig* 13.3.1. So, *fig* 13.3.2 compares the collocations resulting from their sum. However, these are not the most typical mediatic ones (which are shown in *fig* 13.2.1-7), just to make comparisons easier. The paraphonic effect is still perceptible, although less than in broader mediatic accents.

As can be seen, all elements are more or less centralized. Many of them do not need new simbols, to show their difference: just a (proper) diacritic would be sufficient, in the figures. But, a few of them move to other boxes. Mostly /i, y, ɯ, u, ɛ/, especially when lengthened by a full T-3, /ˌ/ [ˌ]: [ii → iɪ], [yy → yʏ], [ɯɯ → ɯɯ], [uu → uʊ], [ɛɑ → ɜɜ], and possibly [ʌʌ → ɐɐ], [aa → ɐɐ] (with a slightly different starting point), [ɑɑ → ʌʌ], too.

We can also find: [ʌɜ → ɐɜ], [aɜ → ɐɜ] (again, with a slightly different starting

point), [aɒ → ɐɒ], and [ɑɒ → ʌɒ]. In addition, [ɛ → ɐ], and [ˌəɪ → ˌəə], [ˌɤʊ → ˌʁʊ], [ˌɚ → ˌɜ], [ˌɻ → ˌɻ], and [ˌɪ → ˌə], [ˌɨ → ˌə], as well.

fig 13.1.3. Pekingese mediatic pronunciation: vowel reduction (in two vocogram, resulting from the combination of *fig* 6.3-4), for comparison.

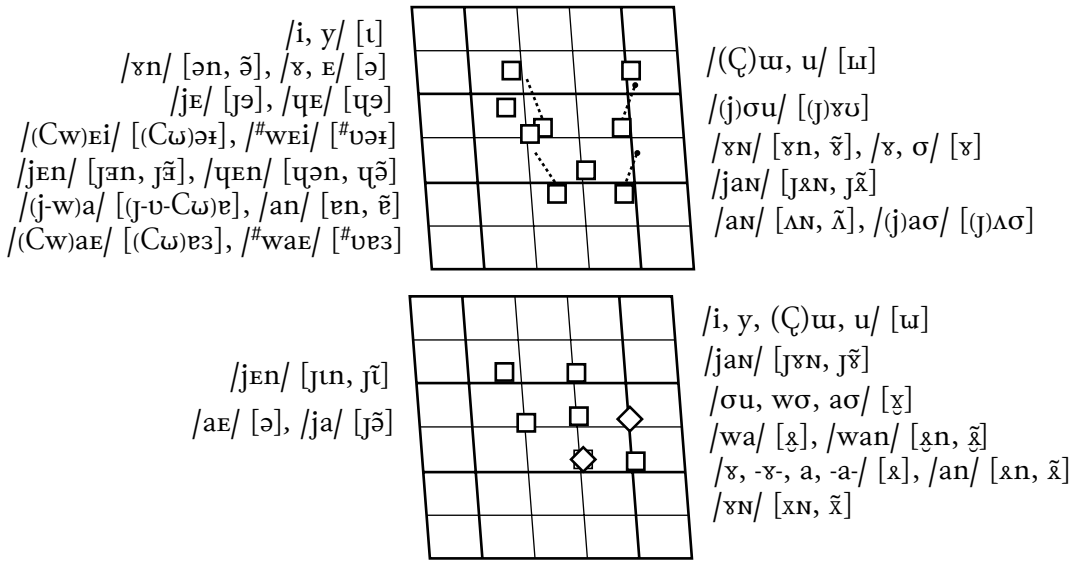


fig 13.2.1. Pekingese mediatic pronunciation: /(j-ɥ-w)Vʔ/.

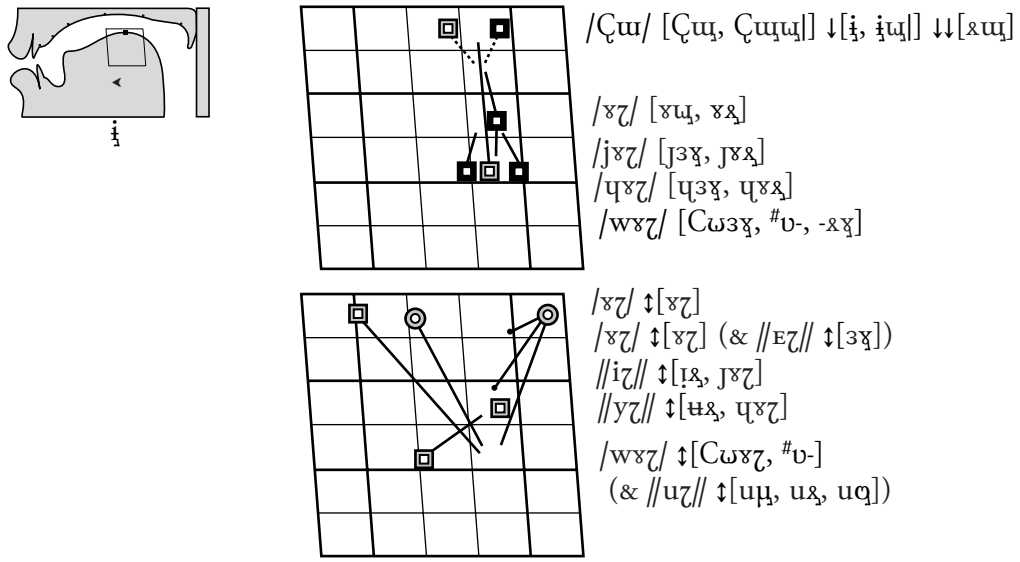


fig 13.2.2. Pekingese mediatic pronunciation: /(j-ɥ-w)V(ɴ)ʔ/.

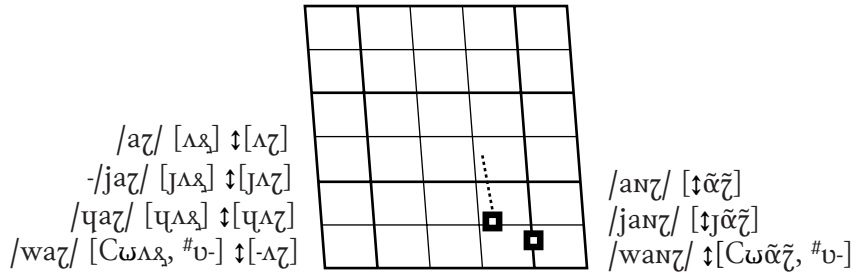
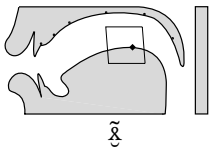
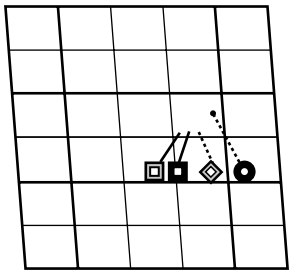


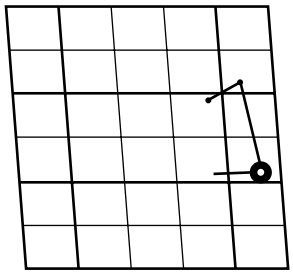
fig 13.2.3. Pekingese mediatic pronunciation: /(j-ɥ-w)Vɴz/.



