

Luciano Canepari & Francesca Miscio (2017<sup>2</sup>)

# *Japanese Pronunciation & Accents*

*Geo-social Applications of the Natural Phonetics & Tonetics Method*

9	1. Foreword
11	<i>Why do Phonetics?</i>
13	<i>Typography &amp; canIPA symbols</i>
17	2. Pronunciation & Phonetics
20	<i>The Phonotonic Method</i>
29	3. The phono-articulatory apparatus
33	<i>The vocal folds</i>
38	<i>Resonators (five cavities)</i>
39	<i>The lips</i>
43	4. The classification of sounds
47	5. Vowels & vocoids
53	6. Japanese vowels
58	<i>Japanese Diphthongs</i>
59	<i>Vowel nasalization</i>
59	<i>Vowel devoicing</i>
61	<i>International Japanese pronunciation</i>
62	<i>Yàkuza Japanese pronunciation</i>
63	<i>A note on 'jeipòppu' pronunciation</i>
63	<i>Comparing [CV] sequences with true diphthongs [NV]</i>
65	7. Consonants & contoids
66	<i>Places &amp; manners of articulation</i>
69	8. Japanese consonants
69	<i>Nasals</i>
72	<i>Stops (&amp; stopstrictives, or 'affricates')</i>
75	<i>Constrictives (or 'fricatives')</i>
76	<i>Approximants</i>
77	<i>'Trills'</i>
80	<i>'Palatalization'</i>
80	<i>Loanwords – gairaiō</i>
83	9. Structures
83	<i>Consonant gemination</i>
85	<i>Japanese accent – àkuseñto</i>
88	<i>The very nature of Japanese accent</i>

92	<i>Japanese stress – sutòresu</i>
98	<i>Japanese sentences</i>
101	10. Intonation
103	<i>Tunings</i>
104	<i>Protunes</i>
104	<i>Tunes</i>
107	<i>Parentheses and quotations</i>
108	<i>Japanese intonation</i>
113	<i>Some general paratonetic considerations</i>
115	<i>Some kindlier tune usages</i>
117	11. Texts in phonotonic transcription
118	<i>The North Wind and the Sun</i>
120	<i>Two conversations</i>
128	<i>A very famous 17-mora poem – haiku</i>
129	12. Japanese Mini-phono-dictionary
139	13. Japanese regional accents (& maps)
143	<i>Hokkaido</i>
144	<i>Tohoku (north-west)</i>
145	<i>Tohoku (mid-east)</i>
146	<i>Tohoku (south-east)</i>
147	<i>Kanto (north-east)</i>
149	<i>Kanto (south-west)</i>
151	<i>Chuubu (north-east)</i>
152	<i>Chuubu (south-east)</i>
153	<i>Chuubu (north-west)</i>
154	<i>Chuubu (south-west)</i>
155	<i>Kansai (mid)</i>
157	<i>Kansai (north)</i>
158	<i>Chugoku (north)</i>
159	<i>Chugoku (mid)</i>
160	<i>Shikoku</i>
162	<i>Kyuushuu (north-east)</i>
163	<i>Kyuushuu (mid)</i>
164	<i>Kyuushuu (north-west)</i>
165	<i>Kyuushuu (south-west)</i>
166	<i>Ryuukyuu Islands (Okinawa)</i>
167	14. Foreign accents of Japanese
173	<i>English</i>
175	<i>French</i>
178	<i>German</i>
180	<i>Spanish</i>
185	<i>Italian</i>
188	<i>Portuguese</i>
190	<i>Greek</i>
191	<i>Russian</i>
193	<i>Arabic</i>
195	<i>Turkish</i>
197	<i>Hindi</i>
199	<i>Burmese</i>
200	<i>Malay</i>

201	<i>Filipino</i>
202	<i>Chinese</i>
203	<i>Korean</i>
204	<i>Fijian</i>
205	15. Two extra phonopses
205	<i>Ainu</i>
206	<i>Old Japanese</i>
207	16. Japanese accent of English
211	17. Lexical combinations
219	18. Shin-Meikai Nihongo Akusento Jiten – Appendixes
245	19. Phonopses of 26 languages
263	20. Bibliography
265	<i>Official IPA chart</i>

The *place names* we indicate will be fully transliterated (using our system) only when given in *italic*. They also serve as language examples. Instead, in general titles and expositions, treating them as international words, we will omit accent (eg *à* → *a*), ‘long’ vowels (eg *aa* → *a*), voiceless *i* and *u* (eg *i*, *u* → *i*, *u*), moraic *n* (eg *n̄* → *n*), nasal *g* (eg *ḡ* → *g*). For instance: Tokyo, Nagasaki, Nihon, but *Tookyoo*, *Naḡasaki*, *Ni-hòñ*...

fig 13.0.3. The 20 pronunciation koinés (with some places, including *Okinawa*, in the *Ryūkyū Islands*, between *Taiwàn* and *Kyūshū*, in southern Japan).



# 6. Japanese vowels

6.1. Japanese has only five vowels, which can be distinctively short or 'long' (or rather *doubled*, being realized as *monotimbric diphthongs*), and they can combine into different kinds of sequences (as happens in a slow and careful pronunciation of Spanish in *Saavedra* or *La Habana* [ʃaa'βɛ'ðra, laa'βa'na]).

However, too often Japanese pronunciation is hastily 'described' as having the vowels of Spanish and the consonants of English. We will see that this is not the case at all.

fig 6.1 gives the orograms of the articulations of these vowels: *i*, *e*, *a*, *o*, *u* [i, ɛ, ɐ, σ, ɯ], *ii*, *ee*, *aa*, *oo*, *uu* [ii, ɛɛ, aa, σσ, ɯɯ]; fig 6.2 shows their vocogram, which is much more precise. In addition, fig 6.3 presents their labiograms and palato-grams, to give an overall view.

fig 6.1. Neutral Japanese vowels: orograms.

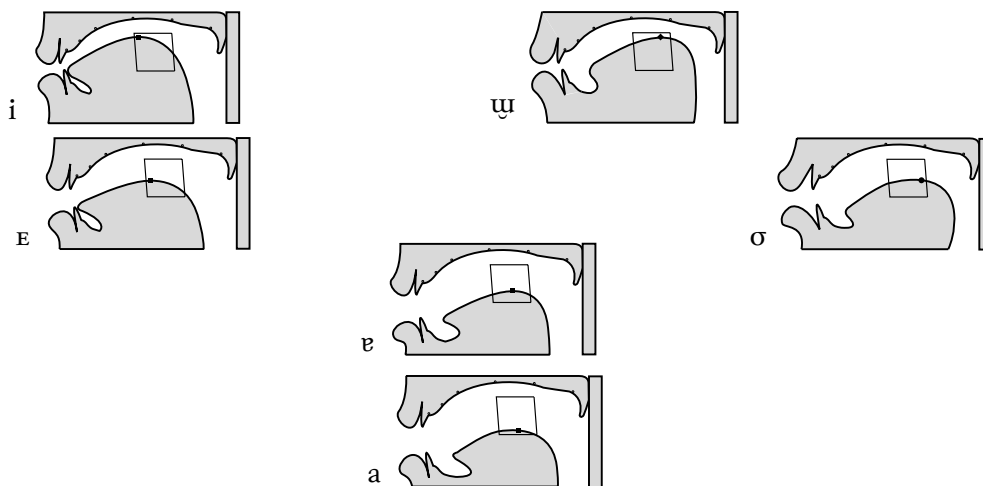


fig 6.2. Neutral Japanese vowels: vocogram.

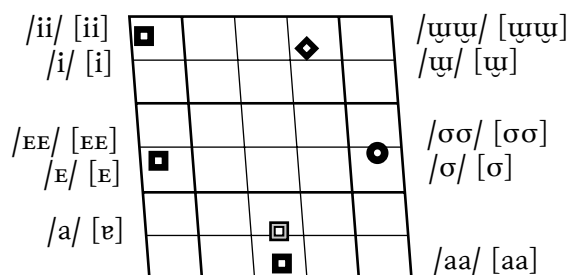
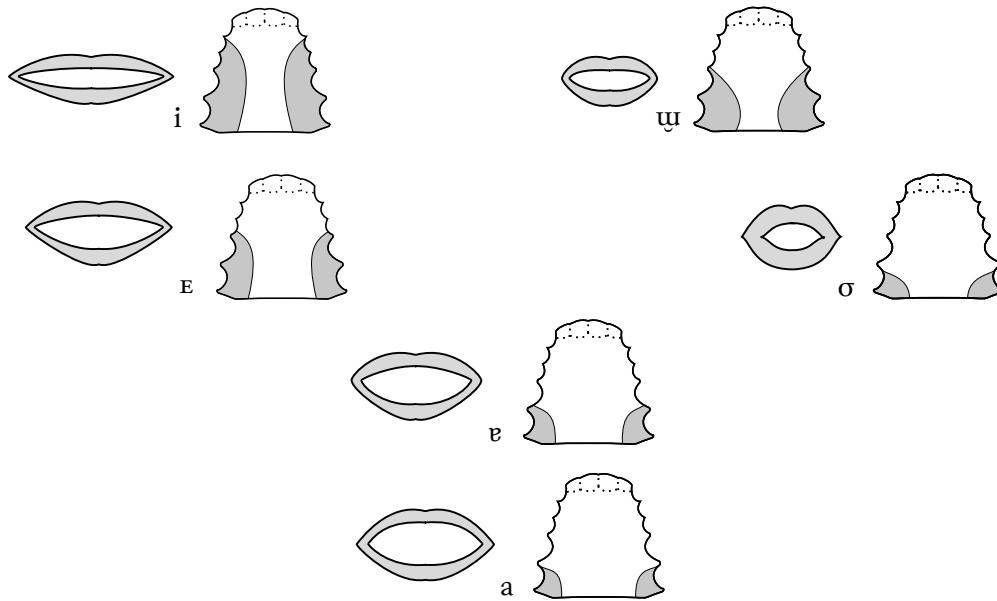


fig 6.3. *Neutral* Japanese vowels: labiograms & palatograms.

6.2. Since ‘long’ vowels are phonemic sequences (and phonetic geminations), the marker of (short) /a/ [ɐ] is grey, seeing that it is articulated as a less open vowel (even in stressed syllables).

In neutral Japanese pronunciation, the most problematic vowel (of course for non-native speakers, but the most variable socially and regionally even for native speakers) is *u* /ɯ/ [ɯ], which lacks the typical and full lip rounding of [u], so widespread in many languages. But let us notice that in neutral Mandarin-Chinese pronunciation we find exactly /ɯ/ [ɯ] (articulated with completely spread lips, cf fig 5.1-5, especially fig 5.5) in the syllables which *pīnyīn* writing shows as *ci*, *zi*, *si* (although we indicate them as *cī*, *zī*, *sī*, in a friendlier way – again, cf fig 5.1-5).

Let us carefully notice that a sadly false Chinese tradition, instead of [ɯ], still uses the *non-IPA* symbol ‘[ɿ]’, which in the *canIPA* system, more coherently, stands for the main taxophone of *r*: /ɿ/ [ɿ, ɿ̃].

Besides, [ɯ] is articulated with the tongue dorsum in a fronter position than for the [u] used in most languages; in fact, it is back-central, not simply back (thus corresponding to the tongue position of a rounded [ɯ]).

Let us now see some examples for each vowel phoneme: *iè* [i-ɛ], *giñkoo* [-ɰiŋ̃·kɔσ], *sabishii* [sɛ-bi.çĩi], *eieñ* [-ɛi·ɛŋ̃], *ègao* [-ɛ.ŋɛσ], *karate* [kɛ-ɾɛ·tɛ], *ài* [-ɛi], *zasshì* [-dzɛ̃·ç̃·ç̃·ĩ], *mato* [mɛ-tɔ], *nèko* [-nɛkɔ], *okìru* [σ-ci.ɿ̃], *otokò* [σ-tɔ·kɔ], *sumì* [sɯ-mi], *àkuma* [-ɛ\_kɯ.mɛ], *mushì* [mɯ-çĩ].

6.3. For tautomorphemic *ei* and *ou* (belonging to the same lexeme) –but the latter is generally transliterated as *oo* (as the former can unsatisfactorily be *ee*)– it is ‘normal’ to have [ɛɛ, σσ]: *señsèi* [sɛ̃·sɛ̃], *suiei* [sɯ-iɛɛ], *eikyoo* [-ɛɛ·ç̃σ], *doshuu* [-σσ·ç̃jɯɯ]; *oogōñ* [-σσ·ŋ̃], *ooeñ* [-σσ·ɛ̃]. However, spelling and education encourage more and more that *ei* /ɛɛ/ should become exactly *ei* /ɛi/, in modern neutral pronunciation.

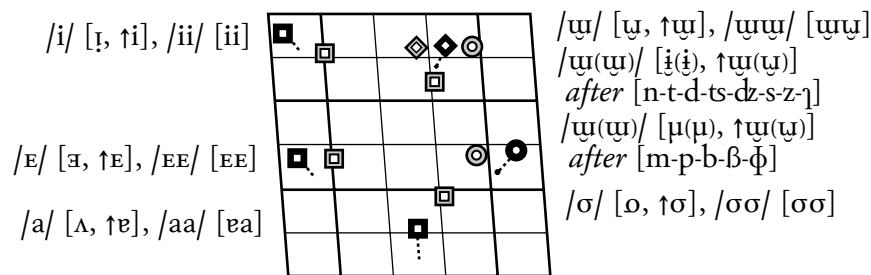
Instead, for heteromorphemic *ei* and *ou* (in more or less simple lexical compounds), it is certainly normal to have only [ei, σɥ]: *keito* [-kei·tσ], *sèi* [-sei], *arasòu* [-e·ɣe-sσɥ], *sou* [-sσɥ].

There are also some *colloquial variants* of the Japanese vowels. As fig 6.4 shows, the short ones have two possibilities: they can remain more or less as in the normal neutral accent (black markers, indicated as ↑, because they can be used intentionally, sometimes with firm commitment). Or they can be centralized (grey markers). Of course, oscillations between the two positions are quite possible, for each element.

Thus, we have: /i/ [ɪ, ↑i], /e/ [ɛ, ↑e], /a/ [ʌ, ↑ə], /σ/ [ɔ, ↑σ], /ɥ/ [ɥ, ↑ɥ]. And /ii/ [iiʌ], /EE/ [EEʌ], /aa/ [əa], /σσ/ [σσʌ], /ɥɥ/ [ɥɥ].

Notice that, for /ɥ/ [ɥ, ↑ɥ] and /ɥɥ/ [ɥɥ], we have to indicate the possibility of two additional taxophones. In fact, after [m, p, b, β] (including the very frequent intervocalic variants of /b/ [β, β̥, β̃]), we often find fully labialized vocoids: [μ(μ)]. On the contrary, after [n, t, d, ts, dz, s, z, ɲ, ɺ] (and also [ɺ, r, l, ɺ, ɺ̥, ɺ̃], and others, for /ɺ/ as we will duly see), /ɥ(ɥ)/ can often become [ɿ(ɿ)] (central, rather than back-central [ɥ(ɥ)]).

fig 6.4. *Colloquial* vowel variants.

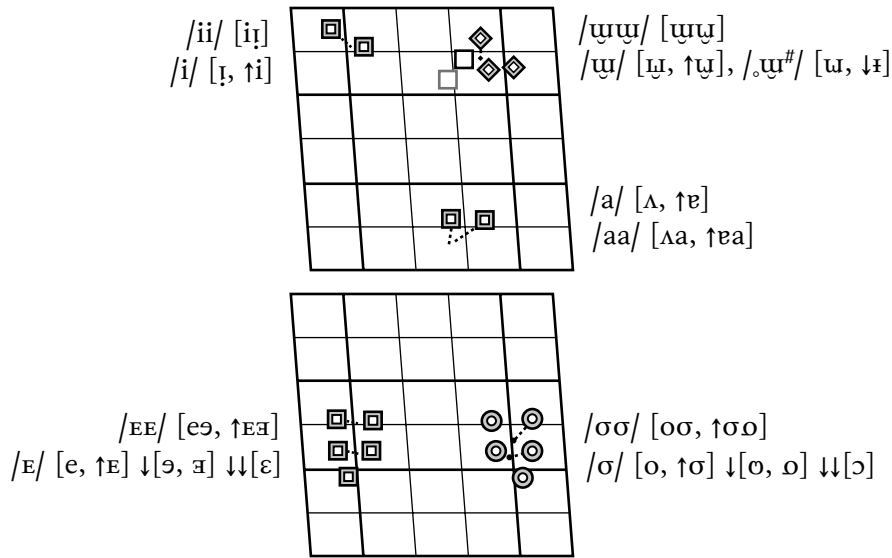


6.4. In *mediatic pronunciation*, mostly the younger generations articulate /ɥ/ in a fronter position, as high central –with partial or full rounding, [ɿ] or [ɥ], respectively. But we always notate [ɥ], since those are youth characteristics which generally change into [ɥ], sooner or later. It is interesting to note that this peculiarity continues in relative time, without really changing the structure of neutral pronunciation, in spite of any colloquial *or* mediatic trends. Obviously, real non-neutral pronunciations are a different thing.

As fig 6.5 shows, in *mediatic pronunciation*, there are other differences in comparison with both *neutral* and *colloquial* pronunciations. In fact, the short vowels can be even more centralized, as can be seen. But, most of all, they can be realized by means of more different vocoids, and greater variations, including some socio-phonetic peculiarities.

Let us carefully look at fig 6.5, and consider where each vocoid is placed, guiding ourselves with the arrows: positive ↑, negative ↓. It can also be seen that unstressed final /ɥ#/ has completely spread vocoids: [ɥ, ↓ɥ]. In addition, let us pay particular attention to the great variation and peculiar realizations of /e/ [e, ↑e] ↓[ə, ɛ] ↓↓[ɛ] and /σ/ [o, ↑σ] ↓[ɔ, ɔ] ↓↓[ɔ]. Ditimbric narrow diphthongs realize the five ‘long’ vowels.

fig 6.5. Principal *mediatic* variants.



6.5. Passing, now, to fig 6.6, it is easy to find further and broader taxophones, including gender and age sociophonic ones for /ɰ/, /ɰɰ/, /aa/. So, we find that especially *male* speakers can have a backer /ɰ/ [ɰ], %or lower realizations: ↓[ɰ, ɰ]; also /aa/ [↓↓ʌʌ], /a/ [↓ʌ].

Perhaps, the most peculiar variation is to be found for /aE/ and /ai, si/, which can have some very marked overlapping realizations: [↓↓ɐɛ, ↓↓↓aɛ] for the three of them; including [↓↓↓EE], in common with /EE/ [EE, ↓EE]. All these are shown in fig 6.6.

fig 6.6. Further and broader *mediatic* variants: *m(ale)*, *f(emale)*, *y(oung)*.

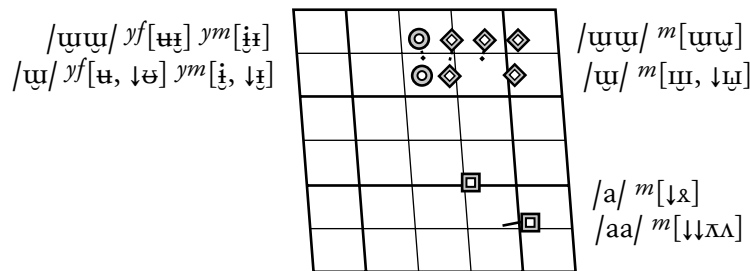


fig 6.7. *Orograms & labiograms* of some high vocoids, for comparison between neutral, colloquial, mediatic, and international Japanese accents (including some foreign elements, too).

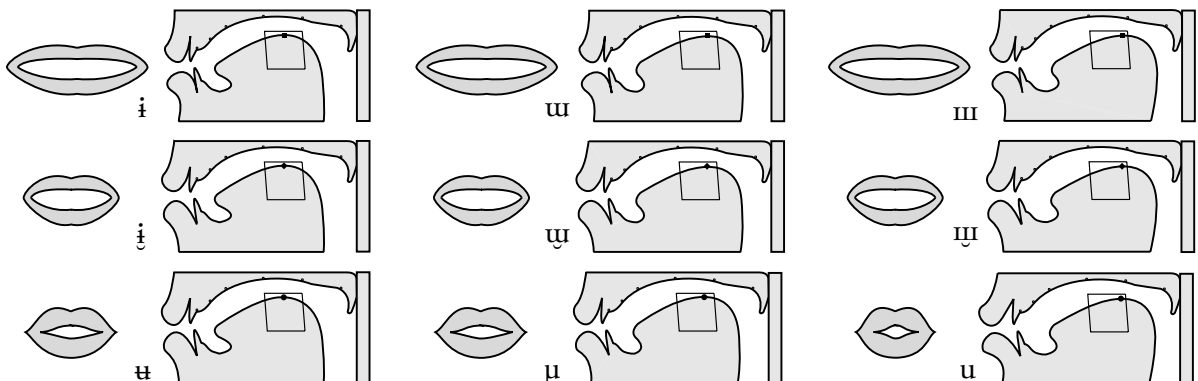




fig 6.8.1. Neutral Japanese diphthongs (see next page).

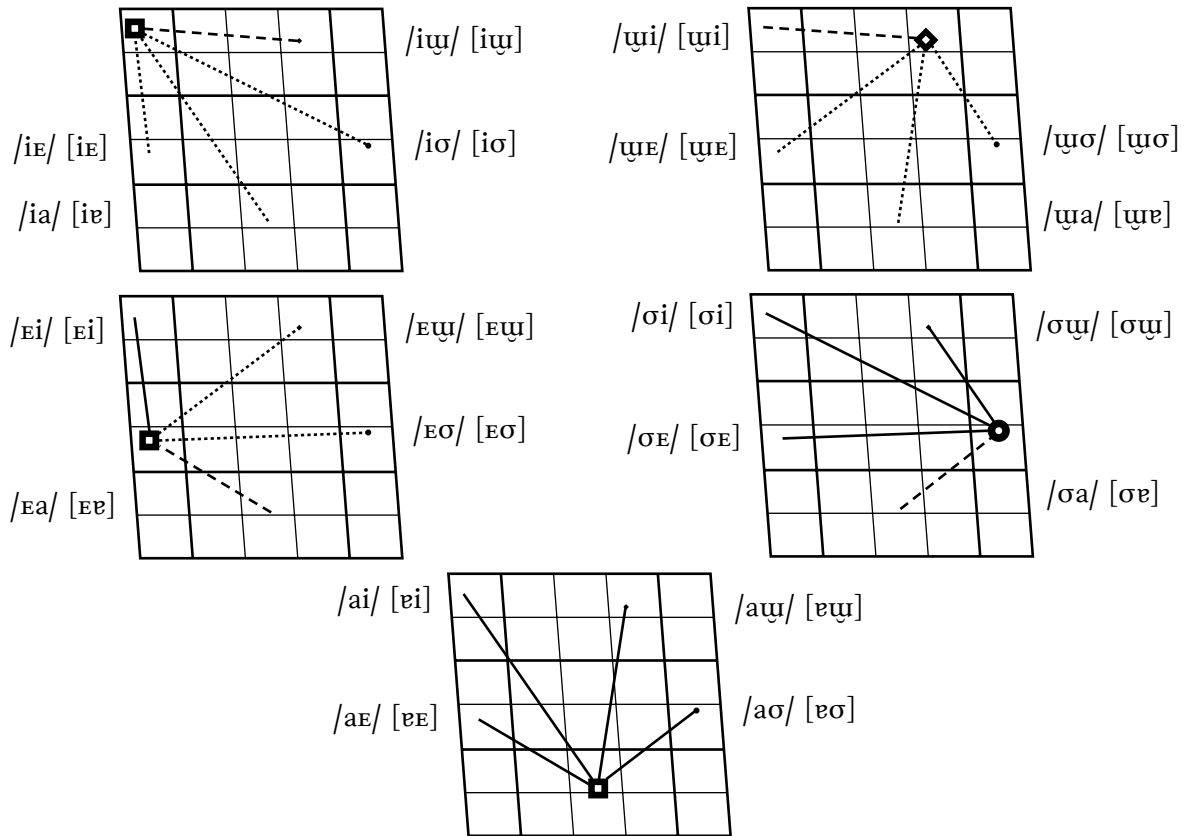
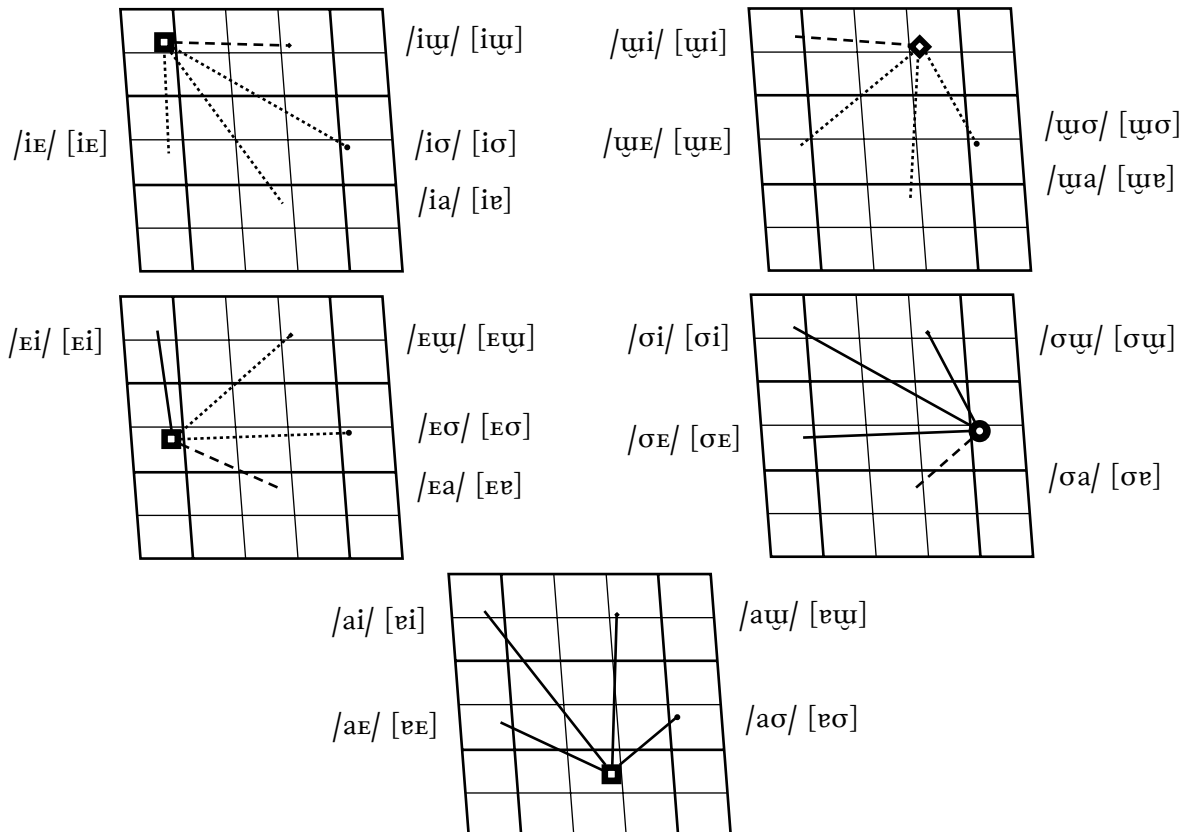


fig 6.8.2. Possible normalized colloquial diphthongs.