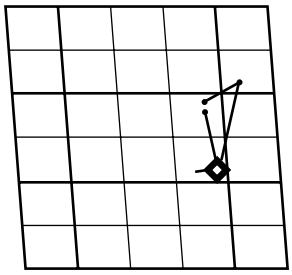


/ɤɴz/ [ɪx̃ˠ̃z̃]
/jɤɴz/ [jx̃ˠ̃, jɜ̃ˠ̃] & [ɪjx̃ˠ̃z̃]
(& /iɴz/ [ɪĩˠ̃])
/wɤɴz/ [Cωx̃ˠ̃, ʰʊ-] & [ɪ-x̃ˠ̃z̃]
/ɥɔɴz/ [ɥõˠ̃]
(& /jɤɴz/ [jx̃ˠ̃, ɪjx̃ˠ̃z̃])
/wɔɴz/ [ωõˠ̃]
(& /wɤɴz/ [Cωx̃ˠ̃, ʰʊ-, ɪ-x̃ˠ̃z̃])

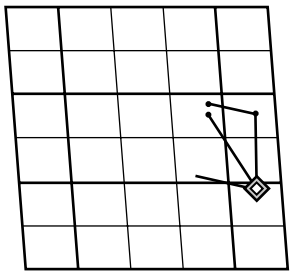
fig 13.2.4. Pekingese mediatic pronunciation: /ɔuz/.



/ɔuz/ [ɔɥ] (& [ɔɤ̯, ɔʊɥ])



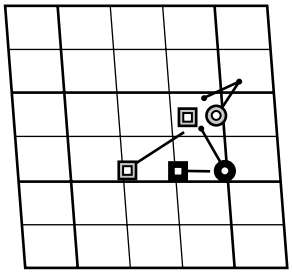
[ɤ̯ɥ] (& [ɤ̯ɤ̯, ɤ̯ʊɥ])



[ɪx̃ɥ] (& ɪ[x̃ɤ̯, x̃oɥ])

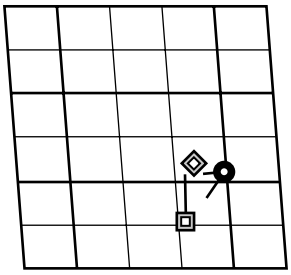
fig 13.2.5. Pekingese mediatic pronunciation: /j-ɥV(V)z/.

/jɛz/ [jɜɤ̯] ɪ[jɤz, jɜɤ̯]
(& ɪ/iz/ [ɪi])
/ɥɛz/ [ɥɜɤ̯] ɪ[ɥɤz, ɥɜɤ̯]
(& ɪ/yz/ [ɪʉɤ̯])

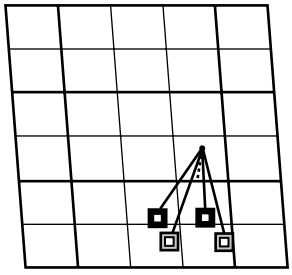


/jɔuz/ [jɒɥ, ɥ-] ɪ[jɒʊɥ]

fig 13.2.6. Pekingese mediatic pronunciation: /(j-w)V(V)z/.

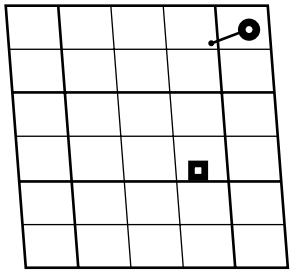


/wɔz/ [ωɒɤ̯,
ωɒʌ] & ɪ[Cωɤ̯z, ʰʊ-]
& /ɪwaz/ [Cωʌɤ̯, ʰʊ-]
/m-p-f/+/wɔz/
[Cωɒɤ̯] & [ɪCωɤ̯z]

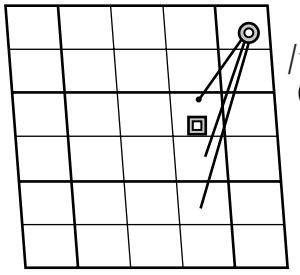


/jaɔz/ [jɐɒɤ̯, jaɒɤ̯]
/aɔz/ [ʌɒɤ̯, ɑɒɤ̯]
(& /az/ [ɐɤ̯])

fig 13.2.7. Pekingese mediatic pronunciation: /uz, wɤz/.



/uz/ [uɻ]
/wɤz/ [Cωɤ̯, ʰʊɤ̯]



/uz/ [uɤ̯, uʏ, uɒ]
(& /wɤz/ ɪ[Cωɤz, ʰʊ-])

fig 13.3.1. Pekingese mediatic pronunciation: typical reduced-vowel paraphonic setting (implying a reduction in the vocoid space within the mouth, although here shown in a generic way, with only extreme vowels, even if not peculiarly Chinese ones).

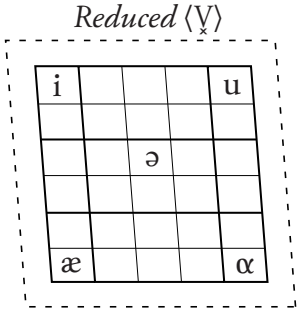
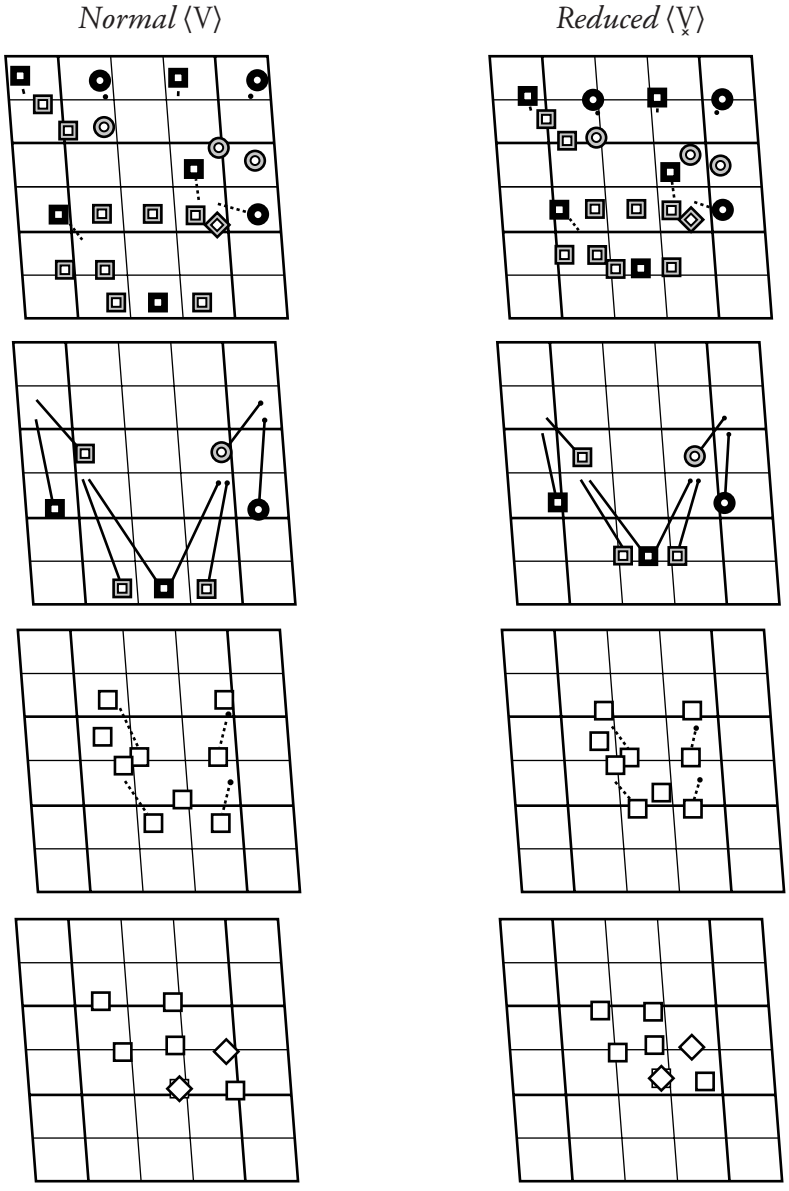


fig 13.3.2. Pekingese mediatic pronunciation: normal & reduced vowel settings.