**Japanese diphthongs** (including the preceding page)

6.6. Each Japanese vowel can combine with any other, thus forming actual *diphthongs*: [VV] (as fig 6.8.1 shows). Also clearly notice that each vowel can even form a monotimbric diphthong by combining with itself, as we have seen above: *ii, ee, aa, oo, uu* [ii, ee, aa, oo, uu]. Thus, as already said, for various reasons, we are firmly convinced that –for Japanese– the best notation (either phonic or graphic) is [VV] VV (also to unambiguously show on which mora an *accent* may be).

In fig 6.8.1-3, diphthongs are shown by means of different lines: *unbroken* lines for *current* diphthongs; *broken* lines with *larger* strokes for *rare* diphthongs; *broken* lines with *shorter* strokes for only *possible* diphthongs.

fig 6.8.3. Possible normalized *mediatic* diphthongs.

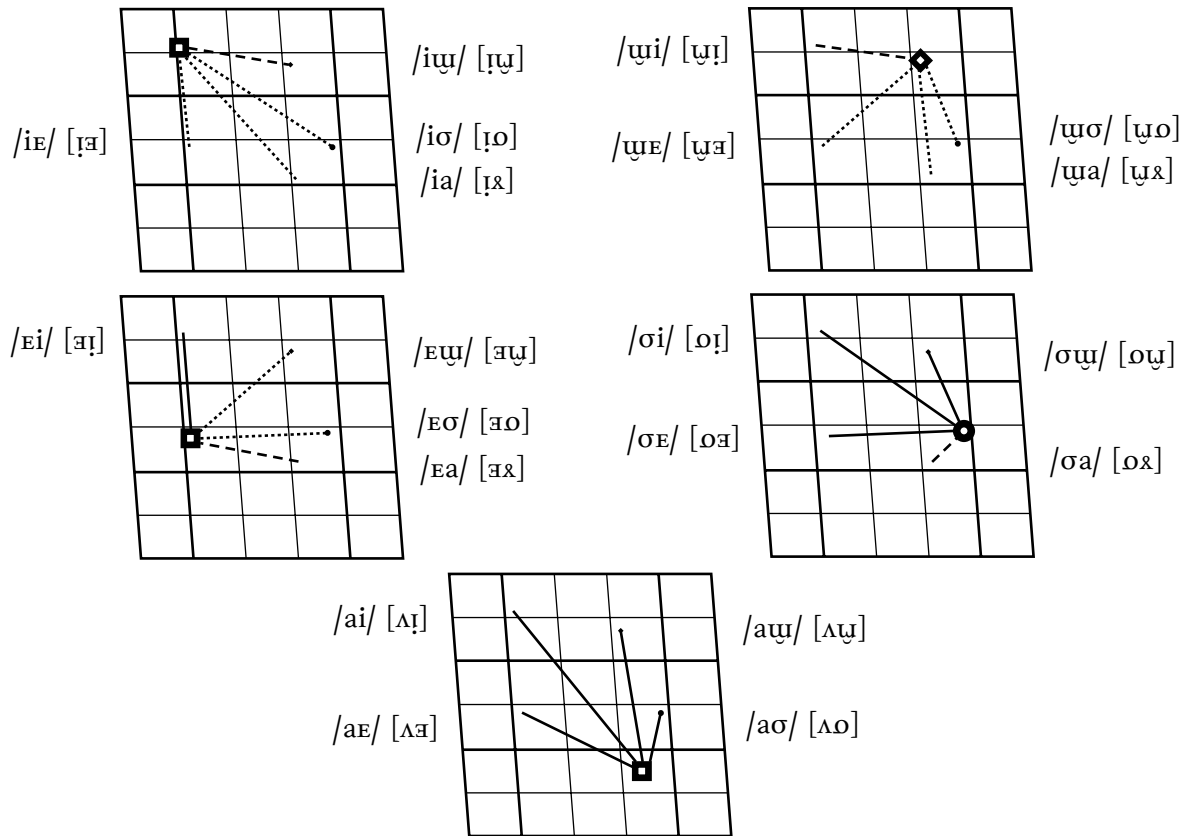
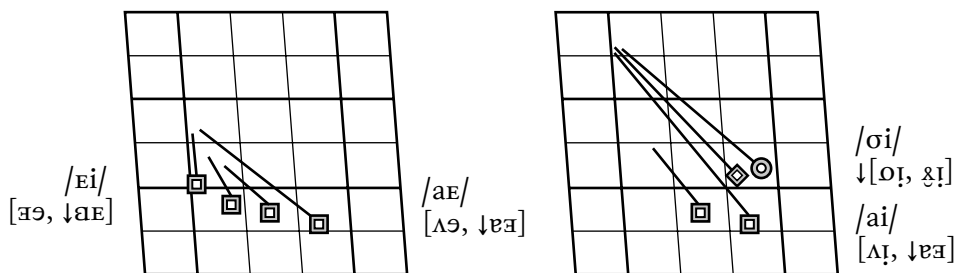


fig 6.8.4. Broader *mediatic* diphthong variants.



### Vowel nasalization

6.7. Besides, in neutral Japanese the last vocoid in a checked syllable in /ŋ/ is slightly nasalized: [Ṽ̥], but not fully nasalized, [Ṽ̃]. (Notice that our notation is more logical and natural than the official *IPA* one, which does not have such a possibility, and –recklessly– uses ‘[V̥]’ to transcribe creaky voice, corresponding to our [V̥]).

The production of *oral* vocoids, in comparison with seminasalized and fully nasalized vocoids, can be clearly seen in fig 5.8. Of course, for oral segments, the velum is raised and closes the passage to the nose (1st row). For (fully) *nasalized* segments, the velum is lowered and the passage is free (3rd row), while, for *seminasalized* segments, the velum is only partially lowered, so that the passage to the nasal cavity is very narrow, thus only a little amount of air can go through the nose.

As we know, neutral French pronunciation has *four* nasalized vowel phonemes, realized as *six* fully nasalized vocoids, with no nasal contoid after them, even in stressed syllables (fig 5.9 shows them with their correspondent oral articulations).

French examples: /ã/ [ã̃, ̃ã], /œ/ [œ̃, ̃œ], /ɔ̃/ [ɔ̃], /õ/ [õ̃], as in *incertain* [ã̃sɛʁˈtã̃], *un brun* [œ̃bʁœ̃], *pendant* [pɑ̃dɑ̃], *mon nom* [mɔ̃nɔ̃]. In stressed checked syllables, we have [Ṽ̃]: *peindre* [pɑ̃dʁ̃], *humble* [œ̃b̃], *lampe* [lɑ̃p̃], *nombre* [nɔ̃bʁ̃]. Here is a ‘funny’ example: *un bon vin blanc* [œ̃bɔ̃ ˈvã ˈblã̃].

These phonemes can even be followed by nasal contoids (originally belonging to a different syllable): *ennui* [œ̃nɥĩ], *non-être* [nɔ̃nɛtʃ̃], *emmener* [œ̃mˈnẽ] (*em-mener* [œ̃mœ̃nẽ]). By the way, let us notice that a Japanese word like *hòn* [hɔ̃ŋ̃] must not become simply [õ̃]!

In *colloquial* and *mediatic* Japanese pronunciations, the nasalization of vocoids is stronger. In fact, instead of becoming simply [Ṽ̥ŋ̃] (seminasalized), they are fully nasalized, [Ṽ̃ŋ̃], and diphthongs, instead of [Ṽ̥Ṽ̥ŋ̃], become [Ṽ̃Ṽ̃ŋ̃]. In addition, also in free syllables, either stressed or not, we hear seminasalized vocoids and diphthongs when immediately followed by a nasal contoid: [Ṽ̥#N] and [Ṽ̥Ṽ̥#N].

Furthermore, in *colloquial* and *mediatic* pronunciation, even a tautosyllabic nasal contoid partially nasalizes the vowel it precedes: [NṼ̥(#)] and [NṼ̥Ṽ̥(#)]. This occurs both within words and phrases, or at the end of them.

### Vowel devoicing

6.8. Besides, neutral Japanese has the peculiarity of presenting *vowel devoicing* (cf fig 3.4). Thus *voiceless vocoids* are produced, *ie* with no vibration of the vocal folds, as also happens with consonants, such as (voiceless) [s] /s/ in comparison with (voiced) [z] /z/. More appropriately, these voiceless vocoids are lenis too (or lenited), *ie* only the arytenoids are open, as for [h] (again, cf fig 3.4.F+).

In neutral pronunciation, vowel devoicing concerns only [ĩ̥, ũ̥] between voiceless consonants, or between a voiceless consonant and a pause, generally when in low-pitch syllables, but also in mid-pitch syllables, however never on ‘accented’ morae (after which the pitch becomes low), and never in interrogative tunes

(which produce a raising of the basic pitch): *kīshī* [ci̇-çi̇], *kīshī* [-ci̇çi̇], *hītōtsu* [h̥i̇-tōts̥u̇], *kuchibiru* [k̥u̇-t̥çi̇-bi̇ɾu̇], *takusān* [t̥ėk̥u̇-s̥ėn̥].

Among the most frequent cases of /i, ɥ/ devoicing we certainly find the following grammeme elements: *dēsū* [-d̥ės̥u̇], *-masū* [-m̥ės̥u̇], *-shīta* [-çi̇t̥ė]. From a true phonetic point of view, they surely undergo resyllabication, mostly becoming: *dēsū* [-.d̥ės̥], *-masū* [-.m̥ės̥], *-(V)shīta* [-.(V)çi̇t̥ė].

Of course, the same is true within lexemes, as in the case of those just seen above, which can certainly be heard as: *kīshī* [-ççi̇], *hanashī* [h̥ė-n̥ėççi̇], *hītōtsu* [-h̥tōts̥], *kuchibiru* [k̥t̥çi̇-bi̇ɾu̇], *takusān* [t̥ėk̥s̥ėn̥], *akūhyō* [̥ėk̥u̇-h̥j̥ȯ].

Let us notice that the first three never become anything like ‘[-k̥çi̇, h̥ė-n̥ės̥, -h̥tōts̥]’. In fact, the contoids that precede them maintain their coarticulations. This surely hints at the vocoids that have been dropped.

We will indicate each case where /i, ɥ/ can be devoiced and dropped. But, depending on the speed of utterance, some of them may be only partially devoiced, in *slower* speech: [i̇, ɥ̇], or even fully voiced [i, ɥ], instead of [i̇, ɥ̇].

6.9. In *faster* speech, more than one vowel (and even in consecutive syllables) can be devoiced or dropped: *bakuchiku* [b̥ė-k̥u̇-t̥çi̇-k̥u̇, ̥b̥ėk̥u̇-t̥çi̇-k̥u̇], and even *tsukukubōshī* [ts̥u̇-k̥u̇-t̥s̥u̇-k̥u̇-b̥ȯççi̇, ̥ts̥u̇k̥u̇], *rekishiteki* [̥ɾėçi̇çi̇t̥ėçi̇, ̥ɾėçi̇], *fukushikikōkyū* [f̥u̇-k̥u̇çi̇çi̇k̥ȯççi̇, ̥f̥u̇k̥u̇, -çi̇çi̇], but with some [i, ɥ], in *slower* speech.

Of course, a syllable with a devoiced/dropped vowel cannot keep its original stress (rather theoretical in neutral pronunciation, by now, more a *traditional* kind of pronunciation), which is shifted, then: *mushikēn* [̥m̥ɥ̇çi̇k̥ėn̥, m̥ɥ̇çi̇k̥ėn̥].

Instead, in (more) *modern* pronunciation –and also in *mediatic* and *colloquial* pronunciations– we can certainly find *mushikēn* [m̥ɥ̇çi̇k̥ėn̥], too.

The same goes for *kitte*, *tsūku*, *kakūsu*, *akikaze* which *traditionally* become [çit̥ė, ts̥u̇-k̥u̇, -k̥ėk̥u̇s̥u̇, ̥ėçi̇k̥ėzė] (*kittle*, *tsukū*, *kakūsu*, *akikaze*), changing stress pattern. However, *modernly* they are [-çit̥ė, -ts̥u̇k̥u̇, k̥ė-k̥u̇s̥u̇, ̥ėçi̇k̥ėzė], compromise realizations, but of course, a partially or totally devoiced [i̇, i̇, ɥ̇, ɥ̇] can be heard, too: [-çit̥ė, -çit̥ė, -ts̥u̇k̥u̇, -ts̥u̇; k̥ė-k̥u̇s̥u̇, --k̥u̇, ̥ėçi̇k̥ėzė, ̥ėçi̇, ̥ėçi̇k̥ėzė, ̥ėçi̇].

In order to better show the peculiarities we are dealing with, in these examples we will use only neutral segments (instead of those shown in fig 6.4-7 & fig 6.8.2-4 & fig 6.10.1-3).

On the other hand, we often find ‘intermediate’ accents between true *neutral* ones and more typical *mediatic* or *colloquial* accents, even in the recordings of many teaching courses, and not only those for general learning, but also some on the pronunciation of Japanese.

Some further examples: *kikū* [çi̇-k̥u̇], *bakufu* [-b̥ėk̥u̇f̥u̇, ̥f̥u̇], *susumu* [s̥u̇-s̥ɥ̇m̥ɥ̇], *chikarā* [t̥çi̇k̥ėɾė], *hitori* [h̥i̇-t̥ȯɾi̇], *kippu* [çip̥p̥ɥ̇], *shīshī* [-çi̇çi̇], *chisso* [-t̥çi̇s̥s̥ȯ], *kikichigai* [çi̇çi̇-t̥çi̇ɾėi̇, ̥çi̇çi̇], *chīshiki* [-t̥çi̇çi̇çi̇, ̥çi̇], *nyūgakūshiki* [̥ɲ̥j̥ɥ̇ɾėk̥u̇çi̇çi̇, ̥ɲ̥j̥ɥ̇ɾėk̥u̇çi̇çi̇], *kishitsu/kishitsu* [-çi̇çi̇-t̥s̥u̇, ̥çi̇çi̇]

·ts̥i̥], *fuhitsuyoo/fuhitsùyoo/fuhìtsùyoo* [·f̥i̥-ɰi̥ts̥ɰ.j̥σσ, ·f̥i̥-ɰi̥-t̥ɰ.j̥σσ, ·f̥ɰi̥-ɰi̥-], *fukù/fùku* [·f̥i̥-k̥ɰ, -f̥ɰ.k̥i̥], *àki* [-e̥.ɰi̥], *kàrasu* [·k̥e̥.ɰe̥.s̥i̥, -k̥e̥.ɰe̥.s̥i̥], *àki-no* [-e̥.ɰi̥.n̥σ, ·e̥.ɰi̥.n̥σ], *kàrasu-wa* [·k̥e̥.ɰe̥.s̥ɰ.ɰe̥, ·k̥e̥.ɰe̥.s̥ɰ.ɰe̥].

6.10. In *colloquial* and *mediatic* pronunciation, we also find another degree of devoicing: *final half-voicing*. It applies, instead, to *all* vowels (either short or long, and stressed or not) before a *pause*, preceded by any consonant (not necessarily a voiceless one; in fact this also occurs in diphthongs, [V̥V]). Therefore, before a pause, the last vocoid is [i̥, e̥, ɸ̥, σ̥, ɰ̥]. Obviously, short /i, ɰ/ can become [i̥, ɰ̥], in the condition of total devoicing: in the context [C̥V].

It is important to state that this type of devoicing is syntagmatic, which means that we have one vocoid –not two– which begins as voiced and ends as voiceless (cf fig 3.4.G). So its first part is voiced, whereas the second one is voiceless (note that a superscript symbol here means ‘half a symbol’): [V̥V] – and this is hinted at by the pause context, even if the actual pause is short.

In simpler terms, we could say that the auditory effect is almost that of a very short [h], [V<sup>h</sup>] or a semiapproximant [Vh]. Indeed, for emphasis, strictly speaking (or for other paraphonic implications connected with states of mind), often an actual [Vh] sequence may be heard: *Sòo dèsu ka?!* [i̥-s̥σσ.d̥e̥.s̥i̥.k̥e̥(h)].

Most typically, in *mediatic* (and *colloquial*) pronunciation, vowel devoicing can also concern the other three (short) vowels, in addition to /i, ɰ/, but less systematically.

In fact, we can find [e̥, σ̥], especially in the first syllable of words, when followed by a syllable containing the same vowel and another voiceless consonant: *katanà* [k̥e̥-t̥e̥.n̥e̥], *hakà* [h̥e̥-k̥e̥], *kakànai* [k̥e̥-k̥e̥.n̥e̥i̥], *torikakaru* [t̥σ̥.ɰi̥.k̥e̥-k̥e̥.ɰ̥], *hokorì* [h̥σ̥-k̥σ̥.ɰi̥], *tokoro* [t̥σ̥-k̥σ̥.ɰ̥], *hosòì* [h̥σ̥-s̥σ̥i̥].

More rarely, also /E/ can be devoiced, [E̥]: *kesanai* [k̥E̥-s̥e̥.n̥e̥i̥], *kesòo* [k̥E̥-ê̥.j̥σσ], *sekkaku* [s̥E̥k̥-k̥e̥.k̥i̥].

In addition, in *mediatic* pronunciation (and *colloquial*, and some *regional* ones, as well), even accented vowels can be devoiced: *hanashi* [h̥e̥-n̥e̥-ê̥i̥] → [h̥e̥-n̥e̥-ê̥i̥̥].

Let us add that, in *mediatic* pronunciation, initial vocoids, between a pause and a voiceless contoid, can also become [V̥] – mostly /i, ɰ/ (but occasionally also /E, a, σ/): *ishì* [i̥-ê̥i̥], *utāgai* [i̥-ɰ̥-t̥e̥.ɰe̥i̥], *akubi* [e̥-k̥ɰ̥.bi̥].

On the contrary, word-final vocoids before a pause are often produced with *creaky voice*, [V̥̥], instead of the partial devoicing seen above, [V̥].

On the other hand, those [V̥] can also be changed into [V̥̥]: instead of assuming creaky voice, they can be followed by a short and anobtrusive [ʔ], on purpose indicated as [ʔ].

It is to be noted that, for *emphasis*, both initial and final vowels in a word are often accompanied by a real [ʔ]: [#ʔV, Vʔ#], even with no pause at all. But, the communicative effect remains quite different and clearly distinguishable. This can also happen when, even without emphasis, we find [V#V].

## International Japanese pronunciation

6.10. Let us say a few words on a type of pronunciation, which differs somehow from all the other kinds we have seen. In fact, while those are actual native accents, this one is a simplified version, most suitable for teaching Japanese pronunciation to foreigners, with [a, μ, μμ]. This can be sufficient; but, of course, the acquisition of proper [ɯ, ɯɯ] would be (much) better.

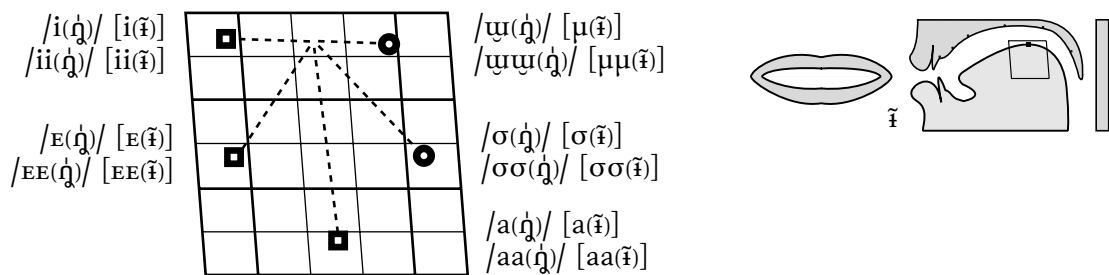
We call it ‘international Japanese pronunciation’: it has only five *stable* vocoids, both short and doubled (as monotimbric diphthongs for ‘long’ vowels): [i, ii; ɛ, ɛɛ; a, aa; σ, σσ; μ, μμ]. Thus, no [ɐ], but rounded [μ, μμ] (back-central, like [ɯ, ɯɯ]), and different from the very widespread [u, uu], found in so many languages, which sound too non-Japanese).

Of course, the international *diphthongs* are the combinations of the basic international vocoids. In this kind of pronunciation, no *nasalization* (even light) is necessary. But, clearly, if someone produces either partial or full nasalization, no native speakers will object.

As regards vowel *devoicing*, in slow <sup>or</sup> poor Japanese, its absence does not bother much, including hesitations and false starts. But, of course, native hearers would appreciate much more a fluent and sure rate. Thus, it is certainly a good idea to imitate native speakers as far as possible, especially for *dèsu* and *dèshita*, and other forms in *-shita*, and similar ones.

Of course, instead of actual structures like [-dɛ.ʂt̚ɯ̄], [-dɛ.ç̣ị.t̚ɛ̄], with more or less clearly perceptible two or three syllables, although with [CV̄], it is more than sufficient to simply utter [-dɛʂ], [-dɛʂ:], [-dɛʂ:], [-dɛç̣.t̚ɛ̄], [-dɛç̣:t̚ɛ̄], [-dɛç̣.t̚ɛ̄], &c.

fig 6.9. *International* Japanese vocoids (also in /V̄/ sequences).



## Yàkuza Japanese pronunciation

6.11. Let us also add a few words on a peculiar type of pronunciation, which –again– differs from the typical kinds of pronunciation that foreign learners are supposed to find among native speakers: how *yàkuza* [-jɐ.kɯ̄.zɐ̄] members are heard on Japanese films.

Thus, fig 6.10.1-3 show the most typical realizations of the vowels and diphthongs of this kind of pronunciation. Of course, some oscillations towards mediatic, or colloquial or even neutral accents, are possible, as also are some towards re-

gional accents, too. For the *consonants*, a few observations will be found about *r*, in § 8.2.5 & fig 8.13. Here, let us add that two paraphonic settings are quite typical: a lowered larynx (·:), restricted tongue (Y). In addition, short final vowels have various possible, and typical, realizations, too: /V<sup>#</sup>/ [V<sub>Δ</sub><sup>#</sup>], [Vh<sup>#</sup>], [V<sub>o</sub><sup>#</sup>], [Y<sup>#</sup>].

fig 6.10.1. *Yakuza* Japanese vowels.

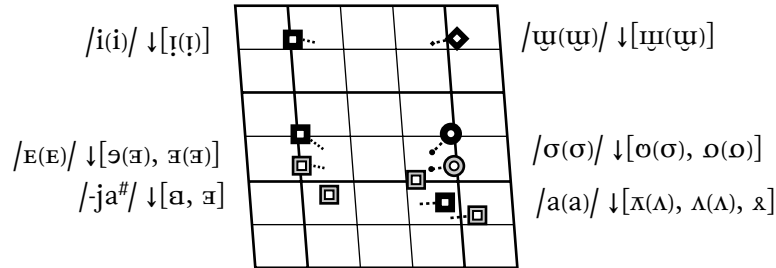


fig 6.10.2. *Yakuza* Japanese diphthongs.

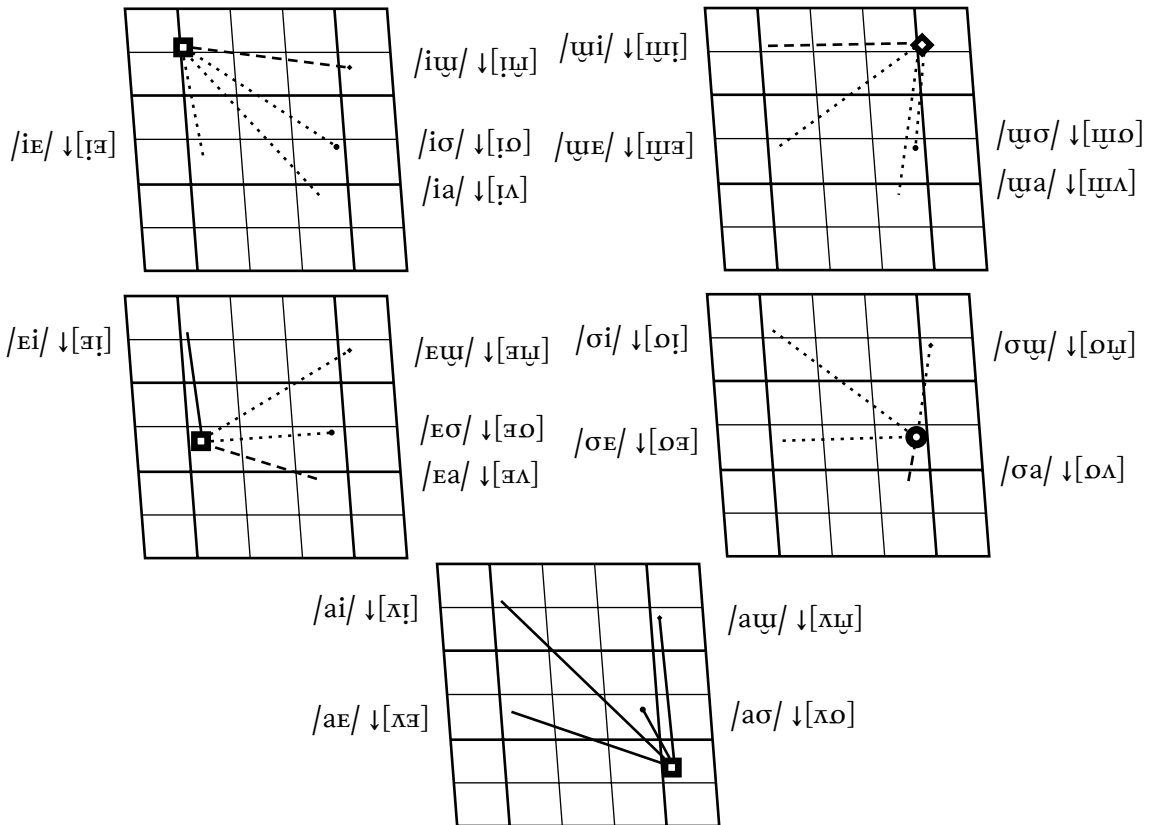
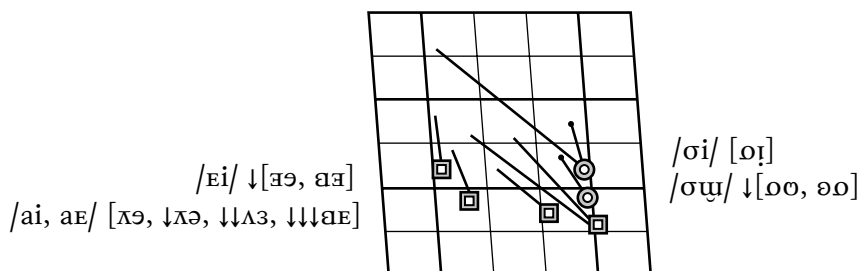


fig 6.10.3. *Yakuza* Japanese broader diphthong variants.