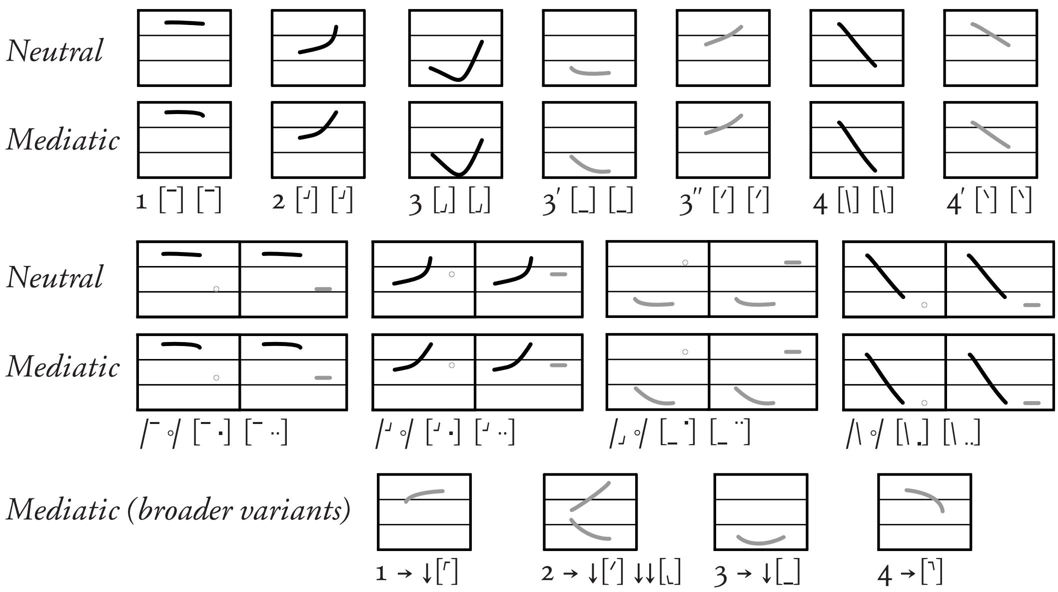


13.8. As far as *stress* and *intonation* are concerned, there are no significant differences with respect to neutral Chinese. However, a relatively large number of common words ending with a marked toneme, /ˈ$#/, feature a 'zero toneme' in Peking-accented Chinese: *pángxie* [˧pʰẽŋ̩·ʑɟɜ] 'crab', *wéiqün* [˧ʋɩi·tɕɣ̃ʀ] 'apron', *nángua* [˧nã̃ŋ̩·ɡ̊ɯʌ] 'pumpkin'.

These can coexist with their 'normal' realizations, as presented in most dictionaries: *pángxiè* [˙pʰaɴ\ɕjɛ̩], *wéiqǘn* [˙wɘɪ˧tɕɦʏn], *nánguā* [˙nʌŋ˥k̩wa]. Similarly, post-verbal monosyllabic pronouns *wǒ*, *nǐ*, *tā* are more systematically reduced to /˳$/ than in neutral pronunciation (unless emphasis is needed – se, however, §6.5.5): *wǒ wèn nǐ* [˳ɯɣ̩\ʋɜ̃ŋ̩n̩ɪ̩] 'I ask you', *gēn tā zǒu* [˙ɡ̊ɜ̃ʀ·tɐ˩t̩ʂɤ̥ɤ̥ʋ̩, -ɣ̩ɣ̩] 'to go away with him/her'.

The four marked *tonemes*, and the unmarked one, have only tiny differences in comparison with the neutral ones, as can be seen in fig 13.4.

fig 13.4. Pekingese mediatic pronunciation: neutral ton(em)es compared with milder and broader mediatic ones.



21.
Phonetic 'transcriptions' in different handbooks

21.1. We are perfectly aware that such highly detailed phonotonetic transcriptions, as those presented in this book, 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 systematic 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 highly unfriendly '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 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 there is a clear need for specifically prepared pronunciation handbooks, for those who want a more complete perspective, and aim at acquiring a native –or nearly-native– pronunciation.

This book, in fact, can be certainly used as a pronunciation handbook, as it provides a detailed description of all phonotonetic 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 enthu-

siast (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.

21.1.1. 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', which is natural for native speakers, but not necessarily so for foreign learners. Namely, our choice fell on [C̬h, C̥], for stops and stop-strictives, and on [C̬], for constrictives. Here follows a complete list:

b	[b̥]	*p*	[ph]	*f*	[f]	*m*	[m]		
d	[d̥]	*t*	[th]			*n*	[n]	*l*	[l]
z	[d̥z]	*c*	[tsh]	*s*	[s]				
j	[d̥ʑ]	*q*	[tɕh]	*x*	[ç]			*y*	[j] & [ɥ]
zh	[d̥ʐ]	*ch*	[ʈʂh]	*sh*	[ʂ]			*r*	[ʐ]
g	[g̊]	*k*	[kh]	*h*	[ʁ]	*ng*	[ɴ]	*w*	[w]

21.1.2. For [(J)V, (J)VV, (J)VN], a similar criterion is followed, but without renouncing realism: thus, a diphthong like [ɑʊ] will not be simplistically reduced to '[ɑu]', even though [ʊ] is not used individually elsewhere. In other words, simplicity must be combined with realism (*fig* 21.1).

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

I

i	[i, #ji]	*ï*	[ɯ, ʐ̍]	*ui*	[wᴇɪ]		
in	[(#j)ɪn]	*ing*	[(#j)ɩɴ]				

E

e	[ɤ]	*ie*	[jᴇ]	*üe*	[ɥᴇ]	*ei*	[ᴇɪ]
en	[ɜn]	*er*	[ɤʐ]	*eng*	[ɤɴ]	*ueng*	[wɤɴ]

A

a	[a]						
ai	[ᴀᴇ]	*uai*	[waᴇ]	*ao*	[ɑɑ]	*iao*	[jaɑ]
(w)an	[(w)ᴀn]	*ian*	[jɛn]	*üan*	[ɥɛn]		
(i)ang	[(j)aɴ]	*uang*	[wɑɴ]				

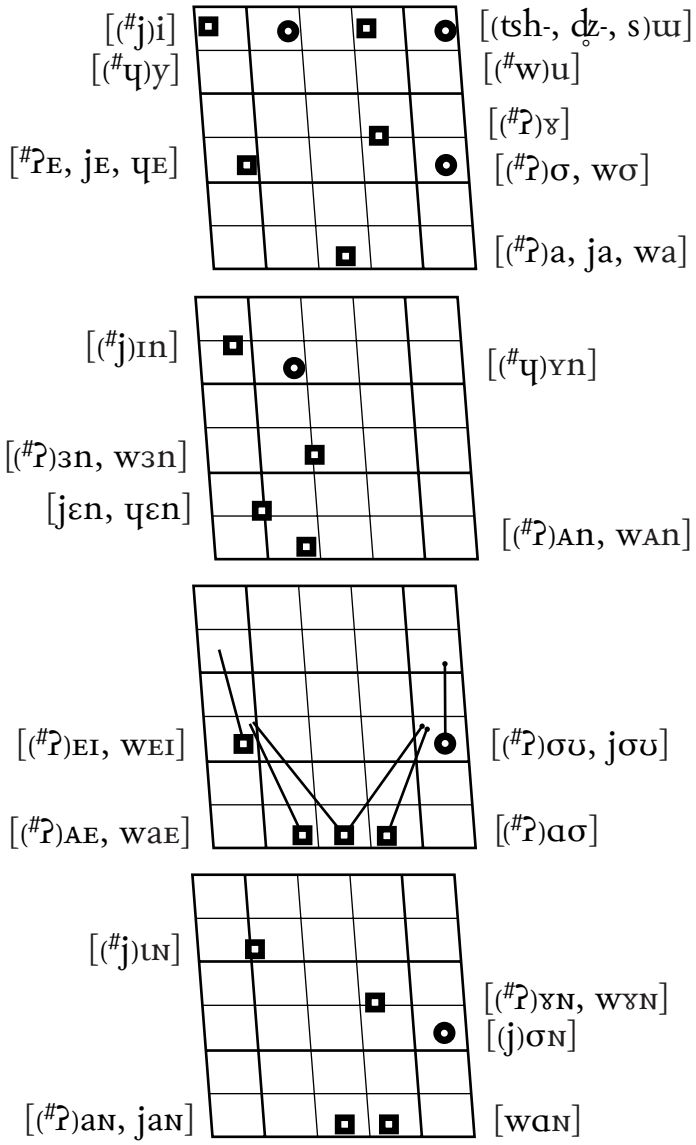
O

o	[ɑ]	*ou*	[ɑʊ]	*uo*	[wɑ]
ong	[ɑɴ]	*iong*	[jɑɴ]		

U

u	[(#w)u]	*ü*	[(#ɥ)y]	*iu*	[jɑʊ]
un	[wɜn]	*ün*	[(#ɥ)ʏn].		

fig 21.1. Simplified vowel inventory, for teaching purposes.



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 [jæ̱n] 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 unstressed [ə]: no panic, let them use [ə], at first.

Listening to audio files *and* practising kinaesthesia, with a constant ear feed-

back, will enable those really interested in acquiring [ɜ] –and, at a more advanced stage, also [ɜ]– 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 neither [ᴀ], nor [ɑ]. Articulating [a] everywhere is fine, and phonemically correct; yet it is preferable, for the transcription itself, to keep using three symbols, anyway.

In fact, there will be some students who, being more sensitive than others to the quality of vowel sounds, want to know *where* [ᴀ] 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 our 'normalized' neutral 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 booklet *Pronuncia cinese per italiani* –Chinese Pronunciation for Italians– which was intended, primarily, for those who simply want to be understood, irrespective of how foreign their 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.

21.1.3. As far as *érhuà* is concerned, we have devised a reasonable scheme, derived from that proposed in § 7.7, with just minor adjustments. Thus, we have:

A. [aʐ]: *huà-huàr* [˙ʁwa\ʁwaʐ] *n*[˙ʁwa\ʁwɐ̢ʁ̝];
[jaʐ]: *chuāngliánr* [¨ʈʂhwɑɴ˩ljaʐ] *n*[¨ʈʂhwɑɴ˩ljɐʁ̝];
[ɥaʐ]: *yüánqüānr* [˙ɥɛŋ¯ʨhɥaʐ] *n*[˙ɥaŋ¯ʨhɥɐʁ̝];
[waʐ]: *hǎowánr* [‥ʁɑʊ˩waʐ] *n*[‥ʁɑ̢ɔ̢˩wɐʁ̝].

B. [ɤʐ]: *bèikér* [\b̥ᴇɪ˩khɤʐ] *n*[\b̥ᴇ̢ɪ̢˩khɤʁ̝];
[jɤʐ]: *shǐjìnr* [‥ʂʐ̩\ʥ̊jɤʐ] *n*[‥ʂɯ̢\ʨ̊jɤ̢ʁ̝];
[ɥɤʐ]: *qūqür* [¯ʨhy·ʨhɥɤʐ] *n*[¯ʨhy·ʨɥɤʁ̝];
[wɤʐ]: *kǒuwèir* [‥khɔʊ\wɤʐ] *n*[‥khɔ̢ʊ̢\wɤ̢ʁ̝];
[jɜʐ]: *shíjiēr* [˙ʂʐ̩¯ʥ̊jɜʐ] *n*[˙ʂɯ̢¯ʨ̊jɜʁ̝];
[ɥɜʐ]: *mùjüér* [\mu˩ʥ̊ɥɜʐ] *n*[\mu̢˩ʨ̊ɥɜʁ̝].

C. [ɔʐ]: *tǔdòur* [‥thu\d̥ɔʐ] *n*[‥thu̢\ʈɔ̢ʁ̝];
[jɔʐ]: *zhīliúr* [¨ɖ̥ʐ̩˩ljɔʐ] *n* [¨ɖ̥ʐɯ̢˩ljɵʁ̝];
[ɑɔʐ]: *nǎimáor* [‥nᴀᴇ˩mɑɔʐ] *n*[‥nᴀ̢ɜ̢˩mɑʁ̝];
[jaɔʐ]: *màimiáor* [\mᴀᴇ˩mjaɔʐ] *n*[\mᴀ̢ɜ̢˩mjaʁ̝];
[wɔʐ]: *shānpōr* [¨ʂᴀm¯phwɔʐ] *n*[¨ʂᴀm¯p̂hɔʁ̝].

D. [ãʐ̃]: *piānchǎngr* [¨phjɛɳ‚ʈʂhããʐ̃] *n*[¨phjɛɳ‚ʈʂhã̢ã̢ʁ̝̃];
[jãʐ̃]: *dòujiāngr* [\d̥ɔʊ¯ʥ̊jãʐ̃] *n*[\d̥ɔ̢ʊ̢¯ʨ̊jãʁ̝̃];
[wãʐ̃]: *tiānchuāngr* [¨thjɛɳ¯ʈʂhwãʐ̃] *n*[¨thjɛɳ¯ʈʂhwãʁ̝̃].

E. [ɤ̃ʐ̃]: *bǎndèngr* [‥b̥ᴀn\d̥ɤ̃ʐ̃] *n*[‥b̥ᴀ̢ɳ\ʈɤ̢̃ʁ̝̃];
[jɤ̃ʐ̃]: *qǐmíngr* [‥ʨhi˩mjɤ̃ʐ̃] *n*[‥ʨhi̢˩mjɤ̃ʁ̝̃];
[wɤ̃ʐ̃]: *xiǎowèngr* [‥ɕjaɔ\wɤ̃ʐ̃] *n*[‥ɕja̢ɔ̢\wɤ̢̃ʁ̝̃].