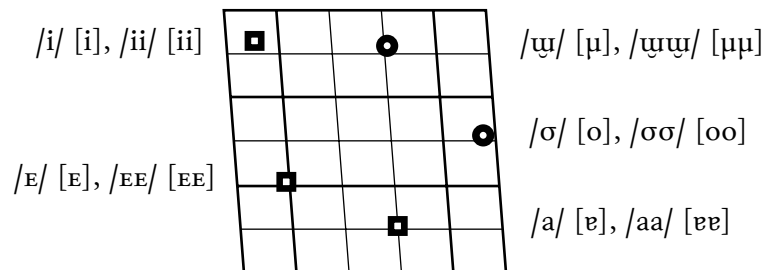


### A note on *jeipòppu* pronunciation

6.12. Certainly not for any cultural interest, but for some concise observations on its kind of pronunciation, let us say a few words about ‘j-pop’ (also containing unconnected English or French phrases, with phones which try to come closer to the original pronunciations).

fig 6.11 shows its *vowels*, implying that its diphthongs are the combination of the vocoids given there. As for the *consonants*, let us observe that very often we hear *b* [β], *t*, *d* [t, d], *r* [l]. But, the most striking thing is the *prosodic* treatment of its phrases. They are almost always stressed on the last syllable, and even on the last mora of diphthongs, like *iranai* said with [-e<sup>i</sup>#]. While its *pitch-accent* is practically absent, due to the music.

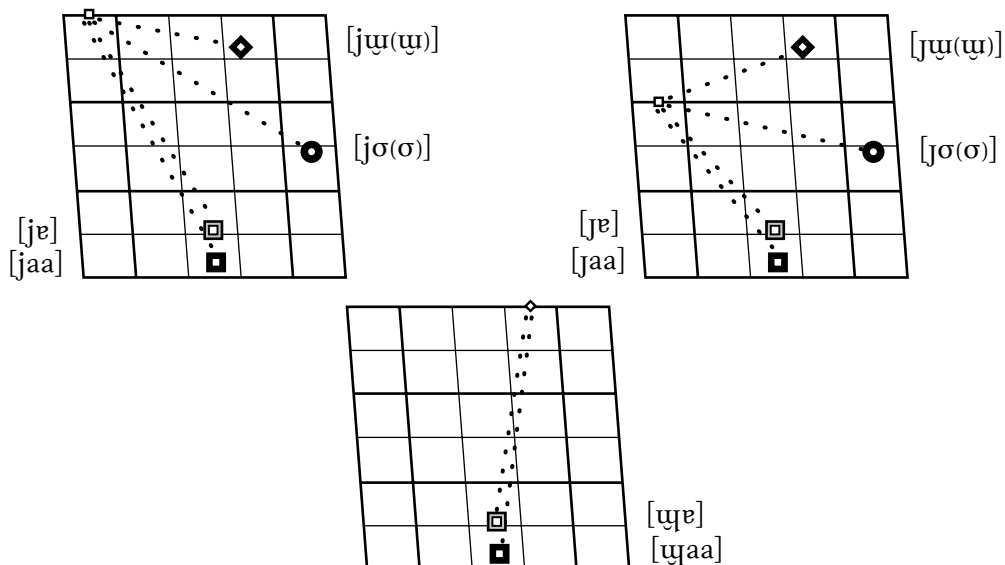
fig 6.11. *Jeipòppu* vowels.



### Comparing /CV/ sequences with true diphthongs /VV/

6.13. Here is an iconic way to show neutral sequences of approximant (or semi-approximant) consonants and vowels, which are not ‘diphthongs’. Let us carefully compare fig 6.12 with fig 6.8.1, which gives actual neutral Japanese diphthongs.

fig 6.12. Neutral Japanese /CV/ sequences.





# 8. Japanese consonants

8.2.0. As we have already said, it is currently thought that Japanese consonants are pronounced like those of English. We have to correct this false information, which is further reinforced by the widespread use of *ch, j, sh, f, ts* in transliteration (as we do ourselves, for the sake of simplicity). In fact, *ch, j, sh* stand for [t͡ɕj; d͡ʒj; ʑj; ɕj], with no absorption of [j] (in spite of possible *mediatic* realizations with no [j]).

Besides, *f* is not [f], but [ɸ]. Finally, there is [ŋ̚], that –as we will see– is an in-tense (‘syllabic’) provelar semi-nasal: /ŋ̚/ realized as [ŋ̚] after /ɛ(ɛ), a(a), σ(σ), ɸ(ɸ)/, but as [ɲ̚], after /i(i)/, in /i(i)ŋ̚/, or before /h(h)j/.

The table of fig 8.0 gives the consonantal articulations of Japanese, which are necessary for an adequate *neutral* pronunciation of this language, including (β), as a possible realization in some loanwords, though not really necessary.

fig 8.0. Table of neutral Japanese consonants (for (β) see fig 8.5 & § 8.2.7).

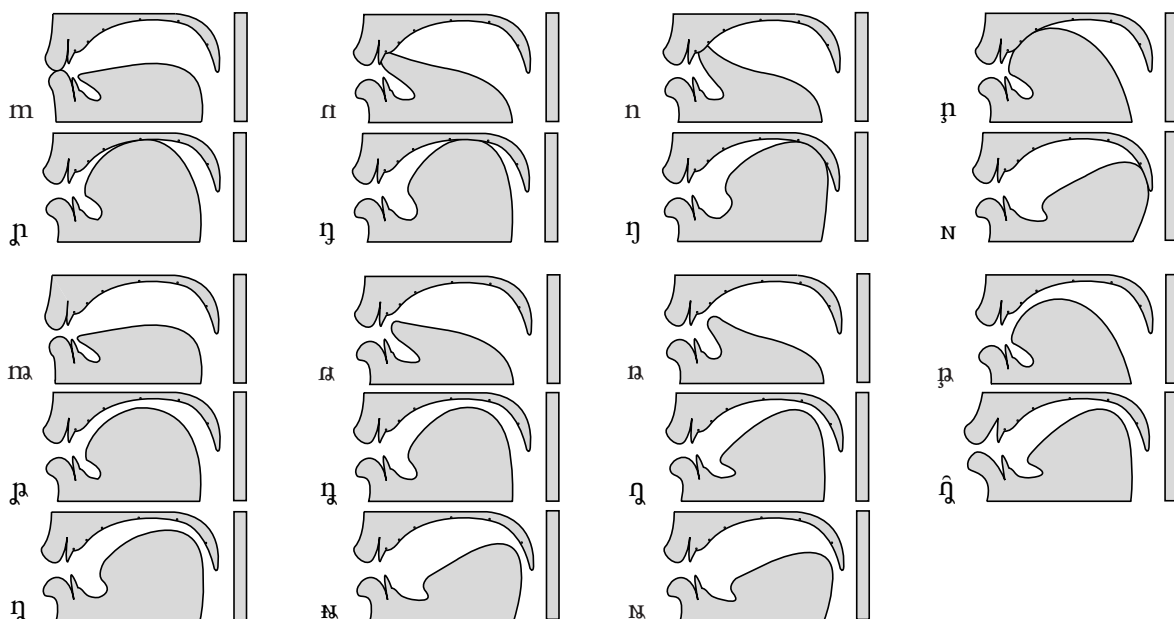
	bilabial	dental	alveolar	prepalatal	bilabialized prepalatal	palatal	postpalatal	prevelar	provelar	bilabialized provelar	velar	laryngeal
N	[m]	[ɸ]	[n]	[ɲ]			[ɲ̚]	[ŋ̚]	ɲ̚	[ŋ̚]	[ŋ̚]	
N	m		n	[ɲ]			[ɲ̚]	[ŋ̚]	ɲ̚	[ŋ̚]	ŋ̚	
K	p b	t d					[c ɟ]	[k ɡ]			k ɡ	[ʔ]
KS		[ts dz]			[t͡ɕ d͡ʒ]							
X												
S		s z			[ɕ ʑ]							
J	[ɸ] (β)					j	[ɰ]			ɸ		h [ɦ]
L			ɾ [ɻ]									

## Nasals

8.2.1.1. At the beginning of syllables, Japanese has three possible nasal phonemes: *m* /m/ [m] (bilabial), *n* /n/ [n] (alveolar; realized, however, as prepalatal, [ɲ], before /i, j/, by assimilation): *nimono* [ɲi-mσ-nσ], and *g̃* [ŋ, ŋ̚, ɲ̚] (tendentially velar): *aḡaku* [ɛ-ŋ̚ɛ.kʰ̚] (less friendly rendered as a plain *g*).

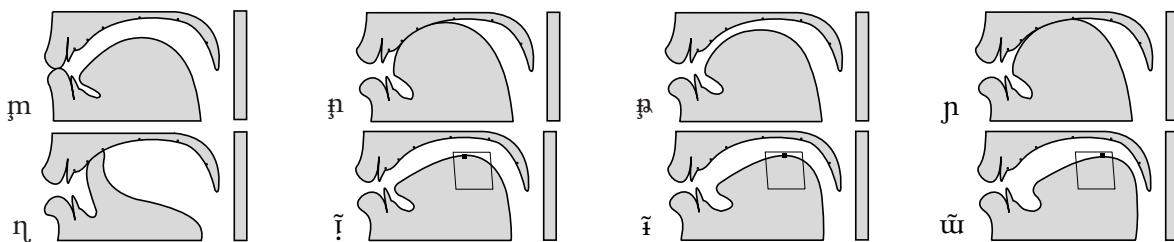
But we must add at once that, within words, *g̃* /ŋ/ [ŋ, ŋ̚, ɲ̚] can be systematic

fig 8.1. Japanese nasals (first 2 rows) & seminasals (last 3 rows), including the much extolled feigned ‘typical Japanese’ uvular ‘[ŋ]’.



only in the most neutral type of pronunciation, after /N, ŋ/. In fact, nowadays the oscillation between [ŋ] and [g] is very widespread, but with much fluctuation between people and words. In *mediatic* and *colloquial* pronunciation, where [g, ɡ, ɰ] predominate, there is also an intermediate stage: the correspondent constrictive velar, prevelar, and postpalatal contoids, [ɣ, ʝ, ɰ], respectively (fig 8.2).

fig 8.2. Japanese *mediatic* variants (and possible *colloquial* vocoidal variants).



However, no neutral native speaker systematically has only [g, ɡ, ɰ]. In (sentence, phrase, lexeme, or grammeme) initial position, [g, ɡ, ɰ] occurs, even for speakers who possess [ŋ, ɲ, ɳ]; for *-gā* we normally find [ŋɐ]; for the conjunction, we have *ga* [gɐ], but *gā* [ŋɐ] is possible too, even after a pause. Certainly, the word *gō* [-gɔ̄] ‘five’ always has [g], whether in its plain form or in compounds.

Again, we find [g, ɡ, ɰ] –above all– in loanwords, in imitative expressions, in Chinese reduplicate words, and after certain proclitics (which are obviously heterorganic): *ḡgā* [-ɰi.ŋɐ.], *ḡn̄gā* [-ɰi.ɲ̄.ŋɐ.]; *mīgi* [.mi-ɲi-], *māguro* [.me-ŋɯ.ɣɔ̄], *mōngèn* [.mɔ̄ŋ̄.gɛ̄ŋ̄.], *māgò* [.me-gɔ̄].

*Summing up systematically*, we have: initial *g* is /g/ [g, ɡ, ɰ], but *ḡ* /ŋ/ mostly in *-gā* (*hàru-gā*), *-ḡurai*, *ḡòtoshì*: *gaikoku*, *gīnkoo*, *guai*, *geta*, *gòmu*. Internal *g* is *ḡ* /ŋ/: *yanāgi*, *ūḡis̄u*, *tsūgeru*.

*Exceptions.* Imitative words: *gàñ-gañ*, *gèra-gera*, *gòro-goro*. Number *gò* ‘five’ and its compounds, like *jùugo*, *gohyakujùu*. Words with honorific *o-*: *o-gyògì*, *o-gèñki*. But, notice that with non-honorific prefixes both versions are mostly accepted: *hi-gòori* and *hi-gòori*, *fu-gòokaku* and *fu-gòokaku*.

We have *g*, at the beginning of the second lexeme in *compounds*, when their link is weak, but not when it is strong: *kootoogàkkoo*, *NihoñGìñkoo*, *madogàrasu*; but: *chuu-gàkkoo*, *shiñgòo*. We also find *g̃* in compounds from original *k*: *kabushiki-g̃aisha*, *ku-chi-g̃èñka*.

Loanwords, mostly from English. Long-standing loans with an original /ŋ/ phoneme have *g̃*. While, more recent loans have /tŋ/ or /g/: *kìñgù*, *bòkushiñgù*, *bìru-dìñgù*, *oru-gañ*, *asuparàgasu*. With geminate *gg* /gg/: *hañdobàggu*.

8.2.1.2. An even bigger (but interesting) phonic problem –regarding nasals– arises from a fourth Japanese nasal phoneme, ie moraic *n̄* /ŋ̄/, which always occurs in a syllable coda and has a prevailing articulation as ‘semi-provelar’ (ie provelar *seminasal*, with no actual contact between the dorsum and the velum). It is phonetically more energetic, ie intense [ŋ̄].

As already said, should somebody want –or have– to avoid writing *n̄*, in translations, and use *n*, they should also use *n'* before *y* (or before a vowel) in the same word, but belonging to a different syllable. Of course, *n̄* is much better –although, obviously– the macron does not hint at a contoid of longer duration. However, by thinking about it as something ‘bulkier’ than a simple nasal contoid, certainly helps in not considering –nor uttering it– as a plain nasal contoid.

In fact, although it is not longer, it surely has an auditorily greater prominence, since it is *intense* (or ‘syllabic’). But, let us immediately state that it is quite wrong to claim that ‘it always has a full syllable beat of its own’. Clearly, a sequence like /Vŋ̄/ does *not* contain *two* syllables, but just *one*, [Vŋ̄] (or its several taxophonic variants, which we will see shortly).

Things are quite different for English *cotton* [ˈkɒtʌn] ([ˈkɒtʌ-n, ˈkɒtʌ-n]), or German *sagen* [ˈzaːgən], where the last segment really *is* a syllable in itself (together with [g] in the German example: [ˈzaː-gən]).

Another still frequently told ‘tale’ is that the Japanese moraic *n̄* /ŋ̄/ is a ‘uvular’ nasal, [N], ie [N̄] (also see fig 8.1). In actual fact, in absolute final position, /ŋ̄/ is normally [ŋ̄]; but in isolated words, said as examples, ‘committed’ speakers may use intense pre-uvular or uvular *seminasals*, [N̄, N̄] (cf fig 8.1). They are also able to exhibit a true uvular intense nasal, [N] (given in brackets in fig 8.1), which they would not ever use in everyday speech.

Clearly enough, in Mandarin Chinese, we actually find /N/ [N]: *kǒngxì* [kɒŋ̄-xì]. It is rather easy to usefully compare it with all the Japanese taxophones (by the way, let us also compare Chinese [ç] with neutral Japanese [ç̄]).

Of course, English speakers (and ‘phoneticians’) feel that it is different from their [ŋ̄] in *taken* [ˈtʰeɪkən]. However, in its basic form, the correct Japanese taxophone is [ŋ̄] (intense provelar *seminasal* contoid): different from both [ŋ̄] and [N̄].

In fact, it is slightly fronter than [ŋ̄], and perceptibly fronter than [N̄]. But,

most of all, it has no firm contact between the tongue and the roof of the mouth, or soft palate, as can be seen in fig 8.1.

The phoneme /ŋ/, realized exactly as [ŋ] (or as postpalatal, [ŋ̟], after /i, ii, ei/, or before /j, hj, hi, hii/) occurs in three positions:

- (1) before continuous consonants (ie those produced with an incomplete occlusion of the oral cavity), that is /s/ [s, ç̣], /z/ [z, ç̣̣], /j/ [j], /ɥ/ [ɥ], /h/ [h, ɦ, φ];
- (2) before a vowel (which is heterosyllabic, of course), and
- (3) in final position before a pause (or, again, before continuous contoids, or vowels).

Some examples: *eñsañ* [·ɛ̃ŋ̣-sɛ̃ŋ̣], *hañsha* [-hɛ̃ŋ̣·ç̣jɛ̃], *uñyu* [-ũŋ̣·jɥ̃], *kañwa* [-kɛ̃ŋ̣·ɥ̃ɛ̃], *seisàñhi* [·sɛ̃ɛ̃-sɛ̃ŋ̣̣·ɦ̣ĩ], *sèñhyaku* [-sɛ̃ŋ̣̣·ɦ̣jɛ̃·kũ̃], *gosèñfu* [·gɔ̃-sɛ̃ŋ̣̣·φũ̃], *tañi* [·tɛ̃ŋ̣-ĩ].

However, by assimilation, /ŋ/ has other taxophones, too, as already said. In fact, it is articulated as a nasal contoid (with full contact, and homorganic to a following consonant, which has a complete oral closure, and /ŋ/ is always intense):

- (1) [ṃ, ṇ, ɲ̣, ŋ̣] (and [ŋ̣, ŋ̣̣]), before the correspondent stops [p, b; t, d] and /k, g/ [k, g] (and [ḳ, g; ç̣, ç̣̣]): *sañpo* [-sɛ̃ŋ̣·pɔ̃], *shiñbuñ* [-ç̣ĩṃ·bũ̃ŋ̣̣], *uñteñ* [-ũŋ̣·tɛ̃ŋ̣̣], *tèñki* [-tɛ̃ŋ̣̣·ç̣ĩ], *tàñka* [-tɛ̃ŋ̣̣·kɛ̃], and

- (2) [ɲ̣, ɲ̣̣] before dental or labialized prepalatal stopstrictive taxophones of /t/ [ts, tç̣] and /z/ [dz, dç̣̣]: *bañzài* [·bɛ̃ŋ̣̣·dzeĩ], *eñchoo* [-ɛ̃ŋ̣̣·tç̣jɔ̃], *niñjiñ* [-ɲ̣ĩŋ̣̣·dç̣jɲ̣̣]; and, naturally, [ŋ̣] also before the frequent *mediatic* stopstrictive (and stop-semi-strictive) variants of *ka* /ka/ [kɛ̃], respectively [kxɛ̃, kɥɛ̃]: *tàñka* [-tɛ̃ŋ̣̣·kxɛ̃, -kɥɛ̃]. In addition, we have

- (3) [ṃ, ṇ, ɲ̣, ŋ̣] (and [ŋ̣, ŋ̣̣]), before the nasals, /m, n, ɲ/ (*m, n, ç̣*), and again [ɲ̣] before *r* /ɲ/ [ɲ̣]: *ùñmei* [-ũŋ̣̣·mɛ̃ɛ̃], *añnài* [·ɛ̃ŋ̣̣·nɛ̃ĩ], *niñniku* [-ɲ̣ĩŋ̣̣·ɲ̣ĩ·kũ̃], *añç̣ai* [-ɛ̃ŋ̣̣·ɲ̣ɛ̃ĩ] (with its secondary variant *añç̣ai* [-ɛ̃ŋ̣̣·ɲ̣ɛ̃ĩ]), and *añraku* [·ɛ̃ŋ̣̣·ɲ̣ɛ̃·kũ̃].

## Stops (& stopstrictives, or ‘affricates’)

8.2.2.1. Japanese has three diphonic pairs for stops: the voiceless are *p, t, k* /p, t, k/ [p, t, k] (as already seen, in *mediatic* pronunciation, *k* /k/ often becomes [kx, kɥ] + /a(a)/; in *yàkuza* pronunciation, *k* /k/ can become [g, ç̣, ç̣̣, ɣ, ç̣̣̣, ɶ]. Occasionally, in *mediatic* pronunciation, *t* /t/ becomes [th] + /a(a), σ(σ)/; and often *p* /p/ is [ph], especially in the context between /ŋ/ and /i(i), a(a)/: *kañpai* [-kɛ̃ŋ̣̣·peĩ] <sup>m</sup>[-kxɛ̃ŋ̣̣·kɥɛ̃ŋ̣̣, -phɛ̃ĩ], *tàko* [-tɛ̃kɔ̃] <sup>m</sup>[-thɛ̃], *kippu* [·ç̣ĩp̣·pɥ̃], *watà* [·ɥ̃ɛ̃-tɛ̃] <sup>m</sup>[-thɛ̃], *kèñpoo* [-kɛ̃ŋ̣̣·pɔ̃]. Word-initial /p/ only occurs in loanwords and onomatopoeic forms. In emphatic speech, there is more ‘aspiration’, which can –then– be heard in *neutral* pronunciation, too, as [ph, th, kh] (fig 8.3).

The greatest ‘oddity’ regards *t* /t/ [t] which, by assimilation, before *i, y* /i, j/ [i, j], is realized as a prepalatal stopstrictive [tç̣] (which is vertically bilabialized, not protruded, as indicated by the diacritic [̣]). See fig 8.4 for some *mediatic* and *foreign* substitutions.

It is transliterated as *ch*, even if /j/ [j] remains, not being absorbed or dropped, as it might be suggested both by its simplified writing, that we ourselves use, and

fig 8.3. Japanese stops and taxophones (including foreign articulations, in brackets).

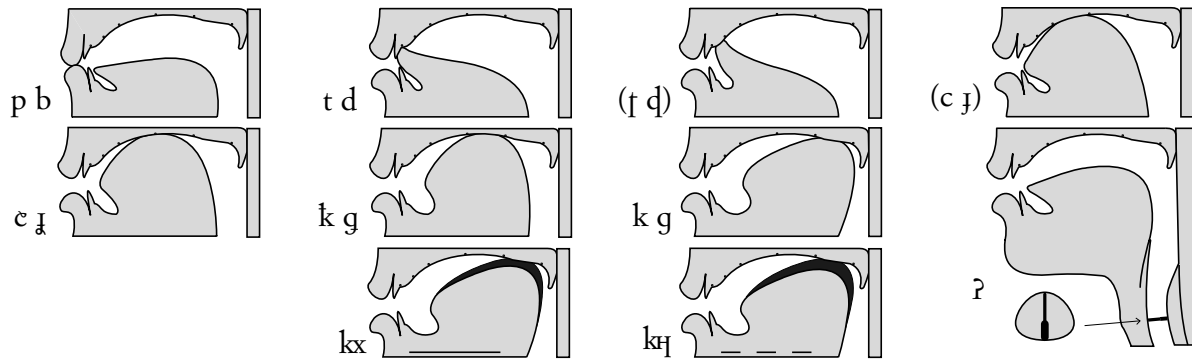
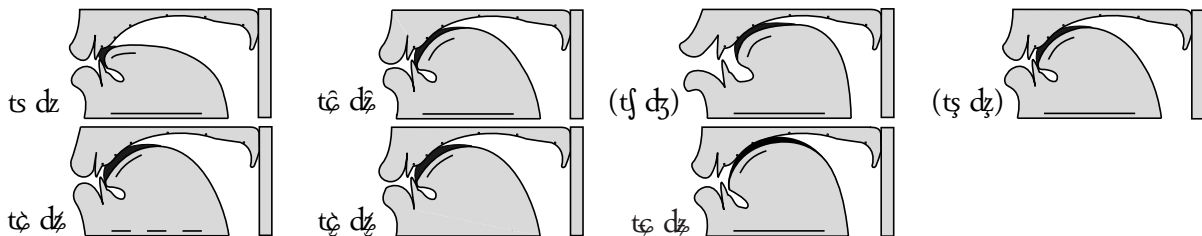


fig 8.4. Japanese stopstrictive realizations of /t/ (including non-neutral or mediatic ones, second row), and main foreign-people substitutes, given in brackets.



the poor phonetic renderings, that we still find in too many books, as ‘[t̚̚]’, or even ‘[t̚̚]’. This last one is not fit at all, being strongly protruded, too – but, see below for the *mediatic* and *colloquial* possibility of /Cj/ → [C̚̚].

However, it is important to note that, before /i, j/, by assimilation, its vertical lip rounding is reduced to something like [t̚̚], but the phone remains different from [t̚̚], which is completely unrounded, or, better, with spread lips).

So, for neutral Japanese, it is much better to use the symbol [t̚̚], also to usefully differentiate it from the Mandarin Chinese one, [t̚̚], which has a (more) normal vertical rounding. A few examples: *chìchì* [t̚̚i-t̚̚i], *chǎnto* [-t̚̚jɛŋ-t̚̚].

In addition –and even more ‘strangely’– *t* /t/ [t], before /ɰ/ [ɰ] is realized as a dental stopstrictive [ts̚̚] (transliterated as *ts*): *tsuǵi* [ts̚̚ɰ-ni].

Let us observe, once and for all, that the so-called ‘palatal’ realizations of /k, g, ŋ/, ie [c, ɟ, ɲ] are rather ‘postpalatal’ (or retracted palatal), and are more satisfactorily represented by means of the symbols shown, instead of [c, ɟ, ɲ], although clearly stating that they are realized in the rear part of their articulatory space, that we define ‘postpalatal’, which in any case remains distinct from the ‘prevelar’ articulation, [k, g, ŋ], typical of most languages before front vowels and [j], as in English *key* [ˈkʰri], *regular* [ˈɹɛɡ-jələ] (Am. Engl. [ˈɹɛɡ-jəlɹ]), *singing* [ˈsɪŋ-ŋ].

Thus, we have: *kippu*, *miǵi*, *tsuǵi* [c̚̚p-p̚̚ɰ-, mi-ni-, ts̚̚ɰ-ni-], provided [ni] remains different from [̚̚ni]. Therefore, /ni/ [̚̚ni] should not be rendered as if it were ‘[̚̚ni]’, but, in *mediatic* pronunciation, we can actually find /ni/ [̚̚ni], too, although oscillating with [̚̚ni]. The same happens with /ki, gi/ [ci, ɟi], and [ci, ɟi], as well.

Let us notice that, in neutral pronunciation, *ky, gy, ǵy* /kj, gj, ŋj/ are [c̚̚j, ɟ̚̚j, ɲ̚̚j], sequences of postpalatal contoids and a palatal approximant (which is partially de-

voiced after voiceless stop phonemes). Thus, we also find: /pj, bj, mj; tj, dj, nj/ [p̚j, bj, mj; t̚j̚, d̚j̚, n̚j̚].