F. [õʐ̃]: *tōukòngr* [¨thɔʊ\khõʐ̃] *n*[¨thɔʊ\khõ̢ʁ̝̃];
[jõʐ̃]: *kūqióngr* [¨khu˩ʨhjõʐ̃] *n*[¨khu˩ʨhɥõʁ̝̃].

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

21.2. 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 Chinese works. Instead, tones and intonation shall always be presented in full notation, since any simplification could easily jeopardise mutual intelligibility. Here is an example of connected speech, transcribed according to the simplified pronunciation described above (*cf* § 12.2):

⟨∴ ʏ̬⟩ [˩ljɔʊ̣.jɔʊ˙ɥy\d̥i˙.|||
\ja̤ma\ɕʏɴ˩ʁɤ ˙ʐɤ\d̥ᴀᴇ ..ɥy˙lɪŋ¯d̥ʐɔɴ·‖ _jɔʊ\ji_d̥ʐɔɴ· ˪\d̥ʑjaɔ·| ..ji˩ɕjɔɴ·d̥ɤ ˙d̥ɔɴ\wu˙˩|||
\d̥ʐɤ.d̥ʐɔɴ ..ji˩ɕjɔɴ·| \ʂʐ̩·¦ ˙ʂʐ̩\d̥ʑjᴇ..ʂan,| ˙d̥zwᴇɪ\d̥a.d̥ɤ· ˩ʂʐ̩.ji\ʂɔʊ˙.||| ..mᴇɪ¯thjɛn ˙jaɔ¯ʈʂhʐ̩.‖ ¯ji-·d̥jaʐ· ˙ljɔʊ\wᴀŋ.d̥ʐʐ̩ ˙ma˩ji˙.|||
˙ʐᴀn˩ʔɤʐ̩.| \ʐan˩ʐɜn· ‖d̥a˩wᴇɪ ¨d̥ʑɪɴ\thaɴ.d̥ɤ \ʂʐ̩.‖ ..ji˩ɕjɔɴ.| _jɔʊ \ji..d̥ʐɔɴ˙ɕi\ɕɪɴ.||| ¯tha-¨ʈʂhʐ̩ ˙ma_ji ˩ʂʐ̩.‖ ˙d̥ʑɥᴇ\b̥u·| ..d̥ʐaŋ̥\d̥ʑɪŋ ¨ʂa˩d̥ʑɥᴇ˙.||| ¨d̥ᴀɴ¯tha ˙phɑɔ¯khᴀᴇ ..mᴇɪ˩ji··g̊ɤ·| _jɔʊ ˩ʈʂhɤɴ¨tɕhjɛŋ˙ʂᴀɴ\wᴀŋ.d̥ʐʐ̩ ˙ma_ji¨d̥ɤ· _ji˙ɕɥᴇ˩ʂʐ̩.‖ ¯tha\d̥ʑjɔʊ..d̥ʐʐ̩¯ʈʂhʐ̩ ˙tɕhi¯d̥ʐɔɴ· \ji..ɕjaɔ-\b̥u..fɜn˙.||| ˥\d̥zwᴇɪ¯d̥wɔ˙.˥¦ ˩wᴇɪ·¦ ˊwu..b̥ᴀᴇ¯d̥ʐɯ̨.||| ¨tha\d̥wᴇɪ· ˙ʂɤɴ\ɕja..d̥ɤ· ˙ma_ji.¦ \b̥u·wɜm-˙b̥u\wɜn˙.||| \d̥ʑɪɴ˩d̥ʐʐ̩·¦ \tɕhy ˙ɕʏŋ_d̥ʐɑɔ· \ɕja.ji··g̊ɤ \mu¯b̥jaɔ.||| ˪..khɤ\d̥ʑjɛn˙.˩ ..ji˩ɕjɔɴ·¯tɕhɪɴ·ʈʂhu··d̥ɤ ¯d̥ʐʐ̩··d̥ɑɔ˙.||| \jaɔ _ʂʐ̩ \d̥zɯ_d̥ʑi̧d̥ɤ ..d̥ʐɔɴ˩tɕhʏŋ˩ ¯ʂɤɴ¨tshwɜŋ̥· \ɕja.tɕhy˙.‖ \d̥ʑjɔʊ\b̥i¯ɕy· _jɔʊ ˊma.ji¨d̥ʑja˩d̥zu··d̥ɤ·| ¯ʂɤɴ¨ʂɤɴ\b̥u¯ɕi˙.|||
..ji˩ɕjɔɴ˙ \ʂᴀɴ_tɕhjᴇ ¯d̥ʐʐ̩··d̥ɑɔ· ˩ljɔʊ.jɔʊ˙ɥy\d̥i˙.||| ˙d̥ʑi\ʁwa̧· ..ʂʐ̩\jɔɴ ¨d̥zɯ˩ɥɛn˙.‖ ˙ʐɜl\lᴇɪ˙.| ¿\b̥u¨g̊ᴀᴇ ¨sᴀn¯sɯ˙.| ¿¨ʔɤʐ\ʁɔʊ ˩ɕɪɴ·ma˙.||||].

21.3. As we have just demonstrated, introducing phonetics in a common Chinese course is neither 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.

Chinese teachers *must* know (at least) the basics of their students' mother tongues, including their phonemic and phonetic systems, 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 Marco (Cerini) 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 'iden-

tical' to our Chinese teacher. He could not help insisting on the 'aspiration' factor, for *p*, *t*, *k* /pɧ, tɧ, kɧ/, totally ignoring the fact that nothing like [Ch] belongs to neutral Italian, unless one is using some special 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, when no reliable and scientifically coherent notational system for tones or intonation is available.

21.3.1. 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 Chao'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 twenties), including quite an extensive tonetic notation, with thick symbols for tones in [ˈ$], and thin ones for tones in [ˌ$] (stressed and half-stressed).

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:

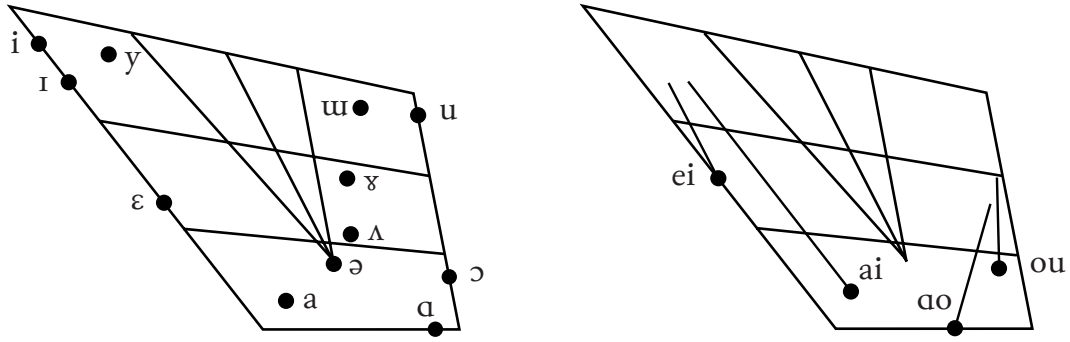
can[b̥, ph]	*off*[b̥, ph]	*can*[d̥z, tsh]	*off*[d̥z, tsh]
can[d̥, th]	*off*[d̥, th]	*can*[d̥ʑ, tɕh]	*off*[j̊, ch]
can[g̊, kh]	*off*[g̊, kh]	*can*[d̥ʐ, tʂh]	*off*[d̥ʒ, tʃh]
can[ʁ]	*off*[x]	*can*[ʐ̊]	*off*[ɹ]*
can[ʂ]	*off*[ʃ]		
can[ɕ]	*off*[ɕ]	*described as a constrictive = *can*[ʐ̧]!	

[V] are treated in a less similar way, yet the two systems remain largely coherent and –so to speak– *mutually intelligible* (ƒig 21.2). Now, let us observe the following list:

can[jᴇ, ɥᴇ]	*off*[iɛ, yə]	*can*[(J)a]	*off*[(V)ɑ]
can[ɤ]	*off*[ɤ]; after [g̊, kh, x]: [ʌ]		
can[wσ]	*off*[uo]; after [ph, b̥, m, f]: [ɔ]		
can[ɯ]	*off*[ɯ]	*can*[ʐ̩]	*off*[ɹ]
can[(w)ᴇɪ, (w)ᴀᴇ]	*off*[(u)eɪ, (u)aɪ]		
can[ɑσ, jaσ, (j)σʊ]	*off*[(i)ɑo; ou, iu]		
can[ɪn, ʏn, ɪɴ]	*off*[ɪn, yn; ɪŋ]		

can[jɛn, ɥɛn]	*off*[iɛn, yan]
can[(w)ʌn, (j)aɴ, wɑɴ]	*off*[(u)an, (i)ɑŋ, uaŋ]
can[(w)ɜn]	*off*[ən, un; ʌŋ]
can[(j)ʊɴ]	*off*[(i)uŋ].

fig 21.2. Vowel diagrams as presented in *Chinese*, by the Linguaphone Institute (2121 [?], first edition, p. 16). In this reproduction, we eliminated the big black dots that where important to orient readers when Daniel Jones' *Cardinal Vowels* were not so known yet (but may be confusing today).



21.3.2. Clearly, this analysis was produced when no highly technological equipment was 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 devices as the *artificial palate*, modelled out of thin resin, or even papier-mâché; the palate had to be covered with chalk powder and inserted into the mouth to immediately remove it, 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 photographs of the phonatory act (with a much higher exposure to radiation than in modern X-ray procedures).

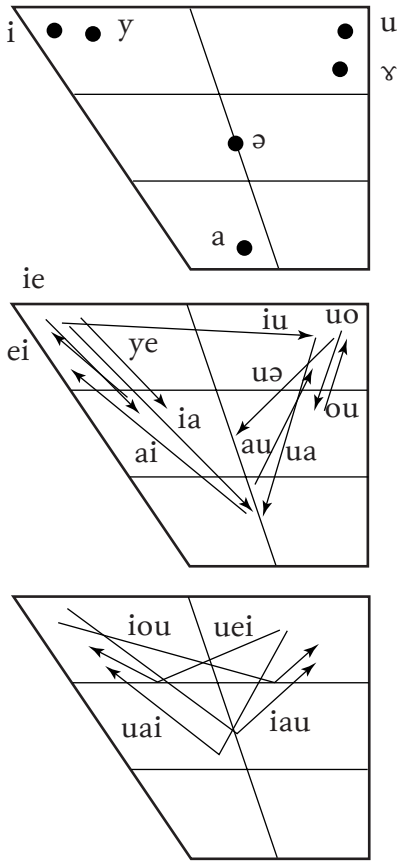
Nonetheless, those pioneers would do their best, not with the purpose of a mere display of scholarly expertise, but to provide their readers with something that could be really useful to learn languages better.

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), until the second series (in the fifties and sixties, on 45 *rpm* records), which had a whole record out of 16 completely dedicated to the pronunciation of the language to be taught, with a sufficient choice of words in *IPA* transcription

for all its phonemes, including cases of main taxophones. 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.

21.3.3. Given such premises, it is quite hard to understand why the *Journal of the International Phonetic Association*, as recently as in 2003, sanctioned an *official* description of Mandarin Chinese, by Lee & Zee, which is embarrassingly inferior to what was available several decades before. Just have a look at the 'official' vowel diagrams (*fig* 21.3.2, that we have graphically improved, not to offend somebody's eyes).

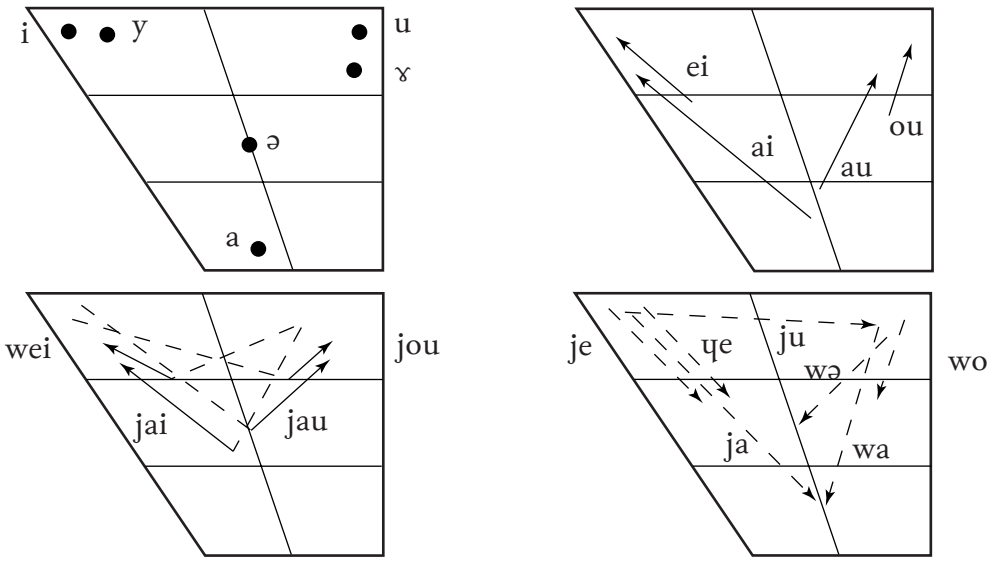
fig 21.3.1. Vowel diagrams as presented in *JIPA* (2003, p. 110).



Apart from the questionable collocation of [ɤ] (^{off}[ɤ]) in the first diagram, the other two trapezia mix up true diphthongs, like '[ei, ai, au]' (^{can}[ɛɪ, aɜ, ɑɔ]), and what should be treated not as diphthongs, or even 'triphthongs', but as [JV(V)] sequences, *eg* '[ye, uei, iau]' (^{can}[ɥɛ, wəɪ, jaɔ]). Regarding '[iau, uai]', we cannot help criticizing the lower part of their 'paths', which should indicate '[a]', but –clearly– lies rather in the area of ^{off}[ə] for '[iau]', while for '[uai]', beside a more than probable ^{off}[ə], again, even with the best will in the world, it might be rendered as a kind of ^{off}[ɜ], but nothing lower (and more appropriate) than that; not even ^{off}[ɐ], which would appear less absurd).

There, 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,

ʃiɡ 21.3.2. Improvement of the official diagrams (and *offIPA* symbols) given in ʃiɡ 21.3 (but with no corrections as far as the movements shown).



ʈʂ, ʂ, ʐ/ appear as '[tʃ̲ʰ, tʃ̲, ʃ̲, ɻ̲]', maybe to avoid that someone might interpret '[ʈʂʰ, ʈʂ, ʂ, ɻ]' as the corresponding *apicopalatal* contoids, used –for instances– in many Dravidian languages (*can*[ʈʂʰ, ʈʂ, ʂ, ʐ]).

We shall not repeat here what we have said about this official treatment of /ɯ/ [ɯ], rendered as [ɨ]! 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.

21.3.4. There is, nonetheless, something that must be said, loud and clear. What is most annoying in many official analyses is not the use of *offIPA in se*, but rather its *misuse*. Undeniably, *offIPA* is structurally simpler than *can*IPA: from one point of view, this is a practical advantage, but also a serious 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 '[iau, uai]' in the *JIPA* diagrams).

*can*IPA 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 loose its readability and –why not?– even its 'typographical appeal'.

Since the introduction of *IPA* to date, AD 2017, many *generic* analyses 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, ^{can}*IPA* 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 (especially for weekend 'phoneticians', even if in charge of university courses).

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.