In *mediatic* and *colloquial* pronunciation, we can generally have oscillation between neutral-like realizations and /kj, gj/ [k̚j̚, k̚j̚] [g̚j̚, g̚j̚] (palatal or postpalatal stop-strictives, with the absorption of /j/); and /pj, bj/ [p̚(j), b̚(j), β̚(j), β̚(j), β̚(j)] (and /mj/ [m̚(j)], as well, with palatalized bilabial contoids, followed *or not* by [j, j]), and /tj, dj/ [t̚(j), d̚(j)] (plain vertically rounded prepalatal stopstrictives, again, with or without [j, j]). Of course, also these *mediatic* and *colloquial* oscillating realizations can be responsible for descriptions positing –for Japanese– plain [Ç], with no [j].

8.2.2.2. For *b* /b/, the only neutral realization is [b], while in *mediatic* and *colloquial* pronunciation, more commonly, it is substituted by a voiced bilabial constrictive, [β], which mostly occurs after vowels, especially in non-slow and non-formal pronunciation. Sometimes, also neutral *n̄b* /h̄b/ [n̄b], in *mediatic* and *colloquial* accents are [n̄β]. Less frequently, between vowels, the corresponding semi-constrictive, [β̥], or approximant, [β̥], can be used: *subèru* [s̄uβ̥-b̄ēj̄ū] (*m*[s̄uβ̥-b̄ēj̄ū; -β̄ē; -β̄ē]), *baikai* [-b̄ēīk̄ēī], *jūn̄bi* [-d̄z̄j̄ūm̄bī] (*m*[-d̄z̄ūn̄β̄ī]). See fig 8.5.

The phoneme *d* /d/ [d] poses no problems, except that genuine Japanese words never have *di*, *dyV* and *du*, substituted by *ji* /zi/ [z̄i, d̄z̄i], *jV* /zjV/ [z̄j̄V, d̄z̄j̄V], and *zu* /z̄u/ [z̄u, d̄z̄u]: *dàijij̄n̄* [-d̄ēīz̄īj̄j̄; -d̄z̄īj̄j̄], *juz̄ù* [d̄z̄j̄ū-z̄ū; -d̄z̄ū].

Its normal distribution, non-emphatic and non-slow, has [Vz̄i; |d̄z̄i, n̄d̄z̄i] and [Vz̄j̄V; |d̄z̄j̄V, n̄d̄z̄j̄V] (but, in formal pronunciation, [Vd̄z̄i, Vd̄z̄j̄V] are commoner, as they are in *slow speech*, too, and often when talking to foreigners).

Occasionally, *d* /d/ [d] can become a semiconstrictive, [ð], especially in *colloquial* and *mediatic* pronunciations.

*g* /g/ [g, ɣ, ɣ̥] poses also no problems: *gòḡaku* [-ḡō.ŋ̄ē.k̄ū], or *gōḡaku* [ḡō.ŋ̄ē.k̄ū], *geisha* [-ḡēī.ç̄j̄ē], *Gifu* [ɣ̄ī-φ̄ū], apart from a complementary (or alternative) distribution with *ḡ* /ŋ/ [ŋ, ŋ̄, ɣ̄], which is the most frequent and typical neutral realization, although too often riskily shown as simple *g*. Some examples: *māḡuro* [m̄ē-ŋ̄ū.ŋ̄ō], *kàḡiri* [k̄ē.ŋ̄ī.ŋ̄ī], *òñḡaku* [-ō.ŋ̄ē.k̄ū]. In addition, in *mediatic* pronunciation, it often becomes [ç, ç̄, ɣ̄], either for /g/ or /ŋ/.

We do not find /ŋ/ [ŋ], but /g/ [g], in *gò* [-ḡō] ‘five’, also within words. After a prefix, we have /g/ [g] again: *ooḡiri* [ō.ō.ŋ̄ī]. In onomatopoeic and mimetic words, as well: *gàta-gata* [-ḡē.t̄ē.ḡē.t̄ē], *gù-i-gui* [-ḡū.ī.ḡūi]. The same happens in compound words: *kootooḡàkkoo* [k̄ō.ō.t̄ō.ḡēk̄k̄.ō.ō]; and in loan words: *bàageñ* [-b̄ā.ḡē.ŋ̄], and adapted proper nouns: *Igirisu* [ī.ŋ̄ī.ŋ̄ī.s̄ū].

Normally, also word-initial *g* is /g/ [g]: *gairaīḡo* [ḡēī-ŋ̄ē.ŋ̄ō]. Notice that *-ḡo* [ŋ̄ō] ‘word’ is different from *gò* [-ḡō] ‘five’. In addition, we have enclitic grammemes like: *-ḡa* [-ŋ̄ē], *-ḡo* [-ŋ̄ō], *-ḡoro* [-ŋ̄ō.ŋ̄ō], *-ḡurai* [-ŋ̄ū.ŋ̄ē].

Word-initial vowels, both at the beginning or in the middle of phrases and sentences, are generally preceded by [ʔ] (also shown in fig 8.3), especially for *emphasis*, or to separate vowels of adjoining words.

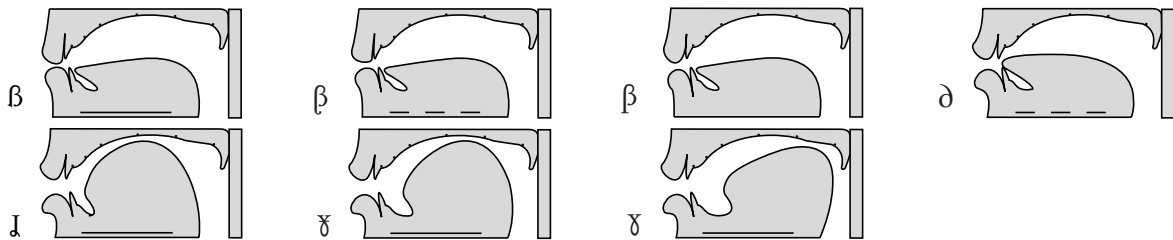
This fact will be indicated prevalingly in connected transcriptions (to avoid people think that current Japanese is quite similar to northern German, with its

predominating [#ʔV] – not only when stressed: [#ʔV]).

In addition, especially (but not only) in *women's* pronunciation, short utterance-final vowels, mostly with a *suspensive* tune, can be followed by [Vʔ(ʔ)ʰ] (ie with or without –an audible– release) as an alternative pronunciation, instead of a possible more ‘normal’ partial devoicing of the last vocoid, more typical of *mediatic* pronunciation: *òke* [-(ʔ)σ.ḱE̞, -(ʔ)σ.ḱE̞(ʔ),], *asù* [(ʔ)ɐ-sũ̞, (ʔ)ɐ-sũ̞(ʔ),].

In the passage in § 11.1.3, a couple of cases are indicated. A common *mediatic* variant is [Vʰ] (with creaky voice), or [Vʔʰ] (as shown in § 6.9).

fig 8.5. Japanese /b, d, g/: non-neutral different realizations. The xenophoneme (β) can be realized as [β, β̥, β̰; v, ʊ, ɸ] ↑[v̥, v̥], commonly [b].



### Constrictives (or ‘fricatives’)

8.2.3. Japanese has a couple of grooved dental constrictives: *s* /s/ [s, ç̣], *z* /z/ [z, ç̣̰]. The voiceless one, /s/, is realized as a bilabialized, but not protruded, prepalatal [ç̣], and /j/ [j] before *i* /i/ is [i], but by coarticulation, it has reduced rounding. Thus, it is transliterated as *sh*, and before *yV* /jV/ [jV], as *shV*. However, /j/ [j] does not disappear, although it may become a semiapproximant, [j̠], in neutral pronunciation (or, actually [∅], in *mediatic* or *colloquial* pronunciation): *shashiñ* [ç̣jɐ-ç̣iɰ̞]. See fig 8.6.

The corresponding voiced phoneme, *z* /z/, is [VzV] (in slow *or* careful pronunciation: [VdzV]) and [dzV, ɳdzV]. Hence, it is realized as a dental constrictive, between vowels, both in words and sentences, either in normal or fast speech. However, after a pause or /ḥ/, it is realized as a stopstrictive: *kaze* [kɛ-zɛ] (slow *or* care-

fig 8.6. Japanese /s/ [s, ç̣], /z/ [z, ç̣̰] (& most frequent unsatisfactory foreign substitutes for [ç̣, ç̣̰]).

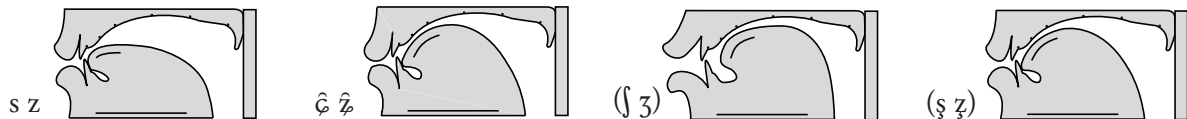
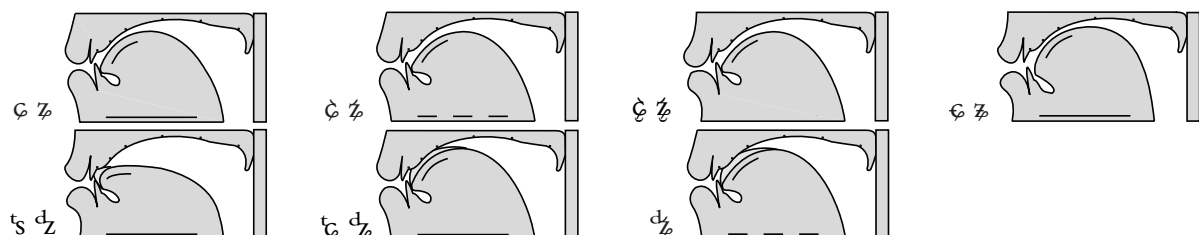


fig 8.7. Japanese not fully neutral realizations of [ç̣, ç̣̰] (and [s]).



ful: [kɛ-dzɛ]), *zuan* [dzɯ-ɛŋ], *sènzo* [-sɛŋdzɔ].

In addition, /z/ [ʒ, dʒ] is transliterated as *j*, before *i* /i/ [i]. The same happens before other vowels in the sequence /zjV/: jV [ʒjV, dʒjV] (although /j/ [j] remains, with the same peculiarities seen for /s/). It is realized as prepalatal rounded (again, with somehow reduced rounding) [Vʒi, VʒjV] (slow <sup>or</sup> careful: [Vdʒi, VdʒjV]) and [dʒi, dʒjV; ɲdʒi, ɲdʒjV]: *Fùji* [-ɸɯʒi], *jiteñ* [dʒi-tɛŋ], *kanji* [-kɛŋdʒi]. Fig 8.7 shows frequent *mediatic* variants of [ʒ, ʒ].

As a quite frequent speech defect, /s/ becomes [ʃ] (a voiceless non-grooved *uni*-lateral constrictive, shown in fig 7.3.1-3, which, in Welsh and Zulu, is not just a ‘strange’ sound, but a full phoneme in itself).

During childhood, instead of this more problematic phone, we can find /s/ articulated as shown in the bottom row of fig 8.4: any of the three phones given there can be used, also alternating between them, [t̪, t̪̥, t̪̥̥]. Usually, this speech defect is spontaneously corrected as soon as children grow up.

## Approximants

8.2.4.1. Japanese has three approximants. The first, *y* /j/ [j], is (voiced) palatal: *sayo(o)nàra* [sɛ-jɔ(ɔ)-nɛ.ɲɛ], *yashiki* [jɛ-ʒi-çi], *yookyuu* [-jɔɔ-çjɯɯ], *shùushi* [-çjɯɯ-çi], *miyako* [mi-ɛ.kɔ].

As can be seen from our examples, it remains unchanged in word-initial position /#jV/ [#jV], whereas it is realized as a phonetic ‘zero’, [∅], when preceded by *i* /i/: /ijV/ → [iV]. In non-slow speech, the same is possible for /ɛjV/ → [ɛV], or –better– [ɛjV]. Thus, we will mark [ɛjV] (by using the palatal semi-approximant symbol): *heyà* [hɛ-jɛ].

As already said, after the voiceless stop phonemes /pjV, tjV, kjV/, /j/ is devoiced, [j̥]: [pj̥, t̪j̥, çj̥]; but it remains [j] after other consonants (even if voiceless, *s*, *h* /s, h/ [ç, ɲ]), and without being absorbed by /t, s, z, h/: [t̪ç, çj, ʒj, dʒj, ɲj] (in spite of current transliterations as *ch*, *sh*, *j*, which, however, can hint at *mediatic* or *colloquial* pronunciations). Sequences such as *yi*, *ye* do not occur.

As fig 8.15 shows, in *mediatic* pronunciation, *ky*, *gy*, *gy* /kj, gj, ɲj/ are postpalatal [kç, gç] stopstrictives or a sequence with /j/ [ç] for /ɲj/ [ɲç] (see further § 8.2.6.1).

8.2.4.2. The second Japanese approximant, *w* /ɥ/ [ɥ] (fig 8.8), which occurs in the syllable *wa*, is voiced vertically rounded provelar, differing from [w], which is velar *and* fully rounded. It has the same relationship with *u* /ɯ/ [ɯ] as happens in English between /w/ [w] and /uu, ɔ, u/ [ɯ, ɔ, u]: *win* [wɪn:], *quick* [khwɪk], *few* [fju:], *too* [tu:] (Am. Eng. [tu:]), *took* [tʰɒk], *to eat* [tʰi:ɪt]. Thus, we have: *watashi* [ɥɛ-tɛ-çi], *kawa* [kɛ-ɥɛ], *uwasa* [ɥ-ɥɛ-sɛ], *deñwa* [-dɛŋ-ɥɛ].

fig 8.8. Two Japanese approximants: /j/ [j, j̥, ∅] (including a ‘zero’ phone for /Vji/), /ɥ/ [ɥ].





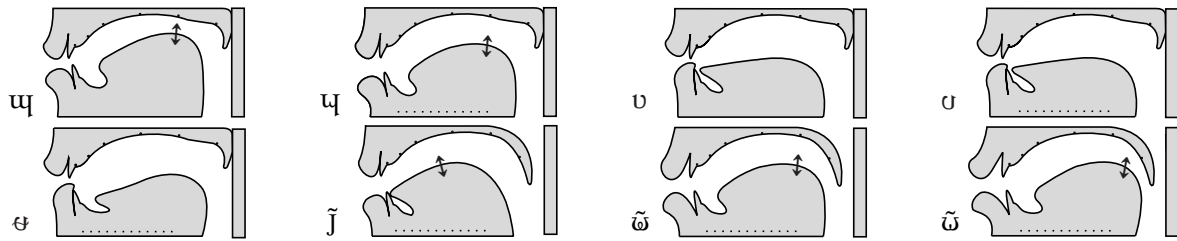
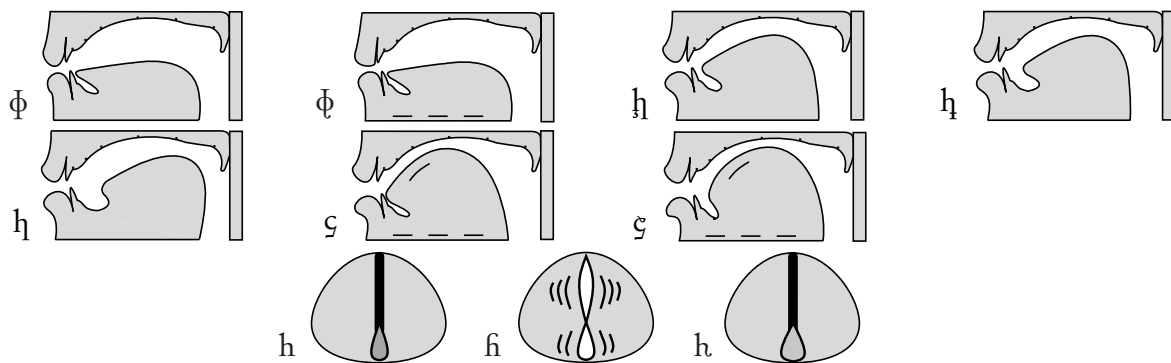
fig 8.9. Japanese /ɥ/: several *mediatic* or *colloquial* variants (including labiodentals).

fig 8.9 also gives three nasalized semi-approximant contours, [ɥ̃, ɥ̃, ɥ̃], which occur in *mediatic* pronunciation, in the following contexts: /ɥ̃<sup>(#)</sup>i, ɥ̃<sup>(#)</sup>e/ [ɥ̃<sup>(#)</sup>ji, ɥ̃<sup>(#)</sup>je], /ɥ̃<sup>(#)</sup>ɥa/ [ɥ̃<sup>(#)</sup>ɥe], /ɥ̃<sup>(#)</sup>o/ [ɥ̃<sup>(#)</sup>ɥo]. In *neutral* pronunciation, we simply have: /ɥ̃/ [ɥ̃] + [h, j, i, #i], /ɥ̃/ [ɥ̃] + [h, ɥ, ɥ, e, e, #ɥ], /ɥ̃/ [ɥ̃] + [o, #o].

A few examples: *teñiñ* [-tɛ̃ɥ̃.iɥ̃] (<sup>m</sup>[-tɛ̃ɥ̃.jiɥ̃]), *Nihòñ-e* [ɥ̃i.hɔ̃ɥ̃.e] (<sup>m</sup>[ɥ̃i.hɔ̃ɥ̃.je]), *deñwa* [-dɛ̃ɥ̃.ɥ̃e] (<sup>m</sup>[-dɛ̃ɥ̃.ɥ̃e]), *Nihòñ-o* [ɥ̃i.hɔ̃ɥ̃.o] (<sup>m</sup>[ɥ̃i.hɔ̃ɥ̃.ɥ̃o]), *sàñu* [-sɛ̃ɥ̃.ɥ̃] (<sup>m</sup>[-sɛ̃ɥ̃.ɥ̃]).

The third approximant, *h* /h/ [h], is voiceless laryngeal. However, in *fast* or *mediatic* pronunciation, [h] can become voiced: [ɦ], after vowels (sometimes also the voiceless semiapproximant [h] may occur). But the most remarkable fact is that, by assimilation, /hɥ/ is [ɥɥ] (a voiceless bilabial approximant, more conveniently transliterated as *fu*), and that in /hi, hjV/ [ɦi, ɦjV] we have a voiceless postpalatal approximant.

In addition, in *mediatic* pronunciation, a voiceless velar approximant, [ɥ], is very frequent for /ha(a):/ *hahòñ* [hɛ̃.hɔ̃ɥ̃] (<sup>m</sup>[hɛ̃.hɔ̃ɥ̃]), *hahèñ* [hɛ̃.hɛ̃ɥ̃] ([<sup>m</sup>hɛ̃.hɛ̃ɥ̃]), *fukòo* [ɥ̃ɥ̃.kɔ̃o], *hiġe* [ɦi.ɥ̃e], *hyakù* [ɦjɛ̃.kɥ̃]. See fig 8.10.

fig 8.10. Japanese /h/ [h, ɦ, ɥ] and *mediatic* variants [h, ɦ, h] (and regional ones [ç, ç]).

### ‘Trills’

8.2.5. Japanese has one phoneme of the ‘trill’ type, which is somehow similar to Spanish *r* /r/ [r], as in *interpretar* [in̩terpreˈtar]. A realization like this could be sufficient for a fairly acceptable and easily comprehensible pronunciation of Japanese, especially if of the *international* kind, all the more so because that phone is indeed one of the possible realizations, also frequent, in *mediatic* and *colloquial* pronunciations.

However, it is better to learn the two most typical articulations given shortly

(which are alveolar, again), and use them instead of [r]. Its more convenient symbol is /ɺ/ [ɺ]. The same phone also occurs in American English, for the diaphoneme /ɾ/, before [ɹ], with lateral contraction of the tongue, eg *better* [ˈbɛɹɹɹ].

In fact, the first and more frequent taxophone of Japanese *r* /ɺ/ is [ɺ], *lateral flap* (or *lateralized flap*, fig 8.11), which occurs after vowels (ie *between* vowels, even within sentences): *kawara* [kɛ-ɸɸɸɸɸ], *eri* [ɛ-ɺi].

fig 8.11. Japanese *r* /ɺ/ [ɺ, ɺ̥, ɺ̥̥] (the two typical phones, and their contexts; and a somehow in-between phone, [ɺ̥̥]).



In *mediatic* pronunciation, a non-lateralized *flap* is possible too, [ɺ̥] (which is the main American phone, for the diaphoneme /ɾ/, in all contexts except with [ɹ]: *Betty* [ˈbɛɹi-i], with no lateral contraction): *kawara* [kɛ-ɸɸɸɸɸ], *eri* [ɛ-ɺ̥i]. In Japanese it is better to use [ɺ], although [ɺ̥] and even [r] are possible, too, as we have already said (though not actually neutral).

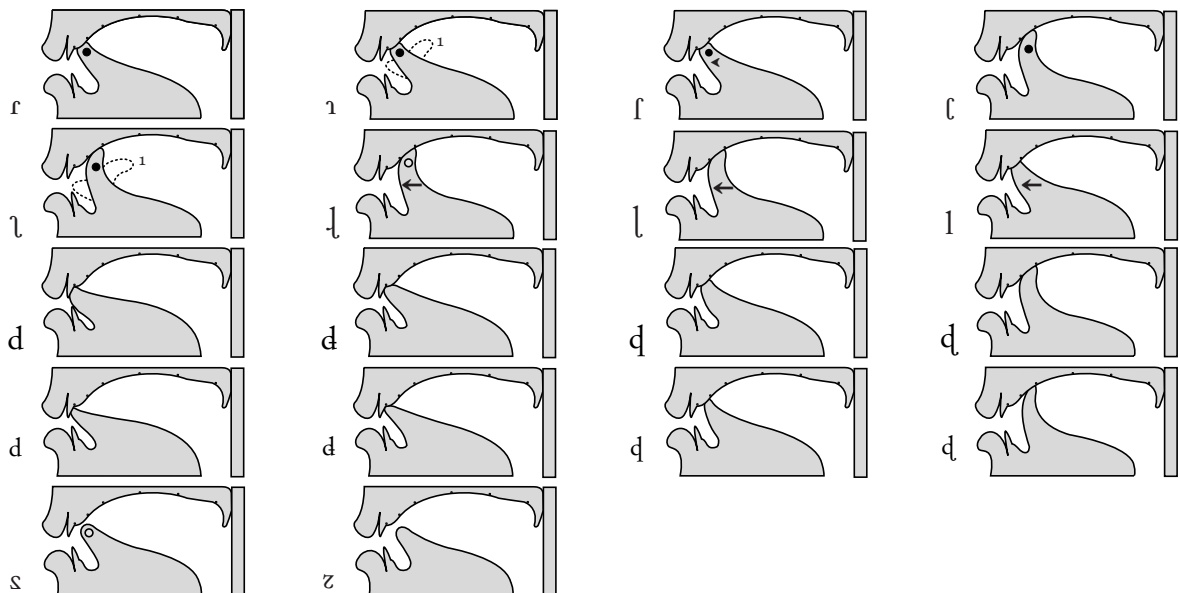
The second important taxophone of *r* /ɺ/ is [ɺ̥̥], *lateral tap* (or *tapped lateral*, fig 8.11), which occurs after /ŋ/ (even within sentences), or after pauses: *bèñri* [ˌbɛŋ̄.li], *ruiji* [ɺ̥̥i-ɺ̥̥i], *rekishi* [ɺ̥̥ɕi-ɕi], *ròñri* [ɺ̥̥ŋ̄.li].

For the sake of simplicity, we could say that the difference between these two realizations consists in different degrees of lateralization. In fact, [ɺ̥] is *less* lateralized, as its lateralization is an added, or secondary, component (let us say 1/3); whereas, for [ɺ̥̥] its lateralization is prevailing, or primary (let us say 2/3).

A compromise phone, [ɺ̥̥̥], a *lateralized tap*, is also shown in fig 8.11. It can be safely used for both [ɺ̥, ɺ̥̥], as it is sufficiently lateralized and short, or quick.

Occasionally, in *mediatic* pronunciation, fully lateral realizations can be heard as

fig 8.12. Japanese /ɺ/: 18 possible partially different *mediatic* or *colloquial* variants.



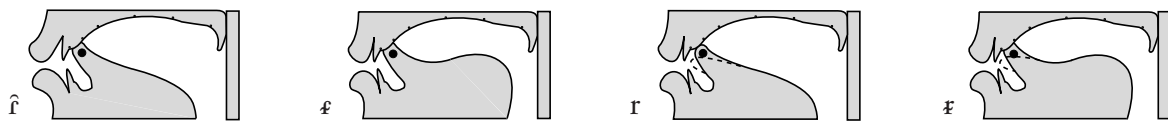
well, [l] (which is still *alveolar*), and also *postalveolar* ones, of various manners of articulation –respectively– *flap*, *tap*, *stop*, *flapped lateral*, *lateral*: [ɭ, ɮ, ɯ, ɺ, ɽ]. See fig 8.12.

Obviously, these phones need not be actively acquired – it is sufficient to be simply able to recognize them, including three more ‘risky’ *mediatic* taxophones: [d̪], a *dental* voiced stop, totally like the *d* /d/ [d] phoneme (!), or [d̠], a *dentalveolar* voiced stop, articulated only in a slightly backer position, and a true *alveolar stop*, [d̥].

A shorter version of these three contoids is also possible, in *mediatic* pronunciation, more frequently in *word-initial* position. We represent them as [d̥̥, d̥̥̥, d̥̥̥̥], and they last about as [ɾ], which can even be no longer than 3 cs (three hundredths of a second). Of course, a fourth possibility is [d̥̥̥̥], as well.

Notice that a typical *yàkuza* pronunciation is well-known mostly for its realizations of /ɽ/, which sound very harsh and rude (cf fig 8.13). They are alveolar, but with some additional features: it can be a *rounded* [ɽ̠], or *velarized* tap [ɽ̠̠], or a trill [ɽ̠̠̠], also with *velarization*, [ɽ̠̠̠̠]. It can also be *lengthened*, with [ɽ̠̠̠̠̠], or [ɽ̠̠̠̠̠̠], in order to sound more ‘frightening’, it seems.

fig 8.13. Japanese /ɽ/ realizations typical of *yàkuza* pronunciation.



Looking carefully at fig 8.14.1-2, it is possible to find some of the most inadequate foreign realizations for *r* /ɽ/: different kinds of uvular phones and others more typical of different English speakers.

fig 8.14.1. Japanese /ɽ/ realizations very different from the many native possibilities.