Let us, now, have a look at *fig* 21.4, which is a real improvement of *fig* 21.3, although we kept the same terrible 'movements', but split the second trapezium in two (becoming the second and fourth). Of course, they are much more readable, in spite of their original defect, but at least we used dotted lines for the consonantal part of the 'diphthongs and triphthongs', representing correct [j, ɥ, w], as shown in the last two trapezia.

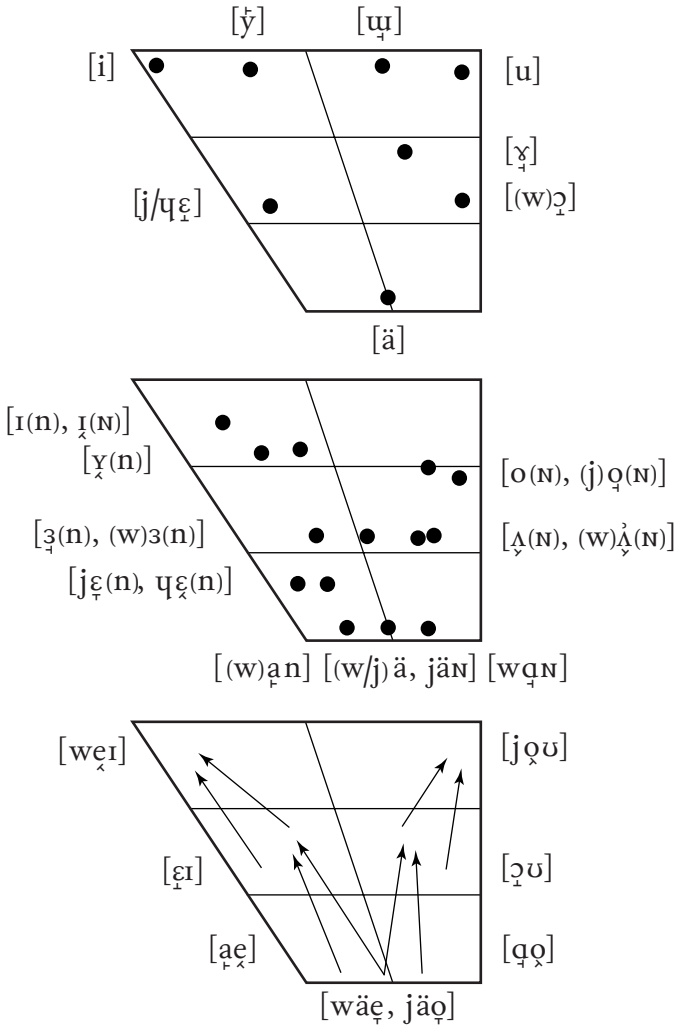
21.3.5. More important than that, we are interested in proving that ^{off}*IPA* can be used with appreciable results, provided everything is done to compensate for its transcriptional limitations. First of all, no transcription should ever become an ugly conglomeration of diacritics. Sure, there are some ^{can}*IPA* symbols with more than one loose diacritic, usually representing peculiar articulations &/or co-articulations; consider, for instance, [x̰̃]: *nasalized* and *laryngealized* (and with the tongue shape modified by *lateral contraction*, which is not an added separate mode of articulation). But, in most cases, unitary symbols are available, and no diacritic is necessary at all, except –logically– when different mechanisms are activated, rightly as the lowering of the velum, or a particular state of the glottis.

When using ^{off}*IPA*, instead, diacritics are inevitable almost everywhere: the point is deciding whether all of them should be printed –or hand-written– at all times, or should they be presented only once, in conjunction with the relevant illustrations. In *fig* 21.4, we give a practical example: only official symbols are used, apart from ^{can}*IPA* expansion and adaptation of [˔ ˂ ˖ ˂ ˕ ˃ ˗ ˃], whose values are intuitively self-evident (rather than squandering [˗] and [˖], for mythic and fictitious 'advanced/retracted tongue root').

Those 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, those diacritics could gradually be 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*: [V̹, n̹, ɴ̹, ʐ̉, ʐ̉]. Also *nasalization* is quite common, *eg* [Ṽ,

fig 21.4. Official vowel diagrams of Chinese 'normalized' pronunciation (with *offIPA* symbols).



z̃]. Vowels may be *partially devoiced*, [V̥], but also *totally devoiced*, [V̊]. We might have to indicate more than one peculiarity at the same time, as in [z̃̊], and to also add *partial labialization*, [z̃̊ʷ]. And if we were to notate all the subtleties involving Chinese [C], some symbols could even need more diacritics!

Now, it is clear that doing the same in *offIPA* would not be viable, especially in transcriptions of continuous speech. 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 travel faster than sound just riding a bike: that is why supersonic jets have been devised! However, a bike still can be a good bike, if we do not hope to use it as a jet. The bottom line is: let us keep *offIPA* true to itself.

21.3.6. In this section, 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: teaching/learning the pronunciation of foreign languages. As such, here, we will not attempt to transcribe neutral Chinese, with all its taxophones and other peculiarities, but the

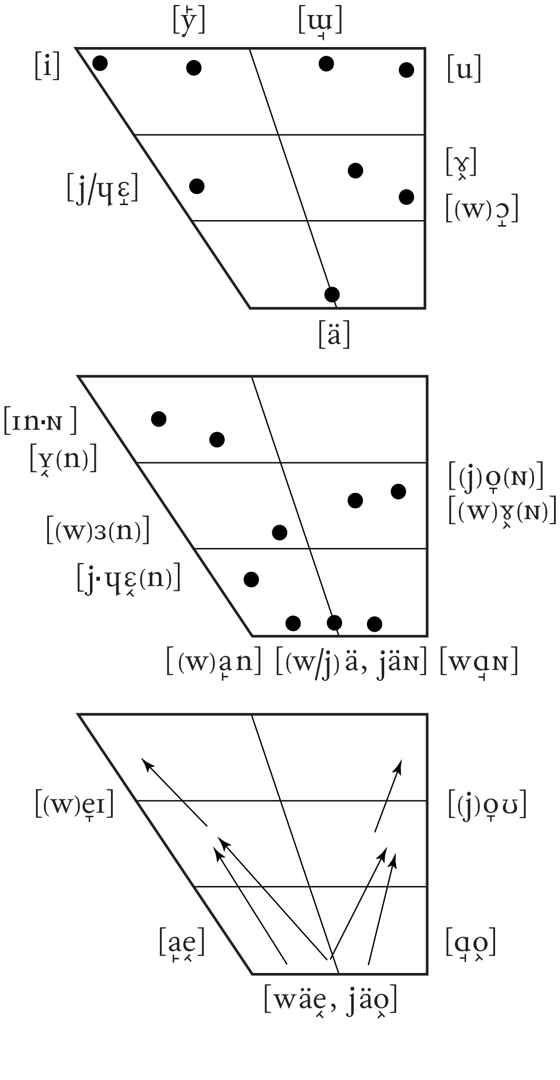
slightly simplified pronunciation described in § 21.1.1-2. Where necessary, the exact value assigned to an official symbol will be specified between ⟦ ⟧, by means of diacritics. Let us start from the consonantal inventory:

b [b̥]	*p* [ph]	*f* [f]	*m* [m]	
d [d̥]	*t* [th]		*n* [n]	*l* [l]
z [d̥z]	*c* [tsh]	*s* [s]		
j [d̥ʑ]	*q* [tɕh]	*x* [ɕ]		*y* [j] & [ɥ]
zh [d̥ʐ]	*ch* [ʈʂh]	*sh* [ʂ]		*r* [ɻ] ⟦ʐ̝⟧
g [g̊]	*k* [kh]	*h* [x] ⟦χ̞⟧	*ng* [ɴ]	*w* [w]

Of course, all the relevant articulatory illustrations (orograms, labiograms and laryngograms) must be provided, as in our Ch 6. As far as the actual symbols are concerned, ^{off}[C̥] may correspond to ^{can}[C̥, C̬]; this can be positively exploited, after all, by presenting ^{off}[C̥] as 'concise' 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 produced, rather, as an approximant: ^{off}⟦h̞⟧ (^{can}[h]).