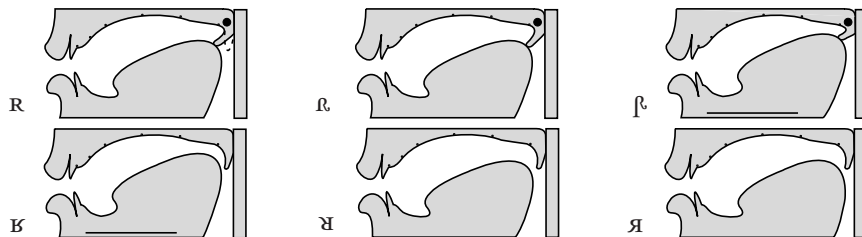
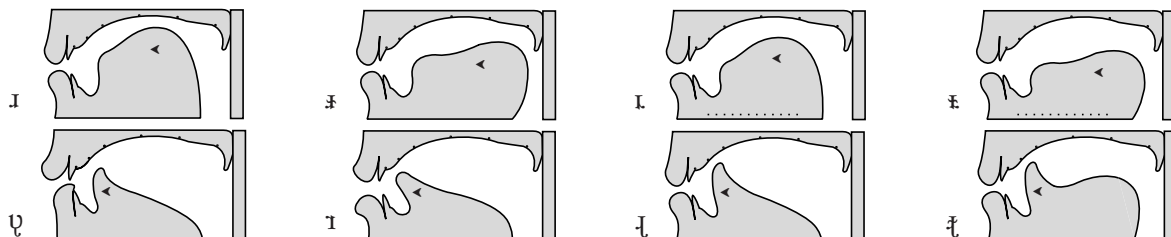


fig 8.14.2. Japanese /ɽ/ realizations typical of different kinds of English speakers.



### ‘Palatalization’

8.2.6.1. Before /i, j/, the phonemes /n, ŋ; t, k, g; s; z; h/ have peculiar but necessary realizations: *ni* /ni/ [ɲi] (pre-palatal), *gi* /ŋi/ [ɲi] (post-palatal); and *chi, ki, gi* /ti, ki, gi/ [tʃi, ci, ɸi]; *shi, ji, hi* /si; zi; hi/ [ʃi; ʒi, |dʒi, ɲdʒi; ɸi].

Some examples: *niñjuu* [-ɲiɲ-dʒjuu], *kağiru* [kæ-ɲiɲu]; *chiè* [tʃi-ɛ], *kieru* [ci-ɛɲu], *giwaku* [ɸi-ɲɛ-kû]; *shimà* [ʃi-mɛ]; *aji* [ɛ-ʒi], *jimi* [dʒi-mi], *màhi* [-mɛɸi].

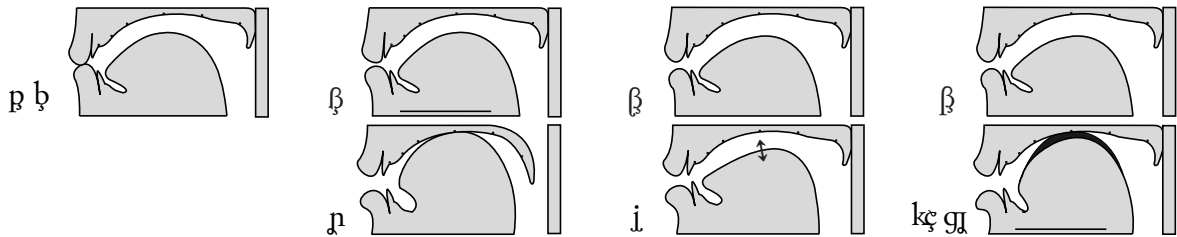
In addition, we find: *ny* /ɲj/ [ɲj], *gy* /ɲj/ [ɲj]; *ch, ky, gy* /tj, kj, gj/ [tʃj, cʃj, ɸj]; *sh, j, hy* /sj; zj; hj/ [ʃj; ʒj, |dʒj, ɲdʒj; ɸj].

Examples: *nyuuğaku* [-ɲjuu-ɲɛ-kû], *añgya* [-ɲjɲ-ɲjɛ]; *chiucho* [-tʃjuuɲtʃjɔ], *kyuu-kyuusha* [cʃjuuɲ-cʃjuuɲcʃjɛ], *gyuunyuu* [-ɲjuuɲ-ɲjuuɲ]; *shiñnyuusei* [ʃiɲ-ɲjuuɲsɛɛ]; *kañyuu* [-kɛɸ-ɲjuuɲ], *jidòsha* [dʒi-dɔɔcʃjɛ], *hyaku* [ɸjɛ-ku].

Again, we must keep in mind that, for /k, g, ŋ/ + /i, j/, the actual articulation is ‘postpalatal’ [c, ɸ, ɲ] (rather than fully palatal, [ç, j, ɲ]); and that [j] remains (in spite of the practical transliteration adopted).

But our phonotonic transcriptions clearly show real neutral pronunciation accurately, although, in *mediatic* and *colloquial* pronunciations, we can certainly find: [kç, ɸç, ɲç], where also [j] is postpalatal.

fig 8.15. Japanese: *mediatic* palatalized and postpalatal phones.



8.2.6.2. All other consonants have no ‘palatalization’ (although certain linguists and phonologists state the contrary, also because they can carry theorism to an excess). So, we regularly have: *ri* /ɹi/ [Vɹi, |li, ɲli]; *mi, pi* /mi, pi/ [mi, pi]; *bi* /bi/ [bi].

Thus: *obieru* [ɔ-bi-ɛɲu], *minori* [mi-nɔɹi], *shìñri* [-ʃiɲli], *risoku* [li-sɔ-kû], *eñpitsu* [-ɛɲp-i-tsû].

In addition, we have /CjV/ [CjV]: *ry* /ɹj/ [ɹj, |lj, ɲlj]; *my, py* /mj, pj/ [mj, pj]; *by* /bj/ [bj]. But, in *mediatic* and *colloquial* pronunciations, we also find /Vbj/ [VBj; Vβj, Vβj]: *eñryo* [-ɛɲljɔ] (*m*[-ɛɲljɔ]), *ryuu* [-ljjuu], *roppyaku* [-ɔp-pjɛ-kû], *byò* [-bjɔɔ], *buñmyaku* [-bûɲ-mjɛ-kû], *zairyò* [dʒei-ɹjɔɔ].

### Loanwords – *gairaiğo* [-ɸei-ɹei-ɲɔ]

8.2.7. As in any language, even in Japanese, loanwords (of which about 10,000 are of English origin) require some adaptation to the syllable structure (which is based on morae in Japanese, as we know) and new phonemic combinations for new sounds, especially for some new consonants.

For these typical adaptations, let us consider three examples: *kùrabu* [kɯ.ɾɛ.bɯ.] ‘club’, *sutoràiki* [sɯ̃:ɾ-ɾɛi.çi.] ‘(workers’) strike’, *sutoràiku* [sɯ̃:ɾ-ɾɛi.kɯ̃.] ‘(baseball) strike’. It is clear how the Japanese syllabic structure changes original monosyllables, with consonant clusters, into actual polysyllables.

Among new combinations, in the traditional taxophonics (which is quite rigid and with a fairly limited number of possibilities), the most common are: *pàatii* [-paa.tii] ‘party’, *dirèkùtaa* [di-ɾɛ.kɯ̃.taa., di-ɾɛ.kɯ̃.taa.] ‘director’, *chèñji* [-tɕ(j)ɛ̃.ɲ.dʒi.] ‘change’, *jètto* [-dʒ(j)ɛ̃.tɕ.] ‘jet’, *shèfu* [-ɕ(j)ɛ̃.ɸ(ɯ̃).] ‘chef’.

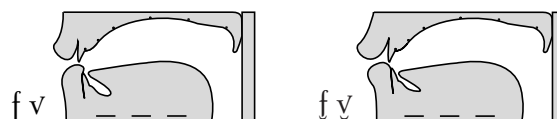
Also (see fig 8.16): *fuirumu*, *fui-* [-ɸɯ-i-ɾɯ-mɯ., -ɾɯ-mɯ., ↑-ɸi-, ↓-ɸi-] ↑↑[-fi-, -fi-] ‘film’, *baioriñ* [.bei-ɕ.ɾi.ɲ., .βei-, .βei-, .βei-] ↑↑[.vɛi-, .vɛi-] ‘violin’, *shìñfonii* [-ɕi.ɲ.ɸɕ-ɲii.] ‘symphony’, *kañtsòone* [kɛ̃.ɲ-tɕɕ.ɲɛ.] ‘canzone’.

Some examples clearly show that, besides placing some phones into new combinations, certain sequences are slightly ‘denipponized’ (as the possible dropping, in these words but not in genuine neutral ones, of [j] after prepalatal –vertically rounded– articulations), becoming slightly more ‘international’.

Certain examples also show, quite clearly, a kind of possible resyllabication, slightly following foreign-language principles.

Here we provide further examples, to complete our small survey: *chèsu* [-tɕ(j)ɛ̃.sɯ̃] ‘chess’, *shèrii* [-ɕ(j)ɛ̃.ɾii] ‘sherry’, *tsàa* [-tsaa] ‘tzar’, *tsetse-bae* [tɕɛ-tɕɛ.bɛɛ] ‘tsetse’, *Èritsiñ* [-ɛ.ɾi.tsi.ɲ] ‘Yeltsin/Ельцин [jɛɭtsɪɲ]’, *tùtti* [-tɯt.ti] ‘tutti’, *tii* [-tii] ‘tea’, *disukù* [-di.sɯ̃.kɯ̃., -di.sɯ̃.kɯ̃.] ‘disc’, *dii-dii-tii* [.dii-dii-tii] ‘DDT’, *Fijii* [-ɸi.dʒii] ‘Fiji’, *fèrii* [-ɸɛ.ɾii] ‘ferry’, *fèaa* [-ɸɛaa] ‘fair’, *fà* [-ɸɛ] ‘F, fah’, *fòto* [-ɸɕ.tɕ] ‘photo’, *fyùujii* [-ɸjɯ̃ɯ̃.dʒii] ‘fusee’, *uìndoo* [ɯ̃-i.ɲ.dɕ] ‘window’, *uisukii* [ɯ̃-isɯ̃.çi] ‘whisky’, *uèsuto* [ɯ̃-ɛ.sɯ̃.tɕ., ɯ̃-ɛ.sɕ.tɕ] ‘waist’, *uòotaa* [ɯ̃-ɕɕ.taa] ‘(cold) water’, *kàrutetto/kuàrutetto* [-kɛ.ɾɯ̃.tɕ.tɕ., kɯ̃-ɛ.ɾɯ̃.tɕ.tɕ.] ‘quartet’.

fig 8.16. Japanese: important ‘committed’ sociophonetic realizations of xenophonemes ‘{ɸ, β}’.



# 9. Japanese structures

## Consonant gemination

9.1.1. We must state that a ‘mora’ coincides with a *light* syllable, as the one formed by a short vowel (/i, e, a, o, u/ [i, e, a, o, u], preceded or not by a single consonant, /C/, or /Cj/), or else by /ŋ/ alone. Also the first element of a geminate consonant (/VC-CV/ [VC:C:V] [[C#C:]] counts as a mora in Japanese, although the second part of a geminate is decidedly longer than its first part, and is followed by a vocoid. But, as far as the length of the consonants is concerned, it is their bisyllabic structure which is relevant: [VC#C:V].

A *half-heavy* syllable corresponds to a geminate vowel (/ii, ee, aa, oo, uu/ [ii, ee, aa, oo, uu]), or a diphthong, or to a short vowel followed by /ŋ/ (/Vŋ/ [Vŋ]). Again, the first element of a geminate consonant (whose second element belongs to the following syllable, together with its vowel, as just seen) counts as another mora added to a geminated or diphthongized vowel. Above, we saw /VC-CV/.

Instead, a *heavy* syllable presents a geminate vowel (or diphthong) + /ŋ/ (/VVŋ/ [VVŋ]) or + the first element of a geminate consonant: /VVC(CV)/ [VVC#(C:V)]. Thus, such heavy syllables may be followed by any other syllable, which can be /CV/ [C:V], as just seen, or a simple /V/ [V], or /ŋ/ [ŋ]. Of course, instead of a simple /V/ [V], we can certainly find /VV/ [VV], as well (either as a geminated or diphthongized vowel). A couple of examples: *tòotta* [-tso:tɔ:tɔ:], *Aiñshutàin* (Einstein) [-eiŋ#çũ:teiŋ#].

As we have already said and seen from various previous examples, in Japanese, *vowel length* is distinctive, ie short and ‘long’ or rather geminate (or doubled) vowels oppose significantly: *so* [-sɔ:], *sò* [-sɔ:] (one mora and one syllable), *soo* [-sɔsɔ:], *sòo* [-sɔsɔ:] (two morae, but one syllable); *tòki* [-tso:çi:], *tòoki* [-tso:çi:].

9.1.2. Even *consonant length* is distinctively present. As just seen, a Japanese geminate consonant counts two morae (in two different syllables). Let us consider: *kite* [çi-te:] (from *kiru* [çi-ŋũ:] ‘to wear’) and *kità* [çi-te:] (from *kuru* [kũ-ŋũ:] ‘to come’), both with two morae and two syllables; *kittè* ‘stamp’ (ie *kittè* [çit-te:] or *kittè* [çit-te:]), and *kittè* (ie *kittè* [çit-te:], or *kittè* [çit-te:], from *kiru* [çi-ŋũ:] ‘to cut’), both with three morae, but two syllables; also *gaka* [gɛ-kɛ:] ‘artist’, *gakka* [gɛk-kɛ:] ‘lesson’.

Let us also note examples like: *Màhha* [-mèh.h.ɛ.] ('Mach [number]'), *sytàffu* ('staff') [s.ŷt̚-tɛ̃.ɸ̃:ɰ̃.].

From a phonic point of view, a doubled consonant always consists of two morae: the first one coincides with the first element of the gemination, even if it is actually shorter, [C]. Whereas the second consonant –which is decidedly longer, since actually lengthened, [C:]– constitutes another mora, together with the (simple) vowel that follows it: [çit-t.ɛ̃.] (in spite of its longer duration).

In fact, *sotto* is [sot̚<sup>#</sup>-t:ɔ̃.] 'softly' (while a similar Italian word *sotto* 'under' is [sot̚<sup>#</sup>to] in a tune, or [sot̚<sup>#</sup>to] in a protune, with non-phonemic different length).

Let us observe well –and listen even more carefully to– the difference between [C<sup>#</sup>C] and [C<sup>#</sup>C:]. Both for Japanese *sotto* and for Italian *sotto*, however, we always have two phono-syllables, even if Japanese *sotto* has three morae.

For the sake of clearness, let us add that, in *colloquial* Japanese pronunciation, instead of a systematic use of [VC<sup>#</sup>C:V], we can often hear simply [VC<sup>#</sup>CV], which –however– still counts as *three* morae, and can actually be something like [VC<sup>#</sup>C:V] or [VC<sup>#</sup>C̣:V].

For the sake of completeness, in *mediatic* Japanese pronunciation, in addition to a systematic plain use of [VC<sup>#</sup>C:V], we can often hear two different alternative structures for voiceless and voiced geminate consonants.

In fact, /VC<sub>Δ</sub><sup>#</sup>C<sub>Δ</sub>V/ can also be realized as [VĊ<sup>#</sup>C:V] (where [Ċ] is a glottalized contoid, cf fig 3.4.B), or [V̤<sup>#</sup>C:V] (where [V̤] is a creaky-voiced vocoid, cf fig 3.4.H), and /VC<sub>◌</sub><sup>#</sup>C<sub>◌</sub>V/, as [VĊ<sub>◌</sub><sup>#</sup>C<sub>◌</sub>:V] (where [Ċ<sub>◌</sub>] is a voiceless contoid, [C<sub>◌</sub>] is a partially voiced/devoiced contoid, [Ċ̣] is a voiced contoid).

9.1.3. A 'moraic' *n̄* is always postvocalic, but it can also be followed by a vowel (and transliterated as *Vn̄V* –or *Vn̄'V*– in order to make people realize we are dealing with /N̄V/ [V̄<sup>#</sup>N̄V], not with *VnV* /N̄V/ [V<sup>#</sup>nV], and the same goes for *Vn̄yV* –or *Vn̄'yV*, if we had not the useful *n̄* sign– /N̄jV/ [V̄<sup>#</sup>j̄V] (by normal and natural assimilation, provided the preceding vowel is not a back one, as in the example *hòñya* [-hɔ̃.ŋ̄.jɛ̃.], given below), which is different from *VnyV* /N̄jV/ [V<sup>#</sup>n̄jV]): *tèndoo* [-tɛ̃.ŋ̄.dɔ̃.ɔ̃.] (cf a similar Italian word, *tendo* /'tɛndo/: [tɛn(ɔ̃)do], ie [tɛn:do] in a tune, but [tɛndo] in a protune).

However, in Japanese, in addition to /N̄V/ [V̄<sup>#</sup>N̄V] and /N̄V/ [V<sup>#</sup>nV], as in *àn̄i* [-ɛ̃.ŋ̄.ī], *àn̄i* [-ɛ̃.ŋ̄.ī], we can also have /N̄j̄V/ [V̄<sup>#</sup>j̄V] (which is the combination of /N̄j̄/ and /j̄V/): *eñgi* [-ɛ̃.ŋ̄.ŋ̄.ī] and /N̄j̄n̄V/ [V̄<sup>#</sup>n̄j̄V] too (combination of /N̄j̄/ + /n̄V/): *añna* [-ɛ̃.ŋ̄.nɛ̃.], *àn̄ni* [-ɛ̃.ŋ̄.ŋ̄.ī] (cf Italian: *Anna* /'anna/ ['an(ɔ̃)na], *anni* /'anni/ ['an(ɔ̃)ni]). Also: *dòñna* [-dɔ̃.ŋ̄.nɛ̃.] (Italian: *donna* /'dɔ̃nna/ ['dɔ̃n(ɔ̃)na]). Let us add this –not useless– example: *hòñya* [-hɔ̃.ŋ̄.jɛ̃.]; in fact, let us notice well that the structure of the first syllable, with a back vowel prevails on the initial palatal consonant of the second syllable.

Therefore, the structure of /j̄/ [m̄, n̄, ŋ̄, ɲ̄, j̄, j̄, j̄] (fully nasal intense contoids) & [m̄, n̄, ŋ̄, ɲ̄, j̄, j̄, j̄] (semi-nasal intense contoids) –one mora– is different from the geminate one /C-CV/ (two morae, or three including the first vowel: /N-C-CV/). However, there is no difference for the counting of morae.

**Japanese accent** – *àkuseñto* [ $\cdot\text{v.k}\hat{\text{u}}\text{.s}\text{e}\text{ŋ}\text{.t}\text{o}$ ,  $\text{.v}\text{k}\text{.s}\text{e}\text{ŋ}\text{.t}\text{o}$ ]

9.2.1. Our transcriptions, as we have seen in the previous section too, indicate that, in Japanese, *accent* is actually a pitch accent, phonemically pertinent, while phonetically, also stress is important and peculiar.

We are not faced with real tones (and tonemes), as in Chinese or Vietnamese (languages where even gliding or compound movements on each syllable are prevailing). On the contrary, in Japanese a pattern stretches over whole words, or whole rhythm groups formed by one or more words and by their (strictly connected) clitic functional syllables (grammatical particles).

The term *àkuseñto* indicates the point, ie the mora, after which the pitch is lowered, passing from mid to low pitch. In our transliterations, this is shown by a grave accent on the pertinent vowel (or mora). Any other morae before the *accent* have mid pitch, except for the very first one, which is low.

If a word or rhythm group has no *accent*, the first mora is low, whereas all the successive ones are mid; thus, without going back to low pitch, according to the pattern that follows (which is limited, here, to four morae – cf fig 9.1).

9.2.2. Only in the following table (which relates to fig 9.1), we will show a *tonemic* (A) and a *tonetic* pattern (B), where [ $\sigma$ ] indicates a low-pitched syllable, with the vowel timbre of / $\sigma$ /; whereas [ $\cdot\sigma$ ] indicates mid pitch. We consider them to be more useful and convenient, in order to describe and learn/teach them. We add pattern (C) that is most recommendable in transliterations, which do not ignore *accent*, when no transcriptions are used.

Let us notice that / $\cdot$ / indicates the presence of *accent* in (A); and that [,] and [-] hint at the pitch height that an added particle has, after a given word or rhythm group: low or mid, respectively.

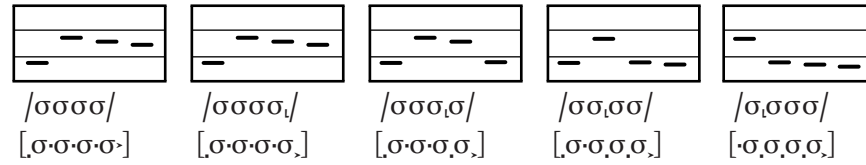
According to a general principle of not explicitly writing unmarked prosodic elements, in our transcriptions, the notation [-] (for mid pitch) could or should be left out. But, it is certainly more useful to show it, all the more so because in actual examples it is much less obtrusive than in the table. In any case, in pattern (H) it has been left out.

For useful comparisons, we add the most widespread patterns used in transliterations (D, E), and the one used in katakana moraic transcription (F) with typically oriental graphic complexities. We also show a phonotonic pattern turned into a more ‘orthodox’ one (G), originating from pattern (E). To indicate any mora, here we use [ $\sigma$ ], / $\sigma$ /,  $\sigma$ ,  $\diamond$  (the last one to ‘symbolize’ katakana writing, in F):

A	/ $\sigma\sigma\sigma\sigma$ /	/ $\sigma\sigma\sigma\sigma$ /	/ $\sigma\sigma\sigma\sigma$ /	/ $\sigma\sigma\sigma\sigma$ /	/ $\sigma\sigma\sigma\sigma$ /
B	[ $\cdot\sigma\cdot\sigma\cdot\sigma\cdot$ ]	[ $\cdot\sigma\cdot\sigma\cdot\sigma\cdot$ ]	[ $\cdot\sigma\cdot\sigma\cdot\sigma\cdot$ ]	[ $\cdot\sigma\cdot\sigma\cdot\sigma\cdot$ ]	[ $\cdot\sigma\cdot\sigma\cdot\sigma\cdot$ ]
C	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$
D	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$
E	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$
F	$\diamond\diamond\diamond\diamond$	$\diamond\diamond\diamond\diamond$	$\diamond\diamond\diamond\diamond$	$\diamond\diamond\diamond\diamond$	$\diamond\diamond\diamond\diamond$
G	[ $\sigma\sigma\sigma\sigma$ ]	[ $\sigma\sigma\sigma\sigma$ ]	[ $\sigma\sigma\sigma\sigma$ ]	[ $\sigma\sigma\sigma\sigma$ ]	[ $\sigma\sigma\sigma\sigma$ ]
H	[ $\sigma\sigma\sigma\sigma$ ]	[ $\sigma\sigma\sigma\sigma$ ]	[ $\sigma\sigma\sigma\sigma$ ]	[ $\sigma\sigma\sigma\sigma$ ]	[ $\sigma\sigma\sigma\sigma$ ]



fig 9.1. Pitch-accent patterns.



9.2.3. In type-A tonetic transcriptions, the first instance ( $/\sigma\sigma\sigma\sigma/$ ) differs from the second ( $/\sigma\sigma\sigma\sigma_i/$ ), because for the latter we also indicate the succeeding lowering (which is hinted at by  $[\cdot]$ , in type-B transcriptions, in opposition to  $[\cdot]$ ). Indeed, it is not really present when no words follow (as we will see shortly). In type-C transliterations, the accent could even be acute ( $\acute{o}$ ), as some authors do, but the grave one ( $\grave{o}$ ) is to be preferred since it can show actual movements better – in fact, the pitch falls (from mid to low).

Type-D and type-E transliterations reflect the first ones in a more abstract way: tonemic and tonetic (A, B); even the katakana transliteration –or ‘transcription’– (F) shows the same characteristic, but in a more abstract way in comparison with real transcriptions (A, B, G, H).

We do not use transcriptions of the type  $/\sigma^1\sigma\sigma\sigma/$   $[\sigma^1\sigma\sigma\sigma]$ , which someone uses though (in the wake of type-D and type-E transliterations), because if syllables or morae were really pronounced on a high pitch, instead of the mid one, the result would not be at all convincing. Let us add that in certain textbooks it is possible to find both types D and E with katakana, and type F with transcriptions.

As far as the indication of pitch variants is concerned, as we have done in § 1.2, for *hirāgānā*  $[\text{ḥi-ṛe.ṛe.ne.}, \text{ḥi-ṛe.ṛe.ne.}, \text{ḥi-ṛe.ṛe.ne.}]$  and *katākāna*  $[\text{kḗ-te.kḗ.ne.}, \text{kḗ-te.kḗ.ne.}]$ , the most important thing is to indicate them (unless there are social or geographical usage differences, which must be explained). Also: *chīkākū*  $[-\text{tḥi.kḗ.kū.}, \text{tḥi-kḗ.kū.}]$  (with stress change) ‘shortly’, but *chīkākū*  $[\text{tḥi-kḗ.kū.}]$  ‘perception’.

9.2.4. Here are some examples taken from Japanese (pitch-)accent dictionaries. They will complete fig 9.1, indicating the possibility of words with six morae (and a particle, to show their pitch movements in actual phrases, even with the particle). Notice, for instance, that *hi-gā* might be placed (and defined) in two different places: either as with its accent on the initial mora of the lexemes,  $/\text{ḥi-gā}/$ , or else on the final one. The same for many other examples. The important thing, here, is to have a general idea of the different possible structures of Japanese (pitch-)accent.

Accentless forms,  $/\text{ḥi-gā}/$ : *hi-gā*, *tori-gā*, *sakura-gā*, *tomodachi-gā*, *tonarimura-gā*, *murasakiiro-gā*

Accent on final mora,  $/\text{ḥi-gā}/$ : *hanā-gā*, *otokō-gā*, *imootō-gā*, *Oshoōgatsū-gā*, *juuichi-gatsū-gā*

Accent on middle prefinal mora (or ‘postinitial’),  $/\text{ḥi-gā}/$ : *kokōro-gā*, *mizuumi-gā*, *watashibūne-gā*, *aiaigāsa-gā*

Accent on middle pre-prefinal mora,  $/\text{ḥi-gā}/$ : *u-gūisu-gā*, *natsuyāsūmi-gā*, *komogōkoro-gā*

Accent on middle post-postinitial mora, /ɲɲɲɲɲɲ(ɲ)/: *onàidoshi-g̃a*, *Shiñkànseñ-g̃a*  
 Accent on middle postinitial mora, /ɲɲɲɲɲɲ(ɲ)/: *omàwarisañ-g̃a*  
 Accent on initial mora, /ɲɲɲɲɲɲ(ɲ)/: *hì-g̃a*, *àme-g̃a*, *inochi-g̃a*, *Fùjisañ-g̃a*, *àkuseñto-g̃a*, *Tookàidoo-g̃a*.

9.2.5. A short –*mono-moraic*– syllable may have two pitch possibilities: *ne* [-nɛ] ‘sound, tone’ (absence of *accent*), *nè* [-nɛ] ‘root’ (presence of *accent*), but tonetically they are both ‘non-low’ (ie said on a mid pitch: [-nɛ]). The same goes for *ki* [-ci] ‘spirit’, *kì* [-ci] ‘tree’ (in isolation: [-ci]), and *ha* [-hɛ] ‘leaf’, *hà* [-hɛ] ‘tooth’ (in isolation: [-hɛ]).

In the case of *two morae*, we can have *sòo* [-sɔɔ] ‘monk’, *soo* [-sɔɔ] ‘villa, inn’, which are monosyllables realized on half-low pitch – but slightly falling or rising, respectively, since they combine mid and low, or low and mid, pitch (cf fig 9.2, and notice that [-], [-] are different from [˘], [˘] –used in other languages– because their movements are less wide than for these last ones).

Also notice that, in *unstressed* bimoraic syllables, we find [˘], as in: *señsèi* [sɛ̃́-ɲ-SEE], *koohii* [kɔ̃́-ɲ-hii]. Notice that [˘] is half-low, and thus different from a true low dot, [˘]. The last one occurs –for instance– in words with *one-mora* first syllable with no *accent* on it, shown in several previous and following examples.

Let us, now, examine the case of a *one-mora* syllable followed by such *particles* as *-g̃a*, *-wa* [ɲɛ, ɲɛ]. These particles are *accentless*, since their pitch depends on what precedes, even if they are obviously [-ɲɛ, -ɲɛ] when pronounced metalinguistically, in isolation. A few examples: *ne-g̃a* [nɛ-ɲɛ] and *nè-g̃a* [-nɛ.ɲɛ], *ki-g̃a* [ci-ɲɛ] and *kì-g̃a* [-ci.ɲɛ], *ha-g̃a* [hɛ-ɲɛ] and *hà-wa* [-hɛ.ɲɛ].

9.2.6. It is fundamental not to believe that Japanese has *two* ‘tonemes’ – a ‘low’ and a ‘mid’ one (too often, misleadingly called ‘high’). As a matter of fact, *accent* is not at all a kind of actual pitch height (as it is not simple stress, either). On the contrary, it is a *pitch fall*. It is a sort of ‘catatonic point’, after which the pitch falls, passing from the mid to the low band, as the examples clearly show.

But, above all, *accent* is either *present* or *absent*. In English, (the position of) *stress* is phonemic (*import* noun, *impórt* verb); whereas it is not so in Japanese. Besides, in English and non-tone languages, pitch depends only on intonation (and paraphonics); whereas, in Japanese, pitch is phonemic and fundamental (although modified by intonation and paraphonics, too, as we will see).

Japanese has mid pitch until an *accent* is inserted, after which the pitch becomes low. If no *accent* occurs, the pitch remains mid (but the first mora of a word or rhythm group is on a low pitch). On the contrary, *stress* in Japanese is not phonemic and depends on a complex interplay of various factors, such as the presence or absence of *accent*, where it is placed, and the syllable structures of their word or rhythm group.

Naturally, the tonograms sufficiently highlight that, in a rhythm group, or in an isolated word, the first mora is low and contrasts with the second one, which

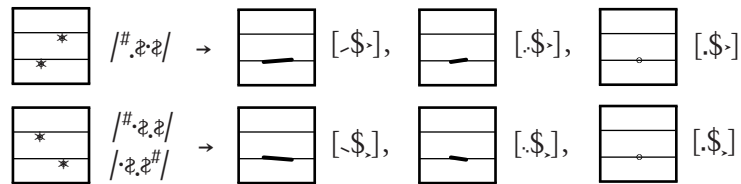
is mid; unless (as we have already said and seen) the first mora itself bears an *accent*, in which case it is mid and all what follows is low.

Again, with two morae, we also have *hàna* [-hɛ.nɛ] ‘edge’, *hanà* [hɛ.nɛ] ‘flower’, *hana* [hɛ.nɛ̃] ‘nose’ (all bisyllabic), as for the monosyllabic examples followed by a particle, as seen above.

As soon as a particle is added, the effect of *accent* is immediately clear: *ne-wa* [nɛ.ɥɛ̃] ‘sound, tone’, *nè-wa* [-nɛ.ɥɛ̃] ‘root’, *hàna-g̃a* [-hɛ.nɛ.ŋɛ̃] ‘edge’, *hanà-g̃a* [hɛ.nɛ.ŋɛ̃] ‘flower’, *hana-g̃a* [hɛ.nɛ̃.ŋɛ̃] ‘nose’; and so on.

Also: *ki-ni kakèru* [ci.ŋi.kɛ̃.kɛ̃ɥ̃] ‘to take a thing to heart’, *kì-ni kakèru* [·ci.ŋi.kɛ̃.kɛ̃ɥ̃] ‘to hang something on the tree’, *kaki-o tabèru* [kɛ̃.ciɔ̃.tɛ̃.bɛ̃ɥ̃] ‘to eat parsimmons’, *kàki-o tabèru* [·kɛ̃.ciɔ̃.tɛ̃.bɛ̃ɥ̃] ‘to eat oysters’.

fig 9.2. Movements in syllables with two morae of different pitch.



### The very nature of Japanese *accent*

9.3.1. In order to ‘naturally’ cope with Japanese *accent*, it is necessary to start from its very nature. As we ‘know’ it is *tonemic*, although not without a stress component, which is also necessary in order to succeed in uttering Japanese words, phrases and sentences.

It must be stated very clearly that we all should do exactly what native speakers and hearers do in a spontaneous way. What matters is the relative height of each syllable (in spite of possible neutral variants, as well).

Of course, all this is a typical Japanese feature, which is not shared even by other tone languages, as Chinese or Vietnamese. It goes without saying that stress-languages are still more different in their way of dealing with prominence. As fig 9.1 shows, each mora has its own pitch, which determines the whole tonetic movement of everything that is said natively.

Also fig 9.2 has to be accurately considered for two-morae syllables of different pitch.

However, the long and short of it is that each Japanese syllable has to be said primarily respecting its pitch. Of course, some stress prominence has to be put on some particular syllable, in order to be able to easily utter the whole sequence.

On the other hand, stress-languages are certainly not without a pitch component. Otherwise they would not sound natural. In addition, all languages (either tone or stress ones) have certain intonation (and paraphonic) structures, which make them sound even more natural.

Thus, the Japanese native-speakers’ brain drives them to pitch each syllable accurately maintaining its intrinsic tone, as shown in fig 9.1-2. Even foreign learners should manage to do something similar, favoring pitch over stress.

9.3.2. Limiting our observations mainly to three- and four-syllable words, it is interesting to note that, in *slow and precise speech*, as in uttering words in isolation, native speakers may exhibit a kind of ‘citation utterance’, which gives more prominence to their third and second last syllables, but the others keep a certain prominence, as well, though less clear.

Looking at fig 9.1, let us try to show these differences by means of [•] for their greater degree, and [◌] for their relatively lesser degree: [◌ ◌ ◌ #], [◌ ◌ ◌ #], [◌ ◌ ◌ #] and [◌ ◌ ◌ ◌ #], [◌ ◌ ◌ ◌ #], [◌ ◌ ◌ ◌ #], [◌ ◌ ◌ ◌ #]. And, any [•] or [◌] can be realized as more prominent, as far as stress is concerned: [-] or [◌].

The choice is rather free, although there are given preferences dependent on particular words and speakers. However, most native speakers do not even realize that there may be this tiny difference, provided the relative pitch on each syllable is respected.

But, at *normal* speed, for words, phrases and sentences (obviously, not declaimed), we find other patterns, which are determined more by rhythm. Thus, we have some weaker syllables alternating with stronger ones: [◌ - ◌], [◌ - ◌], and [- ◌ ◌] or [◌ ◌ ◌], for three syllables; and [◌ - ◌ ◌] or [◌ ◌ ◌], [◌ - ◌ ◌] or [◌ ◌ ◌], [◌ - ◌ ◌] or [◌ ◌ ◌], [◌ ◌ ◌] or [◌ ◌ ◌], respectively, for four syllables.

In actual sentences, the combination of words and phrases avoids having two secondary-stressed syllables in contact. Thus, one [- ◌] or [◌ ◌] is normally changed into [◌ ◌] or [◌ ◌].

9.3.3. The *heaviness* of each syllable depends on the following *prominence scale*, from the lightest to the heaviest one, which gives the following groups:

- 1: [V̥] (fully voiceless), [V̥] (partially voiceless),
- 2: [V̥] (voiced low-pitched [i, ʉ]),
- 3: [V̥] (voiced low-pitched [E, e, σ]),
- 4: [V̥] (‘unaccented’ voiced mid-pitched [i, ʉ]),
- 5: [V̥] (‘unaccented’ voiced mid-pitched [E, e, σ]),
- 6: [V̥] (‘accented’ voiced mid-pitched [i, ʉ]),
- 7: [V̥] (‘accented’ voiced mid-pitched [E, e, σ]),
- 8: [VC̥] (in checked syllables with voiceless consonants),
- 9: [VC̥] (in checked syllables with voiced consonants),
- 10: [VV] (‘long’ vowels, including /aa/ [aa]),
- 11: [VV̥] (true diphthongs),
- 12: [VN̥] (short vowel in checked syllable with an intense nasal),
- 13: [V̥C̥, V̥C̥, V̥V̥, V̥V̥, V̥N̥] (as 8-12 with an *accent* on their first mora),
- 14: [VVC̥] (‘long’ vowel in checked syllable with voiced or voiceless consonant),
- 15: [VVC̥] (diphthong in checked syllable with voiced or voiceless consonant),
- 16: [VVN̥], [VVN̥] (‘long’ vowel or diphthong with an intense nasal),
- 17: [V̥VC̥, V̥VC̥, V̥VN̥, V̥VN̥] (as 14-16 with an *accent* on their first mora).

Let us notice that sets 2-3, 4-5, 6-7 have an important heaviness difference concerning the assignment of stress strength on the syllables of actual words. In fact, /i,

ɥ/, even outside devoicing, are less heavy than /ɛ, a, σ/. Thus, they are more prone to avoid receiving stress prominence when an accent precedes them in trisyllabic words as *dòchira*, *tèkubi*, *màyugē* (with [-..]), than *nìmotsu*, *mègane*, *tèashì*, *Mùraki* (with [·\_]), in spite of cases like *mèushì* (most often with [·\_]), similar to *ruiji* (with [-·]). We will see more on this subject in the following sections.

9.3.4. *Summing up the prosodic structure of Japanese accent.* As we know, Japanese is an agglutinative language, with its strong points and weaknesses. We will present what is already well-known, as in Kawahara (2015), but adapting terms and examples to our own method.

Thus, first of all, we cannot avoid talking about the very many *suffixes*, with all their peculiarities, which are a serious problem not only for foreigners. In fact, suffixes may interact with lexemes in different ways. We find at least seven types of Japanese suffixes.

Let us start with the *weak* ones, like *-nàdo* [-nɛ.dσ] or *-tàra* [-tɛ.ɣɛ], which lose their accent when combined with accented lexemes: *tàbetara* [·tɛ.bɛ.tɛ.ɣɛ] (*tabè* [tɛ-bɛ], also changing accent). But, they maintain their accent when combined with accentless lexemes: *maḡetàra* [mɛ.ɣɛ-tɛ.ɣɛ] (*maḡe* [mɛ.ɣɛ]).

We also find *strong* suffixes, like *-ppòì* [-p.p:σi], which always maintain their accent: *kodomoppòì* [kσ.dσ.mσ.p.p:σi] (*kodomo* [kσ.dσ.mσ]), *netsuppòì* [nɛ.tɥ.p.p:σi] (*netsù* [nɛ-tɥ]), including *Nihoḡppòì* [ni.hσ.p.p:σi] (*Nihoḡ* [ni.hσ]).

9.3.5. In addition, there are three types of *early* suffixes, which tend to shift an accent to accentless lexemes. We have *early weak* ones, like *-shi* [-çi], which puts an accent on the final syllable of accentless lexemes: *Yoshidàshì* [jσ.çi-dɛ.çi] (*Yoshida* [jσ.çi-dɛ]). But accented lexemes maintain their accent: *Mùrakishì* [mɥ.ɣɛ.çi] (*Mùraki* [mɥ.ɣɛ]).

Instead, *early strong* suffixes, like *-kè* [-kɛ], shift an accent to the final syllable of lexemes (either with or without an accent): *Yoshidàke* [jσ.çi-dɛ.kɛ] (*Yoshida* [jσ.çi-dɛ]), *Murakìke* [mɥ.ɣɛ.çi.kɛ] (*Mùraki* [mɥ.ɣɛ]).

Lastly, *early moving* suffixes, like *-monò* [mσ-nσ], put an accent right on the final syllable of accented lexemes: *yomìmono* [jσ-mi.mσ.nσ] (*yòmu* [-jσ.mɥ]). But no accent at all is put on combined accentless lexemes: *norimono* [nσ-ɣi.mσ.nσ] (*noru* [nσ-ɣɥ]).

Another kind of suffixes are the *loose* ones, like *-teki* [tɛ-ci], with no accent of their own, which also change accented lexemes into accentless ones: *roḡriteki* [lσ.ɣi.tɛ.ci] (*roḡri* [-lσ.ɣi]). But note that certain such suffixes are quite peculiar, if not irregular.

For instance, *-no* [-nσ] produces such forms as: *kawa-no* [kɛ-ɥɛ.nσ] (*kawà* [kɛ-ɥɛ] – in accented polysyllables). But: *hà-no* [-hɛ.nσ] (*hà* [-hɛ] – in accented monosyllables), *kòdòro-no* [kσ-kσ.ɣσ.nσ] (*kòdòro* [kσ-kσ.ɣσ] – in non-final accented polysyllables).

In addition, *Nihoḡ-no* [ni-hσ.ɣ.nσ] (*Nihoḡ* [ni-hσ]) – in final accented polysyllables ending in a heavy syllable).

9.3.6. Let us conclude with suffixes, like *-zu* [-dzɯ, -zɯ], which changes accentless forms into accented ones on their first syllable: *Òkamotozu* [·σ·kə\_mσ.tσ.zɯ] (*Okamoto* [·σ·kə-mσ.tσ]).

A few *prefixes*, like *o-* [-σ] and *ma-* [-mɛ], can produce peculiar results. For instance: *o-sùshi* [σ-sɯ.çi] (*sushi* [sɯ-çi], but also *sùshi* [-sɯ.çi]), *o-tègami* [σ-te\_ɲɛ.mi] (*tegami* [te-ɲɛ.mi]), *o-mañjuu* [σ-mɛɲ.dʒjɯɯ] (*mañjuu* [mɛɲ.dʒjɯɯ]), *o-imo* [σ-i-mσ] (*imò* [i-mσ]), *o-sàtsu* [σ-se.tsɯ] (from *satsumaimo* [se.tsɯ-mei-mσ]). And: *mattàira* [mɛt-tei.ɲɛ] (*taira* [-tei.ɲɛ]), *mappìruma* [mɛp-pi.ɲɛ.mɛ] (*hirumà* [ɰi-ɲɛ.mɛ]).

9.3.7. For *compound* words, it is already agreed that the position of accent depends on the length of the second lexeme. So, two types of compounds are considered: those with a *short* second lexeme (ie with one or two syllables, for one or two morae), and those with a *long* second lexeme (ie larger than for the other type).

Thus, the *short* type can either maintain its accent, or put one on the last syllable of the first lexeme (whether accentless or accented on any syllable).

These examples show a maintained accent: *boohañ-bèru* [·bσσ-hɛɲ-bɛ.ɲɯ] (*boohañ* [·bσσ-hɛɲ]), *Perusha-nèko* [pɛ.ɲɯ.çi.ɲɛ-nɛ.kσ] (*Pèrusha* [pɛ.ɲɯ.çi.ɲɛ]).

Here are examples with a final accent on the first lexeme: *minashì-gō* [mi-nɛ-çi.ɲσ] (*minashi* [mi-nɛ-çi]), *mai-gō-inu* [·mɛi.ɲσ\_i.ɲɯ] (*mai-gō* [-mɛi.ɲσ]), *Kana-gawà-shi* [kɛ-nɛ.ɲɛ-ɯɛ.çi] (*Kanà-gawa* [kɛ-nɛ.ɲɛ.ɯɛ]), and also *shijùu-kata* [çi-ʒjɯɯ.kɛ.te] (*shijùu* [çi-ʒjɯɯ]).

9.3.8. For the *long* type, we find an accent on the first syllable of the second lexeme, when it is accentless, or accented on its final syllable: *Minami-Àmerika* [mi-nɛ-miɛ\_mɛ.ɲi.kɛ] (*minami* [mi-nɛ-mi], *Amerika* [ɛ-mɛ.ɲi.kɛ]), *oñna-tòmodachì* [σɲ-nɛ.tσ.mσ-de.tçi] (*oñnà* [-σɲ-nɛ], *tomodachì* [tσ.mσ-de.tçi]), *deka-àtama* [de.kɛ-te.mɛ] (*dèka* [-de.kɛ], *atamà* [ɛ-te.mɛ]).

In other *long*-type compounds, the second lexeme accent is retained: *aka-orèñji* [ɛ.kɛσ-ɲɛɲ.dʒi] (*àka* [-ɛ.kɛ], *orèñji* [σ-ɲɛɲ.dʒi]), *natsu-kudàmono* [nɛ-tsɯ.kɯ-de.mσ.nσ] (*natsù* [nɛ-tsɯ], *kudàmono* [kɯ-de.mσ.nσ]).

But there is great variation when the second lexeme originally has an accent on its second last syllable: *nama-tàmàgō* [nɛ-mɛ-te.mɛ.ɲσ, -te-mɛ.ɲσ] (*nàma* [-nɛ.mɛ], *tàmàgō* [te-ma.ɲσ]), *kamiòmùtsu* [kɛ-mi-σ.mɯ.tsɯ, -mɯ] (*kami* [kɛ-mi], *omùtsu* [σ-mɯ.tsɯ]), *hidari-ùchìwa* [ɰi-de.ɲi.ɲɛ.tçi.ɯɛ, --tçi] (*hidari* [ɰi-de.ɲi], *uchìwa* [ɯ-tçi.ɯɛ]).

Lastly, if the second lexeme is *longer* than four morae (and either accented or accentless), it generally maintains its structure: *Nañkyoku-tañkeñtai* [-nɛɲ.çjσ.kɯ-teɲ-kɛɲ.tei] (*Nañkyoku* [-nɛɲ.çjσ.kɯ], *tañkeñtai* [teɲ-kɛɲ.tei]).

9.3.9. Generally, *verbs* and *adjectives* are either accented or accentless. They can even form minimal pairs (with more or less finely perceptible prosodic differences): *moèru* [mσ-ɛ.ɲɯ] vs *moeru* [mσ-ɛ.ɲɯ], *kìru* [-çi.ɲɯ] vs *kiru* [çi.ɲɯ]; *atsù* [ɛ-tsɯi] vs *atsui* [ɛ-tsɯi], *umài* [ɯ-mɛi] vs *umai* [ɯ-mɛi].

**Japanese stress** – *sutòresu* [ˌsũ̃ːtɔ̃ˌɽ̃ːsũ̃ː, ˌstɔ̃ˌɽ̃ːsũ̃ː]

9.4.1. Although stress is not actually distinctive in Japanese, nevertheless it has an important prosodic function. On the other hand, when acculturated native speakers talk about Japanese *accent*, they surely mean *pitch* accent –*accent*– which is distinctive. However, in an automatic way, even non-acculturated natives –inevitably– use different degrees of stress for the various syllables which form their sentences. They also use different pitch patterns, depending on their own accent: neutral, mediatic, or regional. But, generally, they have no exact idea of what they are doing.

Since stress is not distinctive, it can oscillate and shift in sentences, phrases, and rhythm groups (and even words). This can also depend on communicative, pragmatic, paraphonic, and emotional factors. Sometimes, it can even change according to which monosyllables are added enclitically.

However, we will give some indications about the phenomenon of stress, since we believe it is impossible to continue ignoring it completely; although this is exactly what still happens, also in university teaching, even by native speakers, who never took care to understand it neither for themselves nor for their students. Actually, even *accent* is an unknown object for too many teachers, as well.

9.4.2. Let us, now, begin to put things in order, starting from *monosyllables* (although having up to four morae, and the possibility of *accent* on the second one), by reflecting on the fact that polymoraic words, as the following, are actually monosyllables (in spite of contrary confused indications): *ìi* [ˌiĩː], *àu* [ˌɛ̃ːũː], *bài* [ˌbɛ̃ːĩː], *kòe* [ˌkɔ̃ːɛ̃ː], *kyòo* [ˌk̃̃̃̃ːj̃̃̃̃ː], *bùñ* [ˌbũ̃ːñ̃̃̃̃ː], *oði* [ˌɔ̃ːsĩː], *baai* [ˌbaaĩː], *byooiñ* [ˌbj̃̃̃̃ːsĩːñ̃̃̃̃ː] (this last example has four morae, but not four syllables, rather only one! – cf English *going* [ˌgɔ̃ːɔ̃ːɔ̃ːɔ̃ː]).

In these Japanese examples, a stressed syllable is always half-low, but it is slightly *falling* (since it derives from the combination of mid and low pitch, within the same syllable), except in the last two examples, where it is slightly *rising*, instead (since it derives from the combination of low and mid pitch, tautosyllabically – cf fig 9.2).

Let us see the monosyllabic examples given above, as well. Compare also: *oìi* [ˌɔ̃ːsĩː, ˌɔ̃ːsĩː] ‘throne; hello’, *oði* [ˌɔ̃ːsĩː] ‘much, many’, *ooì* [ˌɔ̃ːsĩː] ‘a cover’.

9.4.3. True ‘problems’ begin with *bisyllables*, though. In fact, there are differences between *ame* [ˌɛ̃ːmɛ̃ː] ‘candy’ & *àme* [ˌɛ̃ːmɛ̃ː] ‘rain’, *neru* [ˌnɛ̃ːɽ̃ː] ‘to sleep’ & *nèru* [ˌnɛ̃ːɽ̃ː] ‘flannel’. The same happens in accentless bisyllables formed with a particle.

*Bisyllabic lexemes of two morae*, ie with two light syllables, are stressed on the second syllable, unless an *accent* is on the first one, which is then stressed (as in the two-mora monosyllables seen above): *hì-gà* [ˌh̃̃̃̃ːĩːgɛ̃ː] ‘fire’, *hì-gà* [ˌh̃̃̃̃ːĩːgɛ̃ː] ‘sun’, *koko* [ˌkɔ̃ːkɔ̃ː], *aji* [ˌɛ̃ː-ʒĩː], *ue* [ˌũː-ɛ̃ː], *iu* [ˌĩː-ũː], *oi* [ˌɔ̃ː-ĩː], and *otò* [ˌɔ̃ː-tɔ̃ː], *ashì* [ˌɛ̃ː-ʃĩː], *murà* [ˌmɽ̃ː-ɽ̃ː], *tsugì* [ˌtsɽ̃ː-ɽ̃ː], *shìo* [ˌʃĩː-ɔ̃ː], *iè* [ˌĩː-ɛ̃ː], *tsumà* [ˌtsɽ̃ː-mɛ̃ː]. But, with an *accent* on the first mora, we have: *dòre* [ˌdɔ̃ː-ɽ̃ː], *àkì* [ˌɛ̃ː-kĩː], &c.

*Three-mora bisyllabic lexemes* are stressed on their heaviest syllables (ie with more morae than others, and reinforced by the presence of an *accent*), although there are

some oscillations that we will indicate. It is important to accurately observe *accent* differences (ie pitch), since sometimes they are the only actual differences (but fig 9.2 must be carefully considered): *omoi* [σ-moi], *omòi* [σ-moi], *kasoo* [kə-sσσ], *kasòo* [kə-sσσ], *irai* [i-ri], *irài* [i-ri], *aoi* [e-si], *aoì* [e-si]; *shiai* [çi-ei], *kinoo* [çi-nσσ], *yotei* [jσ-tei], *hìgài* [hi-ri], *shìmei* [çi-mei], *kùroo* [kɯ-ɾσσ], *bàree* [be-ri], *kìgyoo* [çi-ri].

9.4.4. More: *ikeñ* [i-keñ], *ikeñ* [i-keñ], and *kikài* [çi-kei], *shimeì* [çi-mei], *shìkèn* [çi-keñ], *gòzèñ* [gσ-zèñ], *zùboñ* [dzɯ-bσñ], *koojì* [-kσσ-zi], *kòojì* [-kσσ-zi], *eñgì* [-eñ-ri], *aida* [-ei-de], *hàiru* [-heiri], *Kyòoto* [-çjσσσ], *kaigì* [-kei-ri], *dàiku* [-dei-ku], *bèñri* [-beñ-ri].

With half-heavy syllables, like the first one in [VC<sup>#</sup>C:V]: *àkka* [-ek-kə], *rèssha* [-leçç-je], where the *accent* reinforces the half-heavy structure). And: *miñnà* [-mij-ne], *miñna* [-mij-ne], but: *mittsù* [-mits-tsɯ], *akka* [-ek-kə] (where the geminate voiceless consonant weakens its initial syllable).

Generally, in *four-mora bisyllables*, stress falls on the *first* syllable, unless it is a light one (ie with just one mora), or if there is an *accent* on the second syllable (or if the first is only half-heavy, ie checked by [C<sup>#</sup>C:] and with no *accent*).

Thus: *juubyoo* [-dzjɯɯ-bjσσ], *jùubyoo* [-dzjɯɯ-bjσσ], *hookoo* [-hσσ-kσσ], *jùudo* [-dzjɯɯ-dσσ], *koojoo* [-kσσ-ziσσ], *keñtoo* [-keñ-tσσ], *señtoo* [-señ-tσσ], *sèñtoo* [-señ-tσσ], *kiñeñ* [-çi-ri-keñ], *yuubiñ* [-jɯɯ-bij-ri], *Tookyoo* [-tσσ-çjσσ]. Also with an *accent* on the first half-heavy syllable, thus reinforced: *bòcchañ* [-botçç-je].

But, with a light first syllable, the *second* one is stressed: *ekiiñ* [-e-çi-ri], or with an *accent* on the second syllable: *keñtò* [keñ-tσσ], *señsèi* [señ-sei], *koojò* [kσσ-ziσσ], *taifùu* [tei-φɯɯ], or with the first half-heavy (especially if with voiceless geminated contoids, [C<sup>#</sup>C:], and with no *accent*): *gakkoo* [gək-kσσ], *tokkyuu* [tσç-jɯɯ], *shippai* [çip-pei], [çi-ri]. Let us compare *kañpai* [-keñ-pei], with its more prominent first syllable, but also [-keñ-pei], seen that its difference is very slight.