For /ʈʂ(h), ʂ, ʐ/, the corresponding 'retroflex' ^{off}[d̥ʐ, ʈʂh, ʂ, ɻ] are fine, but the rel-

fig 21.5. Official vowel diagrams of Chinese 'simplified' pronunciation (with ^{off}*IPA* symbols).



evant orograms should help students distinguish Mandarin postalveolars from true 'retroflex' contoids, articulated further back (*canIPA*: *apicopalatals*). The notation '[[ɻ̟̌]]' dispels almost any doubt about the exact articulation of [ɻ̟].

Since our fonts contain unitary symbols for all stop-strictives, we obviously use them: [ʦ, ʣ; ʧ, ʤ; ʈʂ, ɖʐ]. Most *Unicode* typefaces still have '[ʦ, ʣ; ʧ, ʤ; tʃ, dʒ]', 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 [ʦ, ʣ; ʧ, ʤ; ʈʂ, ɖʐ], avoiding circus things like '[t͡ʂ, t͜s]'.

Regarding the vocalic inventory, there is little to add to what has been clearly explained in § 21.1.2. Reference is to be made to the official trapezia in *fig* 21.5, in order to compare them with our *canIPA* vocograms in *fig* 21.1.

I

i	[i]	*ï*	[ɯ, ɨ] [[ᵾ]]	*ui*	[weɪ] [[e̞]]		
in	[ɪn] [[ɪ̞]]	*ing*	[ɪɴ] [[ɪ̞]]				

E

e	[ɤ] [[ɤ̞]]	*ie*	[jɛ] [[e̞]]	*üe*	[ɥe] [[e̞]]	*ei*	[eɪ] [[e̞]]
en	[ɜn]	*er*	[ɤɻ] [[ɤ̞]]	*eng*	[ɤɴ] [[ɤ̞]]	*ueng*	[wɤɴ] [[ɤ̞]]

A

a	[a] [[ä]]						
ai	[ae] [[e̞]]	*uai*	[wae] [[äe̞]]	*ao*	[ɑo] [[ɑ̟o̞]]	*iao*	[jao] [[äo̞]]
(w)an	[(w)an] [[ä]]			*ian*	[jɛn] [[ɛ̝]]	*üan*	[ɥɛn] [[ɛ̝]]
(i)ang	[(j)aɴ] [[ä]]	*uang*	[wɑɴ] [[ɑ̟]]				

O

o	[ɔ] [[ɔ̞]]	*ou*	[oʊ] [[o̞]]	*uo*	[wo] [[ɔ̞]]
ong	[oɴ] [[o̞]]	*iong*	[joɴ] [[o̞]]		

U

u	[u]	*ü*	[y] [[ẏ]]	*iu*	[joʊ] [[o̞]]
un	[wɜn]	*ün*	[ʏn] [[ɤ̝]].		

For /Çɯ/, as said, we use [ɻ̟̌], *ie* *off*[ɻ̟], and the same contoid can be employed to produce *érhuà*: [Vɻ̟]. In order to transcribe more genuine articulations –those with [ɣ]– the official inventory offers no unambiguous solution. Clearly, we cannot employ *off*[ɻ] to render our *can*[ɥ̟], a *postalveolarized provelar semi-approximant* (with lateral contraction, see § 7.3); the only official alternative is *off*[ɹ].

In *canIPA*, the same symbol is –more conveniently– employed for the most typical realization of American English *r*, a *slightly rounded postalveolarized prevelar approximant* (with lateral contraction, as well), which we could render as '[ɥ̟͡ɻ̟ʷ]' in *offIPA*: [ɥ̟] is quite advanced, [[ɯ̟̇]], and the [˞] diacritic –generically described as implying 'rhoticity'– apparently is the only passable solution to mark lateral contraction. (Here, again, [ɻ̟] stands for a simple slit postalveolar approximant: [ɻ̟̌].)

However, since *off*[ɹ] is usually employed as an *alveolar* approximant, a notation like '[Vɹ]', or '[Vɪ]', could lead to legitimate criticism. As a practical solution, '[ɯɻ]' might work well. Thus, for *can*[ɻ], we would have *off*[ɯr, ʁr, ʊr, ɔr] &c. Anyway, *off IPA* always leaves room for 'personal preferences' (though not always correct, including mixed-up symbols)! Therefore, when using a very small type size, we suggest '[Vɹ]' as less hard to decipher than the tiny 'rhoticity hooklet'.

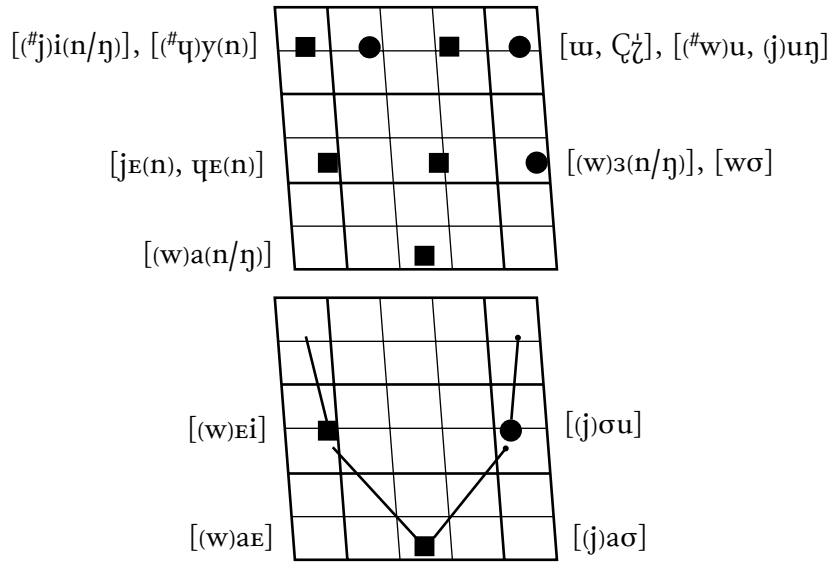
'International' Chinese vowels and consonants

21.3.7. In a book on the pronunciation of Chinese for Italians (*Pronuncia cinese per italiani*), we presented a kind of 'international Chinese pronunciation', which is a simplified, but rigorous, model for teaching purposes. It is sufficient to understand Chinese, and to make oneself understood, because it is based on the native pronunciation, and certainly not a falsification of it.

Only secondary, taxophonic variants are omitted, both for its vowels and consonants (except for a few contoids, which might not be so obvoious for foreigners), while its taxotones are not simplified, because they are much more important, contrary to what common people (but also teachers, even native ones) usually think and do. Even intonation is clearly shown, carefully avoiding mysterious notations (or being completely missing).

fig 21.6 gives the vowels and diphthongs, in two vocograms, and the consonants, in a simple table.

fig 21.6. 'International' Chinese vowels and consonants.



m			n	[ɳ]	[ɲ]			ŋ		
p		t						k		[ʔ]
		ʦ			ʧ	ʨ				
	f	s			ʂ	ɕ				h
					ʐ		j	ɥ	w	
				l						