9.4.5. *Three-mora trisyllables* are stressed on their second syllable (including /VV\$/ words), unless it contains a devoiced vowel, which makes stress *shift* backwards (‘→’), or forwards (‘←’, if an *accent* is there): *kimono* [çi-mσ-nσ], *yubiwa* [jɯ-bi-ɯe], *kochira* [kσ-tçi-ri], *futari* [φû-tɔ-ri], *chikarà* [tçi-ke-ri], *àrashì* [-e-ri-çi], *kaerì* [kə-e-ri], *kaeru* [kə-e-ri], *kaèru* [kə-e-ri], *kàeru* [kə-e-ri], *tàoru* [tə-σ-ri], *uekì* [ɯ-e-çi], *shìkakù* [çi-ke-ku], *kàzoku* [kə-zσ-ku].

It is also possible to have: *dòchira* [-dσ-tçi-ri], *dèguchi* [-de-ri-tçi]; *kèshìkì* [-ke-çi-çi], *ikutsu* [-i-ku-tsù], *pòsuto* [-pσ-sù-tσ]. But: *wafuku* [ɯe-φû-ku], *ashì-tà* [e-çi-tə], *yakusha* [jə-ku-çi] (stressed on their last syllable).

All these can be heard with a stress on their second syllable, but with a tonetically reinforced first syllable, ie *dòchira* [dσ-tçi-ri]. In addition, also the other series shown above can have a stressed fully voiced vowel, ie *wafuku* [ɯe-φû-ku], or also [ɯe-φû-ku].

The important and fundamental thing is that the tonetic patterns are maintained, while the degrees of stress can vary. Native ears are, then, satisfied, in spite of these differences.



*Four-mora trisyllables* are stressed on their first syllable, unless it is a light one, even if there is an *accent* after the first mora: *kaimono* [kɛi·mɔ·nɔ], *Oosaka* [ɔɔ·sɛ·kɛ], *sàabisu* [-saa·bi·sɯ̥], *teñpura* [-tɛ̃·pɯ·ɾɛ], *zèñkoku* [-dzɛ̃·kɔ·kɯ̥], *shoojìkì* [-çjɔɔ·çi·ci].

But, with a light first syllable (whether accented or not), the *second* syllable is currently stressed: *shìrooto* [-çi·ɾɔɔ·tɔ], *butsuzoo* [bɯ·tsɯ·zɔɔ], *sekiniñ* [sɛ·ci·ɲĩ], *àkachañ* [-ɛ·kɛ·tçjɛ̃], *maeuri* [mɛ·ɛɯ·ɾi], *o-iwai* [ɔ·i·ɯɛi], *uekiya* [ɯ·ɛ·ciɛ], *suisèñ* [sɯ·i·sɛ̃], *Nihoñgo* [ɲi·hɔ̃·ŋɔ], *pachiñko* [pɛ·tçĩ·kɔ]. But we have *hatsuoñ* [hɛ·tsɯ·ɔ̃] (and sometimes [hɛ·tsɯ·ɔ̃], too).

Also, with an *accent* after the first morae: *ooàme* [ɔɔ·ɛ·mɛ], *roomàji* [ɾɔɔ·mɛ·çi], *kañgòfu* [kɛ̃·ŋɔ·fɯ̥], *roosòkù* [ɾɔɔ·sɔ·kɯ̥, -sɔ·kɯ̥], *juuḡatsù* [dçjɯɯ·ɲɛ·tsɯ̥], where the *accent* prevails).

In *compounds*, usually we find a seemingly pattern, such as: *bañmeshì* [bɛ̃·mɛ·çi], *gaikoku* [gɛi·kɔ·kɯ̥], *yuudachi* [jɯɯ·dɛ·tçi], *haizara* [hɛi·zɛ·ɾɛ].

9.4.6. *Four-mora quadrisyllables* (independently from the presence or not of an *accent* on any mora) tend to be stressed on the *second* syllable (from the beginning): *kamidama* [kɛ·mi·dɛ·mɛ], *naḡaḡutsu* [nɛ·ɲɛ·ŋɯ̥·tsɯ̥], *norimono* [nɔ·ɾi·mɔ·nɔ], *butaniku* [bɯ·tɛ·ɲi·kɯ̥], *sakanaya* [sɛ·kɛ·nɛ·jɛ], *shiwase* [çi·ɛ·ɯɛ·sɛ], *Amerika* [ɛ·mɛ·ɾi·kɛ], *urikire* [ɯ·ɾi·ci·ɾɛ], *tomodachi* [tɔ·mɔ·dɛ·tçi], *uketsuke* [ɯ·kɛ·tsɯ̥·kɛ], *chikàshitsu* [tçĩ·kɛ·çi·tsɯ̥].

More: *kudàmono* [kɯ·dɛ·mɔ·nɔ], *tebùkuro* [tɛ·bɯ·kɯ̥·ɾɔ], *kanazùchi* [kɛ·nɛ·zɯ̥·tçĩ], *Hiroshima* [çĩ·ɾɔ·çi·mɛ], *hanashitè* [hɛ·nɛ·çĩ·tɛ], *tanoshimi* [tɛ·nɔ·çĩ·mi], *kanemochi* [kɛ·nɛ·mɔ·tçĩ], *inemuri* [i·nɛ·mɯ̥·ɾi], *Kubozono* [kɯ·bɔ·zɔ·nɔ].

Or the *last but one* syllable is stressed (especially with /i, ɯ/ in the second syllable, particularly when it is devoiced): *o-fukuro* [ɔ·fɯ̥·kɯ̥·ɾɔ, ɔ·fɯ̥], *yakusoku* [jɛ·kɯ̥·sɔ·kɯ̥], *kakìkàtà* [kɛ·çĩ·kɛ·tɛ, -kɛ·tɛ], *yotsukado* [jɔ·tsɯ̥·kɛ·dɔ], *toshiyori* [tɔ·çĩ·ɔ·ɾi], *kaminàrì* [kɛ·mi·nɛ·ɾi, -ɾi], *hachiḡatsù* [hɛ·tçĩ·ɲɛ·tsɯ̥], *shìchiḡatsù* [çĩ·tçĩ·ɲɛ·tsɯ̥], *asameshì* [ɛ·sɛ·mɛ·çĩ], *hirumeshì* [çĩ·ɾi·mɛ·çĩ], *hìkidashì* [çĩ·ci·dɛ·çĩ], *kashidashì* [kɛ·çĩ·dɛ·çĩ]; also: *hatamichì* [hɛ·tɛ·mi·tçĩ].

However, in *five-mora quadrisyllables*, a stress generally falls on the *last but one* syllable when it is a heavy one or has an *accent*, or else when the second one contains a devoiced vowel.

9.4.7. Mainly, this also happens with final *-tsu*, *-ri*, or in obvious compounds: *asanèboo* [ɛ·sɛ·nɛ·bɔɔ], *arubàito* [ɛ·ɾi·bɛi·tɔ], *getsuyòobi* [gɛ·tsɯ̥·jɔɔ·bi], *moku-yòobi* [mɔ·kɯ̥·jɔɔ·bi], *birudiñḡu* [bi·ɾi·d̃ĩ·ŋɯ̥], *chokorèto* [tçĩ·kɔ·ɾɛ̃·tɔ].

But, in addition to *jiteñshaya* [dçĩ·tɛ̃·çjɛ·jɛ], *hoñsekìchi* [hɔ̃·sɛ·ci·tçĩ], we also find: *juuniḡatsù* [dçjɯɯ·ɲi·ɲɛ·tsɯ̥], *oñnànoko* [ɔ̃·nɛ·nɔ·kɔ], *saraiḡetsu* [sɛ·ɾɛi·ɲɛ·tsɯ̥].

Or, with *six morae*: *oñḡakùkai* [ɔ̃·ɲɛ·kɯ̥·kɛi], *deñkiḡàisha* [dɛ̃·çĩ·ɲɛi·çjɛ], *gaikokùjìñ* [gɛi·kɔ·kɯ̥·çjɛ̃], we also find: *jiteñshaya* [dçĩ·tɛ̃·çjɛ·jɛ], *hoñsekìchi* [hɔ̃·sɛ·ci·tçĩ]. Let us also consider: *Chuuḡokuḡo* [tçĩ·ɲɛ̃·ŋɔ·kɯ̥·ɾɔ], but: *gaikokuḡo* [gɛi·kɔ·kɯ̥·ɾɔ] (most frequently).

9.4.8. Careful observation of many cases of *quadrisyllables*, however, revealed that the interplay of the different kinds of prominence, often, allows a sort of possible exchange between pitch and stress, so that a structure like [,\$-\$-\$] + [,\$·] (or + [,\$·], or + [,\$·]) can become [,\$-\$-\$] (+ [,\$·], or [,\$·], or [,\$·]), and vice versa, especially in spontaneous sentences.

And this is quite understandable, because, in actual facts, Japanese [-] is generally weaker than the symbol itself may suggest. In fact, we might use [·], which is intermediate between [-] and [·]: ie [-] [·] [·].

On the other hand, also [·] is –so to say– ‘fuller’ than a simple plain [·]; again, it is not simply the weak version of [·] (half-stressed on a mid pitch), but something intermediate between (an unstressed) [·] and [·].

However, the real difference is not merely an intermediate degree of strength, but a kind of fusion of stress and tone: a mid pitch tone with a greater strength than [·]. Let us use [·], which gives the whole syllable an increased prominence, that might even be perceived as being [·]: ie [·] [·] [·].

There is a kind of overlapping of different elements for prominence. In certain languages, even a low pitch, like [ ] [ ] [ ], often sounds as more prominent (ie as somewhat more marked) than [-] [·] [·] (though, generally, less so than [ ] [ ] [ ]).

Thus, it is fairly logical and quite understandable that [·] and [·] can easily be exchanged, producing [,\$·-\$-\$] instead of [,\$-\$-\$], and vice versa: [,\$-\$-\$] instead of [,\$·-\$-\$].

9.4.9. The same can often be true not only of *trisyllables*, but also of *bisyllables* (ie *bañzài!* [bɛŋ.dzɛi, -bɛŋ.dzɛi]), depending on the heaviness of their syllables, generally determined by the presence of heavier structures or, on the contrary, of devoiced vocoids (cf § 9.3.3). All that, again, and even more so, when such words are put into phrases and sentences.

We must keep in mind that the tonetic indications shown –being the relevant and distinctive trait– have to be respected much more than the intensive ones of stress, although the best result is given by a fair interplay between these two elements.

Of course, Japanese native speakers, especially neutral ones, find all this quite natural. On the contrary, foreign speakers, not only those with no tonemes in their language, have a hard time unraveling such ‘mysteries’.

The important thing is that a simple example like *sayonàra* (or its possible variant *sayoonàra*), in addition to a more normal realization like [sɛ-jσ(σ)·nɛ.ɲɛ], [·sɛ·jσ(σ)·nɛ.ɲɛ], can be changed into [sɛ·jσ(σ)-nɛ.ɲɛ], [·sɛ·jσ(σ)·nɛ.ɲɛ].

9.4.10. However, we would not get real Japanese, should we say that as shown below. Notice that, here, we use a conclusive tune. So some tonetic differences with those languages can appear better, also adding [·] for mid-pitch unstressed syllables, and [-], [·] instead of ‘normal’ [ ], [ ] to show mid pitch, on stressed or half-stressed syllables, respectively. Of course, the movements of the tunes partially modify the mid pitch shown for those syllable.

Thus: [sɛə'ɲɛ.ɲɛ, -ɲɛ., sɛ·jɔ'] (English) – or [za·jɔ'na·ɲɛ.] (German), [·sɛ·jɔ-na-

·ra., ·sa·jo·na·] (Spanish), [·sa·jo·na·ra·] (Italian), [·sɜ·ju·na·rɛ., ·sa·jõ·na·] (Portuguese), [·sɑ·jõ·nɑ·βɑ., ·sɑ·jõ·nɑ·βɑ.] (French), [·sɤ·jɐ·nɑ·rɤ.] (Russian), [·sɛ·jɔʊ·nɑ·ʔɑ., ·sɛ·ju·nɑ·] (Arabic), [·sɜ·joo·nɜ·rɜ., ·sɜ·jo·nɐ·] (Hindi), [·sɑ̃·jɔʊ·nɑ̃·lɑ̃.] (Chinese).

Let us note that, in French, [·] indicates a half-high pitched unstressed pretonic syllable. As can be seen, it is different from [·], which we introduced above for Japanese. In fact, [·] is a little higher, but a little weaker than [·], too.

Or also, using the very common Japanese family name *Tanaka* [tɛ·nɛ·kɛ], we have: [·hə·nɑ·kɑ., ·kɛ.] (English) – or [·tʰɑ·nɑ·kɑ.] (German), [·ta·nɑ·kɑ.] (Spanish), [·ta·nɑ·kɑ.] (Italian), [·tɜ·nɑ·kɛ., ·tɛ·nɑ·] (Portuguese), [·tɑ·nɑ·ɕɑ.] (French), [·tɛ·nɑ·kɤ.] (Russian), [·tɑ·nɑ·kɑ., ·tɑ·nɑ·] (Arabic), [·tɜ·nɛ·kɜ.] (Hindi), [·tʰɑ̃·nɑ̃·kʰɑ̃.] (Chinese).

Respectively, something like: *Smith*, *Müller*, *Sánchez*, *Rossi*, *Silva*, *Dupont*, *Ivanov* {Иванов}, *Nasr* {Nasr}, *Sharma* {Sharmaa}, *Chang* {Zhāng}.

9.4.11. In *five-syllable words*, which are generally not simple words, stress decidedly tends to fall on the *last but one* syllable, except for particular compounding or vowel devoicing, as we will see in a while.

Thus, we find, with *five* morae: *katazukèru* [kɛ·tɛ·zɯ·kɛ·ɲɯ.], *inabikari* [i·nɛ·bi·kɛ·ɲi.], *otokònoko* [ɔ·tɔ·kɔ·nɔ·kɔ.], *nejimàwashì* [nɛ·ʒi·mɛ·ɯɛ·çì.], *yamanòbori* [jɛ·mɛ·nɔ·bɔ·ɲi.], *okurimono* [ɔ·kɯ·ɲi·mɔ·nɔ.], *o-kyaku-sàmà* [ɔ·çjɛ·kɯ·sɛ·mɛ., ·mɛ.].

With *six* morae: *Amerikàjìñ* [ɛ·mɛ·ɲi·kɛ·ʒì·ɲi.], *señtakumono* [sɛ·n·tɛ·kɯ·mɔ·nɔ.], *ichinichijuu* [i·tçi·ɲi·tçi·ʒjɯɯ.]; but *omàwari-sañ* [ɔ·mɛ·ɯɛ·ɲi·sɛ·ñ.], *machiàishìtsu* [mɛ·tçi·ɛi·çì·tsɯ.].

With *seven* morae: *o-tètsumadai-sañ* [ɔ·tɛ·tɯ·dɛi·sɛ·ñ., ·ɔ·tɛ·tɯ.], *deñkìsutòobu* [dɛ·ñ·çì·sɯ·tɔ·bɯ.], *Chuuḡokuryòori* [tçi·jɯɯ·ɲɔ·kɯ·ɲjɔ·ɲi.], *dañboosètsubi* [dɛ·m·bɔ·sɛ·tɯ·bi., ·dɛ·m·bɔ·sɛ.], *nyuuḡakushikèn* [ɲjɯɯ·ɲɛ·kɯ·çì·kɛ·ñ.].

With *eight* morae: *koodañjùutaku* [kɔ·sɔ·dɛ·ñ·dɛ·jɯɯ·tɛ·kɯ.]. Of course, even in such cases, there is a real possibility for sequences, mostly including [-] [·], [·] [-] [·] [·], [·] [·] and [·] [·] [·], to exchange their order, using [·] [·], [·] [·] and [·] [·] [·].

9.4.12. *Verbs* in *-ru* generally have their stress on the preceding syllable, which has the *accent*: *oboèru* [ɔ·bɔ·ɛ·ɲɯ.], *kañḡaèru* [kɛ·ñḡ·ɲɛ·ɛ·ɲɯ.], *atsumàru* [ɛ·tɯ·mɛ·ɲɯ.], *shirabèru* [çì·ɲɛ·bɛ·ɲɯ.], *koshìkakèru* [kɔ·çì·kɛ·kɛ·ɲɯ.].

But, other *-Cu* verbs, again with *accent* on the last but one mora/syllable, have their stress on the second syllable: *arawàsru* [ɛ·ɲɛ·ɯɛ·sɯ.], *chìkazùku* [tçi·kɛ·zɯ·kɯ.].

Finally, here are some *accentless* forms: *sawaru* [sɛ·ɯɛ·ɲɯ.], *tsukaru* [tɯ·kɛ·ɲɯ.], *ukèru* [ɯ·kɛ·ɲɯ.].

Usually, *adjectives* are stressed on the last syllable, with  $\hat{V}V^{\#}$ , as in *-ài*; others on the last but one, with  $VV^{\#}$ , as in *-ai*. So, we have: *atatakài* [ɛ·tɛ·tɛ·kɛi.], *yakamashii* [jɛ·kɛ·mɛ·çii.], and: *abunai* [ɛ·bɯ·nɛi.], *oishii* [-ɔi·çii.], *kiroi* [-cii·ɲoi.], *akarui* [ɛ·kɛ·ɲui.], *tsumetai* [tɯ·mɛ·tɛi.], *mushiatsui* [mɯ·çie·tsui.].

But there are even cases like: *mushiatsui* [mɯ·çie·tsui.], *muzukashii* [mɯ·zɯ·kɛ·çii.], *atarashii* [ɛ·tɛ·ɲɛ·çii.], *omoshiròi* [ɔ·mɔ·çii·ɲoi.], *meñdookusài* [mɛ·n·dɔ·sɯ·sei.].

In *iterated forms*, stress and accent belong to the first syllable: *mùzu-muzu* [-mɯ·zɯ·mɯ·zɯ.], *wàza-waza* [-ɯɛ·zɛ·ɯɛ·zɛ.].

9.4.13. Since there are no real and simple rules for *accent* and *stress* in *compounds*, especially for less long ones, we just give some examples, to rouse reflection. We are preparing the *JPD<sup>t</sup>* (*Japanese Pronouncing Dictionary [transliterated]*), with an introduction on how to pass to *canIPA* transcriptions, as in this book, for ‘accent’ (*àkuseñto*), ‘stress’ (*sutòresu*), segments, especially nasal  $\bar{g}$ , vowel devoicing, &c.

Examples: *fuyugeshiki* [·fɯ·jɯ·ŋɛ̄·ç̄i·ç̄i], from *fuyù* [·fɯ·jɯ·] and *keshiki* [·kɛ̄·ç̄i·ç̄i]; *yunyukudàmono* [·jɯ·ŋjɯ·kɯ·dɛ̄·mɔ̄·nɔ̄], from *yunyuu* [·jɯ·ŋjɯ·] and *kudàmono* [·kɯ·dɛ̄·mɔ̄·nɔ̄]; *yuyakezòra* [·jɯ·jɛ̄·kɛ̄·zɔ̄·ɽ], from *yuyake* [·jɯ·jɛ̄·kɛ̄·] and *sòra* [·sɔ̄·ɽ]; *satoğòkoro* [·sɛ̄·tɔ̄·ŋɔ̄·kɔ̄·ɽ], from *sato* [·sɛ̄·tɔ̄·] and *kòkòro* [·kɔ̄·kɔ̄·ɽ].

Others: *niğùruma* [·ni·ŋɯ·ɽ·mɛ̄·], from *ni* [·ni·] and *kuruma* [·kɯ·ɽ·mɛ̄·]; *àsàkaze* [·ɛ̄·sɛ̄·kɛ̄·zɛ̄·], from *àsa* [·ɛ̄·sɛ̄·] and *kaze* [·kɛ̄·zɛ̄·]; *inșutañtokòohii* [·i̇·sɯ̇·tɛ̇·tɔ̇·kɔ̇·ç̄i̇], from *inșutañto* [·i̇·sɯ̇·tɛ̇·tɔ̇·] and *koohii* [·kɔ̇·ç̄i̇]; *gaikokùjiñ* [·gɛi·kɔ̇·ç̄i̇], from *gaikoku* [·gɛi·kɔ̇·kɯ̇·] and *jiñ* [·ç̄i̇]; *gaikokujin̄tòoroku* [·gɛi·kɔ̇·kɯ̇·ç̄i̇·tɔ̇·kɯ̇·], from *gaikokùjiñ* [·gɛi·kɔ̇·kɯ̇·ç̄i̇] and *tooroku* [·tɔ̇·kɯ̇·kɯ̇·].

9.4.14. *Summing up the prosodic structure of Japanese stress.* Unfortunately, native Japanese phoneticians simply talk (and write) about *pitch-accent*, as if *stress-accent* were not an important part in the pronunciation of Japanese. But, it is impossible not to deal with stress, because it would not be serious to do without it.

Note that even syllables with *i*, *u* [i̇ ù̇] are given as present, because they may also be uttered as [i̇, ù̇], or [i, u] (especially in *careful*, or *slow*, or *regional* speech). In fact, even when these two vowels are rendered exactly as [i̇, ù̇], their syllabic beat is present, although, in fast speech, resyllabification can certainly occur, with the result that two distinct syllables actually become just *one* phonetic syllable.

9.4.15. In *bisyllables*, the *unmarked* stress pattern is *final* [·\$·\$·#] – note that in these formulae we indicate *intensity*, independently from *tonality*, which depends on the *accent* of each word. But, if the first syllable is ‘heavier’ (according to the scale given in § 9.3.3), the pattern is *marked*, and *prefinal* [·\$·\$·#].

Examples: *hashi* [·hɛ̄·ç̄i·], *hashì* [·hɛ̄·ç̄i·], *hàshi* [·hɛ̄·ç̄i·]; *esa* [·ɛ̄·sɛ̄·], *otò* [·tɔ̄·tɔ̄·], *kaki* [·kɛ̄·ç̄i·], *kità* [·ç̄i·tɔ̄·], *kesu* [·kɛ̄·sɯ̇·], *keğà* [·kɛ̄·ŋɛ̄·]; but: *kèsa* [·kɛ̄·sɛ̄·], *kùmo* [·kɯ̇·mɔ̄·], *ido* [·i̇·dɔ̄·], *nèko* [·nɛ̄·kɔ̄·].

9.4.16. In *trisyllables*, the *unmarked* stress pattern is *prefinal* [·\$·\$·\$·#]. But, if the other syllables are ‘heavier’ (§ 9.3.3), the pattern is *marked*, and either *final* [·\$·\$·\$·#], or *initial* [·\$·\$·\$·#].

Examples (*prefinal* [·\$·\$·\$·#]): *musume* [·mɯ̇·sɯ̇·mɛ̄·], *tatami* [·tɛ̄·tɛ̄·mi·], *kusuri* [·kɯ̇·sɯ̇·ɽ·], *ushiro* [·ɯ̇·ç̄i·ɽ·], *kodomo* [·kɔ̄·dɔ̄·mɔ̄·], *atama* [·ɛ̄·tɛ̄·mɛ̄·], *shikaku* [·ç̄i·kɛ̄·kɯ̇·], *arùku* [·ɛ̄·ɽ·kɯ̇·], *mègane* [·mɛ̄·ŋɛ̄·nɛ̄·], *nimotsu* [·ni·mɔ̄·tɽ·], *depàato* [·dɛ̄·pɔ̄·tɔ̄·].

But (*final* [·\$·\$·\$·#]): *zashiki* [·dzɛ̄·ç̄i·ç̄i·], *rekishi* [·ɛ̄·ç̄i·ç̄i·], *ashita* [·ɛ̄·ç̄i·tɛ̄·], *wafuku* [·ɯ̇·fɯ̇·kɯ̇·]; or (*initial* [·\$·\$·\$·#]): *tékubi* [·tɛ̄·kɯ̇·bi·], *dòchira* [·dɔ̄·tç̄i·ɽ·], *zènkoku* [·dzɛ̄·ŋ·kɔ̄·kɯ̇·].

9.4.17. In *quadrisyllables*, there are two *unmarked* stress patterns: *postinitial* ([ $\$ \$ \$ \$^\#$ ]) and *prefinal* ([ $\$ \$ \$ \$^\#$ ]).

Examples of [ $\$ \$ \$ \$^\#$ ]: *yorokòbi* [jɔ-ɾɔ-kɔ.bi, ·bi, ··bi], *detarame* [de-te-ɾe-me], *wakamono* [ɥe-ke-mɔ-nɔ], *nomìmono* [nɔ-mi-mɔ-nɔ], *gomìbako* [gɔ-mi-be.kɔ], *sakanaya* [se-ke-ne-je], *Amerika* [e-me-ɾi-ke], *kaḡàkusha* [ke-ɾe.kû.ĉje], *ehàḡakì* [e-he.ɾe.ĉi], *o-yaḡo-sañ* [ɔ-je-ɾɔ-seŋ], *o-imooto-sañ* [ɔi-mɔɔ-to-seŋ], *otooto-sañ* [ɔ-toɔ-to-seŋ].

Examples of the other *unmarked* pattern, *prefinal* [ $\$ \$ \$ \$^\#$ ]: *kashidashì* [ke-ĉi-de-ĉi], *yakusoku* [je.kû-so.kû], *yotsukado* [jɔ-tsû-ke-do], *kakìkàta* [ke-ĉi-ke-te], *myùujikaru* [mjɥu.dĉi.keɾu], *kaminari* [ke-mi-ne-ɾi], *oñnanoko* [ɔ-ɾ-ne-nɔ.kɔ], *shokudòsha* [ĉjɔ.kɥ-do.ĉje], *señtakuya* [seŋ-te.kɥ-je].

Others: *saraiḡetsu* [se-ɾe-ɾe-tsû], *toshokàñiñ* [to-ĉjɔ-ke.ĉi.ĉi], *kootoogàkkoo* [kɔɔ-toɔ-ge.k.kɔɔ], *deñwabàñḡoo* [de.ĉi-ɥe-be.ɾe.ɾe], *kañkoobàsù* [ke.ĉi-kɔ-be-sû], *koohijàwan* [kɔɔ-ĉi-dĉi-je.ɥe.ĉi], *kootsuukòsha* [kɔɔ-tɥu.kɔ.ĉje], *Nihoñ-niñgyoo* [ni-hɔ-ɾi.ĉi.jɔɔ].

Also: *birudiñḡu* [bi-ɾu-di.ɾu], *hañdobàḡgu* [he-ɾ-do-be.g.gu], *chokorèto* [ĉjɔ-kɔ-ɾe-to], *taipuràitaa* [tei-pu-ɾe.itaa].

These two structures can easily be exchanged, especially in non-derived words, producing possible [ $\$ \$ \$ \$^\#$ ] → [ $\$ \$ \$ \$^\#$ ], and [ $\$ \$ \$ \$^\#$ ] → [ $\$ \$ \$ \$^\#$ ]. But, let us notice that we have a *marked, initial* pattern ([ $\$ \$ \$ \$^\#$ ]), in words like: *wàza-waza* [-ɥe.ze.ɥe.ze], *gòmi-gomi* [-gɔ.mi.gɔ.mi].

9.4.18. Let us add some examples with five syllables (more or less clearly compounds): *Chuuḡokuryòori* [ĉjɥu-ɾe-ɾe.kɥ-ɾi.jɔ.ɾi], *deñkisutòobu* [de.ĉi-ci-sû-to.ɔbu], *hakaserònbuñ* [he-ke-se-ɾe-mi-bu.ĉi], *o-tètsudai-sañ* [ɔ-te.tsu-dei-seŋ]; *dañboosètsubi* [de-mi-bɔɔ-se.tsu.bi], *otokònoko* [ɔ-to-kɔ-nɔ.kɔ]; *ookèsutora* [ɔɔ-ke.sû-to.ɾe].

## Japanese sentences

9.5.1. In forming Japanese *sentences*, some modifications are introduced regarding *accent* (and somehow for rhythm-group stress, too). Let us consider the following examples, in order to see how they work, by carefully observing the pitch of the syllables in the second rhythm groups. In fact, normally, the rhythm groups that are not separated by pauses, after a rhythm group with *accent*, remain mid-pitched as the preceding one.

Here are some phrases: *watakushi-no namae* [ɥe-te.kû.ĉi-nɔ-ne-me] (in [ne-me], *na-* becomes [ne], by pitch assimilation to the preceding syllable); *bòku-no señsèi* [-bɔ.kɥ-nɔ-seŋ-sei] (in *señsèi* [seŋ-sei], *-sèi* becomes low, [-sei]); *Nihoñ-chizu* [ni-hɔ-ɾi-ĉi.zu] (without modifications).

More phrases: *àsabañ* [e-se-be.ĉi] (from [-e-se] and *bañ* [-be.ĉi], by unifying everything into *one rhythm group* and lowering *bañ*); *yasashii hòñ* [je-se.ĉi-hɔ.ĉi] (without modifications); *ii jibiki* [ii-ĉi-bi.ĉi] (from *jibiki* [dĉi-bi.ĉi], with lowering of *-biki*, which becomes something like *jibiki*); *suzushii heyà* [su-zu.ĉi-he.ɾe] (in *heyà* [he-ɾe], *ya* is

lowered); *sen̄s̄ei d̄esu* [s̄ēŋ̄·s̄ei·d̄ēs̄ŷ̄, -s̄, -s̄] (with stress reduction on *d̄esu* [-d̄ēs̄ŷ̄, -s̄, -s̄] and lowering of *d̄ē*).

9.5.2. More: *totemo omoshir̄oi h̄on̄* [t̄o·t̄e·m̄o·o·m̄o·ŋ̄i·ŋ̄ōi h̄ōŋ̄] (from *omoshir̄oi* [o·m̄o·ŋ̄i·ŋ̄ōi], with raising of *o*- and lowering of *h̄on̄*, and stress shift); *hijoo-ni fu-kuzatsu-na mōndai* [h̄i·ŋ̄j̄ōo·ni·f̄u·k̄u·z̄a·ts̄u·na·m̄ōn̄·d̄ai] (with raising of *fu*- and *mōn̄*); *m̄otto yas̄ui h̄on̄* [m̄ōt̄t̄ō j̄ē·s̄ui h̄ōŋ̄] (with lowering of *-s̄ui* and *h̄on̄*); *ohayoo gozaim̄asu* [o·h̄ēj̄ōo·ḡo·z̄ēi·m̄ē·s̄ŷ̄, -s̄, -s̄] (with raising of *go*- in *gozaim̄asu* [ḡo·z̄ēi·m̄ē·s̄ŷ̄, -s̄, -s̄]).

Some others: *ar̄iḡatoo gozaim̄asu* [ē·ŋ̄i·ŋ̄ē·t̄ōo·ḡo·z̄ēi·m̄ē·s̄ŷ̄, -s̄, -s̄] (with lowering of *-zaim̄a*:- everything is low, except *-ri*); *d̄oomo ar̄iḡatoo gozaim̄asu* [-d̄ōo·m̄o·ē·ŋ̄i·ŋ̄ē·t̄ōo·ḡo·z̄ēi·m̄ē·s̄ŷ̄, -s̄, -s̄] (with lowering of *-ri* and *-zaim̄a*:- everything is low, except *d̄oo*-, which is half-low falling).

In addition: *uch̄i-kara ēki-m̄ade arukim̄asu* [ū·t̄ŋ̄ī -k̄ē·ŋ̄ē -ē·cī m̄ē·d̄ē | ē·ŋ̄i·cī·m̄ē·s̄ŷ̄, -s̄, -s̄] (in the case of fast speech with no breaks: [ū·t̄ŋ̄ī -k̄ē·ŋ̄ē -ē·cī m̄ē·d̄ē ē·ŋ̄i·cī·m̄ē·s̄ŷ̄, -s̄, -s̄]. But, separately, word by word (notice the stress change in the first rhythm groups, mainly due to /i/ devoicing): [ū·t̄ŋ̄ī, k̄ē·ŋ̄ē, -ē·cī, -m̄ē·d̄ē, ē·ŋ̄i·cī·m̄ē·s̄ŷ̄, -s̄, -s̄].

Also: *n̄iji-kara s̄an̄ji-m̄ade jim̄usho-ni im̄asu* [-ŋ̄i·ŋ̄ī·k̄ē·ŋ̄ē | s̄ēŋ̄·d̄ŋ̄ī m̄ē·d̄ē | d̄ŋ̄ī·m̄ū·ŋ̄j̄ō·ŋ̄ī ī·m̄ē·s̄ŷ̄, -s̄, -s̄]; possibly with fewer breaks, in less slow speech: [-ŋ̄i·ŋ̄ī·k̄ē·ŋ̄ē -s̄ēŋ̄·d̄ŋ̄ī m̄ē·d̄ē | d̄ŋ̄ī·m̄ū·ŋ̄j̄ō·ŋ̄ī ī·m̄ē·s̄ŷ̄, -s̄, -s̄]. On the other hand, in separate rhythm groups, we have: [-ŋ̄i·ŋ̄ī, k̄ē·ŋ̄ē, -s̄ēŋ̄·d̄ŋ̄ī, -m̄ē·d̄ē, d̄ŋ̄ī·m̄ū·ŋ̄j̄ō·ŋ̄ī, ī·m̄ē·s̄ŷ̄, -s̄, -s̄] (with modifications on *kara*).

Or: *ȳoji-m̄ade-ni koko-ni k̄ite kudas̄ai* [-j̄ō·ŋ̄ī m̄ē·d̄ē·ŋ̄ī | k̄o·k̄o·ŋ̄ī·cī·t̄ē·k̄ū·d̄ē·s̄ēī] (slower: [-j̄ō·ŋ̄ī m̄ē·d̄ē·ŋ̄ī | k̄o·k̄o·ŋ̄ī·cī·t̄ē·k̄ū·d̄ē·s̄ēī]. Or: [-j̄ō·ŋ̄ī, m̄ē·d̄ē·ŋ̄ī, k̄o·k̄o·ŋ̄ī, cī·t̄ē, k̄ū·d̄ē·s̄ēī], still with stress modifications on *m̄ade-ni*, in comparison with *m̄ade*).

9.5.3. Here are further examples illustrating stress modifications, in forming rhythm groups: *m̄ade-ni-wa* [m̄ē·d̄ē·ŋ̄ī·ŷ̄ē], *Okinawa* [o·cī·n̄ē·ŷ̄ē], *Okinawa-wa...* [o·cī·n̄ē·ŷ̄ē·ŷ̄ē], but sometimes also [o·cī·n̄ē·ŷ̄ē·ŷ̄ē].

Obviously, it is not always easy to readily distinguish between the effect of pitch prominence and stress prominence. On the contrary, when mid pitch coincides with stress, prominence is quite clear.

If all this combines with a fairly heavy syllable, prominence is even more evident. In any case, if several nearby syllables share the same characteristics, it becomes less easy to distinguish them clearly.

However, *pitch* remains the most important element, being the distinctive one, though undoubtedly stress has a considerable role, too. It is important to find an appropriate balance between these elements, although oscillations are quite possible and normal, indeed.

In an example like *tabem̄ash̄ita ka?* [t̄ē·t̄ē·b̄ē·m̄ē·ŋ̄ī·t̄ē·k̄ē], the low pitch and secondary stress on the syllable *ta*- and the mid pitch and weak stress on *-be*- may give a similar prominence effect. But, of course, it is less than that on *-m̄a*-, and decidedly lesser

than that on *-ta*, and even less so than on *-shi-*, with devoicing up to the loss of its syllabicity: [ç̥i, ç̥i, ç̥].

In the case of *loanwords* (or *gairaiō*), the interplay of pitch and stress (as well as of syllabic weight and *accent* placing) can undergo some hierarchic reversal, by moving closer to the original (stress) pattern: ‘hotel’ (*hōteru*) [hō\_tɛɽɽ], and ‘coffee’ (*kōhī*) [kōsç̥hii], but also [ç̥kōsç̥hii] is possible (and even [ç̥kōsç̥ɸɽɽii]).

9.5.4. Obviously, in *current speech*, some *reduction* phenomena occur in Japanese, too. Here we will consider some of the most ‘normal’ ones. The particles *no* and *ni*, in particular, are often reduced to /ŋ/: *kūru-no nara* [-kɽɽŋn̄ -nɛɽɽ, -kɽɽŋn̄s -nɛɽɽ], *kimi-no uchi* [ç̥i-miŋ̄ ·ɽ-tç̥i, ç̥i-mi-n̄s ·ɽ-tç̥i], *gēnki-ni naru* [-gɛŋ̄ç̥iŋ̄ -nɛɽɽɽ, -gɛŋ̄ç̥i-n̄i -nɛɽɽɽ].

In the *negative*, forms with *-rVnai*, generally, change /ŋV/ into /ŋ̄/: *shiranai* [ç̥iŋ̄-nɛi; ç̥i-ŋ̄ɽ-nɛi], *okurenai* [ç̥s̄kɽŋ̄-nɛi; ç̥s̄kɽŋ̄ɽ-nɛi].

More frequently, there can even be contractions such as: *mīte shimau* [ç̥mi\_tç̣ɽɽ; -mi\_tç̣ɽ ç̥i-mɛɽɽ], *yōnde shimau* [ç̣ɽŋ̄\_dç̣ɽɽ; -ç̣ɽŋ̄dɽ ç̥i-mɛɽɽ].

Let us also consider: *watashi* [ɽɽ-tɽç̣i; ɽ-tɽç̣i; -ɽɽç̣i] (instead of its refined form *watakushi* [ɽɽ-tɽkɽç̣i, -ç̣i]), *dewa* [-dɽɽɽ, -dç̣ɽɽ] (also lexicalized as *ja*), *dayo* [dɽ-jç̣, dɽ-ç̣] (even [dɽ-ç̣, d̄i-ç̣, dç̣i-ç̣], with two exceptional phones: [a] and [d] followed by [i]), *haiiro* [-hɛiiç̣, hɛ-iiç̣, -hɛiç̣], *ichinichi* [i-tç̣i-n̄i-tç̣i, i-tç̣iŋ̄-tç̣i], *arāmi* [ɽ-ɽɽmi, ɽ-ɽɽ-ɽmi, ɽ-ɽɽmi], *ume* [ɽ-mɛ, ŋ̄-mɛ, ŋ̄-mɛ, ŋ̄-mɛ].

# 10. Intonation

10.1. For a complete treatment of intonation and prosodic & paraphonic features in language, the readers are invited to see ¶ 12-14 of *Natural Phonetics & Tonetics*, or the correspondent updated sections in the *canIPA* website.

INTONATION is constituted by the relative pitch of syllables forming more or less long sequences of connected speech. These sequences are called TUNINGS and can consist of pause groups (which, in turn, consist of rhythm groups); but they can also consist in a single word – which can even be monosyllabic: *No. – No? – No! – No...*

What is essential is that pitch –through given differences– adds (or, rather, gives) different pragmasemantic nuances –such as ‘statement, question, command’ &c– to phonic sequences which could otherwise be identical.

Thus the difference obtained is not merely semantic, conceptual, as in the case of ton(em)e languages, such as Chinese or Vietnamese.

However, by using the same principles and the same symbols of syllabic-tone notation, one can accurately (and without too many problems) transcribe the characteristics of pitch and strength of the syllables of a whole utterance.

In fact, stress-tonal signs show both the relative pitch and degrees of stress on the syllables before which they are put.

First, let us see (fig 10.1) an iconic and simple way to introduce people to intonation (applied to neutral British English, as recordings are easy to be found): by carefully reading the examples given, and following the heights shown for every grapheme.

After, we can go to fig 10.2, where one can see the whole pitch extension of an utterance, which is called a TUNING (or intonation group). It is divided into a PROTUNE

fig 10.1. ‘Icono-tono-graphic’ representation of neutral British-English intonation.

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

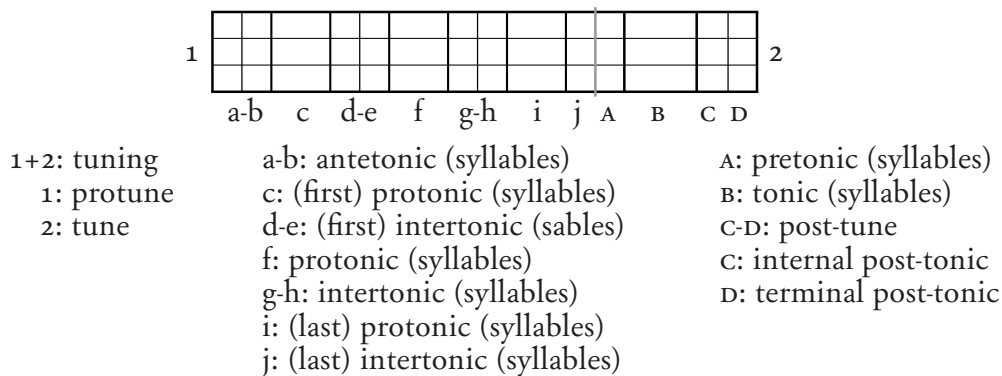


and a TUNE. Here, let us anticipate that a general *tune* consists of three parts: a PRE-TONIC syllable, the TONIC (*ie* the stressed) one, and (two) POSTTONIC syllables.

A *protune* consists of one or more *stressed* syllables and some *unstressed* ones (which are called ‘protonic’ and ‘inter-tonic’ syllables, respectively).

Sometimes, it might be important to refer explicitly to the first or last ‘protonic’ syllable, in the description of certain languages with particular protunes. Usually, the first protonic can be preceded by some ANTETONIC (*ie* initial unstressed) syllables.

fig 10.2. The structure of tunings.



10.2. In anticipation of what will be dealt with presently, one may say that there is a ‘normal’ *protune*, for statements, which has no particular symbol since it is the unmarked one:  $/\cdot/$ . There are, then, three marked protunes: *interrogative* ( $/\dot{\cdot}/$ ), *imperative* ( $/\dot{\cdot}|/$ ), and *emphatic* ( $/\dot{\cdot}|/$ ).

We must make it clear at once that *written* sentences are one thing, while the *spoken* language is quite another reality, often very different, indeed. Naturally, in the spoken language, tunings are much more numerous than ‘simple sentences’ of grammar and syntax, as will be seen below.

But let us consider *tunes*. Generally they are formed by the *tonic* syllable (*ie* the stressed one, which is also the last strong syllable in an utterance, in a sense).

The *pretonic* (*ie* the possible unstressed syllable before it), and the *posttonic* syllables (*ie* the possible unstressed syllables after it).

In the tonetic diagrams (or rather *tonograms*), two posttonic syllables are indicated (*ie* the *internal* and *terminal* ones). Sometimes it is useful to refer to one of them, clearly, in order to highlight typical movements more clearly, above all to distinguish interrogative tunes of the rising type ( $[\cdot' \cdot]$ ), from those of the falling type ( $[\cdot' \cdot']$ ). In any case, the term POST-TUNE may be used to refer to both syllables, collectively.

Let us now consider –concisely (and by looking closely at fig 10.3)– the three marked tunes (of neutral British English): *conclusive* ( $/\cdot/$ ), *interrogative* ( $/\dot{\cdot}/$ ), *suspensive* ( $/\dot{\cdot}|/$ ), and the unmarked: *continuative* ( $/\cdot|/$ ): *On Saturday*  $/\cdot|/$  [ $\text{ɒn}'\text{sæ}\text{t}\text{-}\text{ə}\text{d}\text{e}\text{ɪ}\cdot$ ], *On Saturday?*  $/\dot{\cdot}\dot{\cdot}|/$  [ $\dot{\cdot}\text{ɒn}'\text{sæ}\text{t}\text{-}\text{ə}\text{d}\text{e}\text{ɪ}\cdot$ ], *(If not) on Saturday...* (*then...*)  $/\dot{\cdot}|/$  [ $\text{ɒn}'\text{sæ}\text{t}\text{-}\text{ə}\text{d}\text{e}\text{ɪ}\cdot$ ], (*Perhaps*) *on Saturday*, (*but...*)  $/\cdot|/$  [ $\text{ɒn}'\text{sæ}\text{t}\text{-}\text{ə}\text{d}\text{e}\text{ɪ}\cdot$ ].

10.3. The best way of dealing with the intonation of a language consists in presenting its structures through appropriate and clear diagrams (*ie* tonograms), with clear examples and a simple and sufficiently complete notational system (not a cumbersome and useless one).

First of all, we must repeat that the use and choice of intonation patterns do not depend on syntax at all, but on *semantics* and *pragmatics*, and above all on *communicative goals*.

In fact, even if the syntactic formulation is, in the end, the most evident linguistic rendering (for those who are used to reading and writing), in actual fact it is nothing but a faithful representation of the pragma-semantic way to express concepts and thoughts, peculiar to every language.

If, for instance, one writes –and beforehand says– *I've been looking for this for ages* [æʊnbʌn ˈlɔk-ɪŋ fəʊɪs. fɪˈeɪdʒɪz.], the superficial formulation at hand is only the inevitable result of the mental and linguistic processes that produce, in English, the sentence just seen, although with slight possible variations.

In actual fact, it results from the juxtaposition of different concepts (each one indicated by /./, or [·' ..]) in a single syntactic string, seemingly simple and straightforward, but actually very complex, as is obvious from its prosodic structure, if supported by an appropriate intonation pattern, as indicated by the small but precious signs used.

Let us now examine the intonation structure of neutral British English. However, one must first consider a general scheme which will enable people to really *see* its characteristics.

Thus *fig 10.2* gives the diagram of tunings (or intonation groups). It shows the use one makes –when speaking normally– of pitch heights on the various syllables forming the different possible utterances in a given language.

## Tunings

10.4. Tunings consist (as already seen) of a protune (in our example *I am transcribing the following example* [æməˈtʃrɪbɪŋ ðəˈfɒl-əwɪŋ ɪɡˈzɑːmpəl]) and a tune (*phonetically* [fəˈneɪ-ɪkli.]). In this case, one has a normal protune and a conclusive tune.

The latter is represented, tonemically (in a theoretical way) by /./, and tonetically (in a more realistic way) by [·' ..].

The number of syllables in the example has been calculated on purpose in order to have full correspondence between the tonogram and the syllables of the sentence, to be able to show the characteristics more clearly.

Of course, in normal speech, it is unlikely to find sentences with the same number of syllables; however, the usefulness of the diagram is not compromised, since the actual syllables available (whether more or less than 14) share pitch heights in a fair way.

So they may either compress the movement of several syllables into only one or two, or expand it over a larger number of syllables (cf *fig 10.5* for the tunes).

For instance: *Yes, we do* or *Our aim is to pass on ideas, techniques, and practical activities, which we know work in the classroom* (even if this last example, more realistically, will be divided into more parts, with the addition of the respective tunes, mostly continuative).

Thus: *Our aim is to pass on ideas, techniques, and practical activities, which we know work in the classroom*. In a phono-tonetic transcription, you have: [ɑː̄ ɪ̄ ɪ̄ m ɪ̄ z̄ ɹ̄ ə ˈpʰɑːs ˈbːnː ə ˈdɪːz̄ ɹ̄ hɛkˈniːks̄ ə m̄ ɹ̄ hɪ ˈæ k ɹ̄ ɪ k ɹ̄ ə k ɹ̄ hɪ ˌv ə ɹ̄ ɪ z̄ ɹ̄ wɪ ɹ̄ wɪ ˌn ɹ̄ ɔ ˌw ɹ̄ k̄ ɹ̄ ɹ̄ ʌ n ɹ̄ d̄ w̄ ˌk h l ɑː s ɹ̄ ɹ̄ ɔ m̄ .].

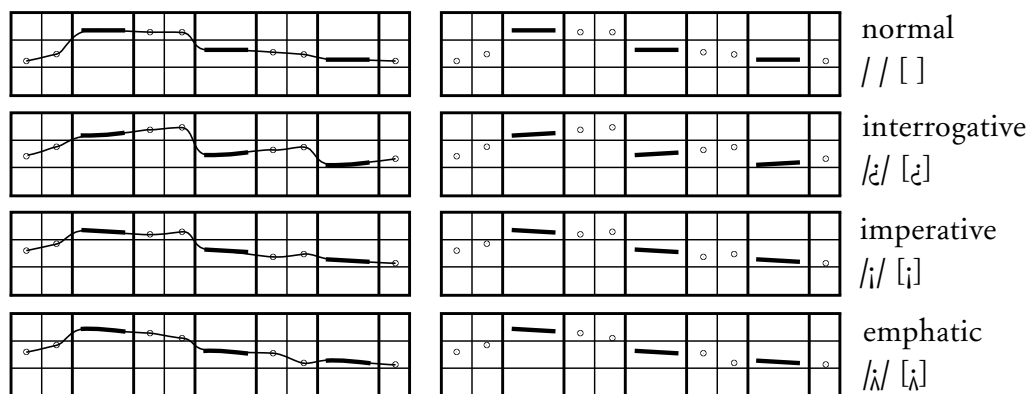
## Protunes

10.5. fig 10.3 shows the four protunes (of neutral British English): one is unmarked, or *normal*, and has no symbol; three are marked: *interrogative* /Ǿ/ [Ǿ], *imperative* /i/ [i] (for instance: *Pay attention!* [ɪ̄ ˌpʰɛɪ ə ˈtʰɛnʃ̄n̄.]), and *emphatic* /λ/ [λ] (*We have to check everything very carefully!* [λ̄ wɪ ˌhæ ɹ̄ ə ˈtʰɛk ˌɛ v̄ ɹ̄ i θ̄ ɪ ɹ̄ ˌv eɪ ɹ̄ i ˈk h eɪ ɹ̄ f i .]).

fig 10.3 shows, on the right, sketchy tonograms; on the left, they are given in a more realistic way. Actually, the schematic diagrams are sufficient indeed (as will be done for Japanese), since these tonograms necessarily generalize and normalize the data, allowing slight differences of realization, as well.

On the contrary, for teaching and learning purposes, these schematic tonograms are decidedly more useful, making comparisons with those of other languages not only possible, but even easier. Furthermore, the schematic tonograms are less distracting, ultimately, than the realistic ones.

fig 10.3. The four protunes of neutral British English.

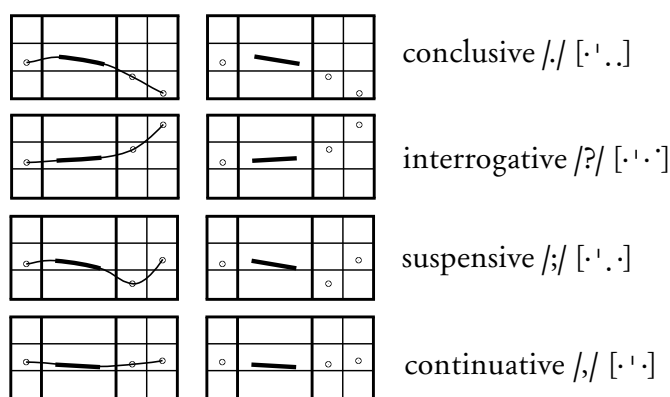


## Tunes

10.6. fig 10.4 shows the three marked tunes (of neutral British pronunciation, again both realistically and schematically) – *conclusive* /./ [ˈ..], *interrogative* /?/ [ˈ.ː], and *suspensive* /:/ [ˈ.ː] – in addition to the unmarked one, *continuative* /|/ [ˈ.ː].

The marked tunes have a functional charge, which is crucial for communica-

fig 10.4. The four tunes of neutral British English.



tion, as they oppose one another distinctively. The unmarked tune – the continuative one – may be considered as the neutralization of the three marked ones (since each of them would be inappropriate in certain – less important – contexts, being too specific and having very definite functions).

The aim of the continuative tune is, above all, to oppose a theoretical ‘zero’ tune. It is quite different from a straightforward and progressive flow of enunciation, without the slightest variations (or breaks), even theoretical or potential.

Its only purpose is to slightly highlight a word, compared to a complete non-occurrence of tunes (as happens within a protune).

Indeed, there is a difference between *I saw six men* [aəˈsɔː ˈsɪks ˈmɛnː.] and *I saw six men* [aəˈsɔː ˈsɪks ˈmɛnː.]; in the latter, of course, *six* is more prominent than in the former, since it has its own tune (although no pause follows it), instead of being a part of the same protune.

At the end of § 10.4, we have seen that a syntactic string does not generally correspond to just one tune; in fact, more or less numerous continuative tunes occur, otherwise the sentence would not sound spontaneous and convincing.

At first, one does not fully realize this internal subdivision, which is completely natural. Its appropriate use goes entirely unnoticed; whereas, its absence would *not* pass unnoticed at all (as happens in unprofessional reading or recitation).

For instance, if one considers an utterance such as *Look! the imprints of a bear*, it is soon realized that it can be said in many ways – apart from actual and paraphonic considerations such as the *fright* taken at the sight, or the *delight* expressed by naturalists, or the *satisfaction* felt by hideous poachers... (all of them are rendered with different nuances, clear and easy to interpret).

Of course, this is different from a unitary sentence such as *Look at the imprints of a bear*, in just one tuning: [ˈlɒk-ʌt ðɪˈɪmpɪnts əvəˈbɛːʒ..].

10.7. Thus, if you go back to the original utterance, what you find is something closer to a natural exposition, as *Look: the imprints of a bear* [ˈlɒk. ðɪˈɪmpɪnts əvəˈbɛːʒ.]; in fact, in the same sentence, there are two pragmatic concepts: the imprints and the sighting of them.

If one then divides it into three parts (of course, with three tunes), the nuances

expressed are more detailed: *Look at the imprints of a bear* [lɒk. ði'ɪmpɪnts. əvə'beɪ.]; in this way, one can manage to separate, conceptually too, imprints of different shapes.

After all, it is possible to use some continuative tunes (ie unmarked /,/ as already seen in the previous section), and this will add something to elocution (in opposition to a unitary utterance, although this is not for emphasis, of course).

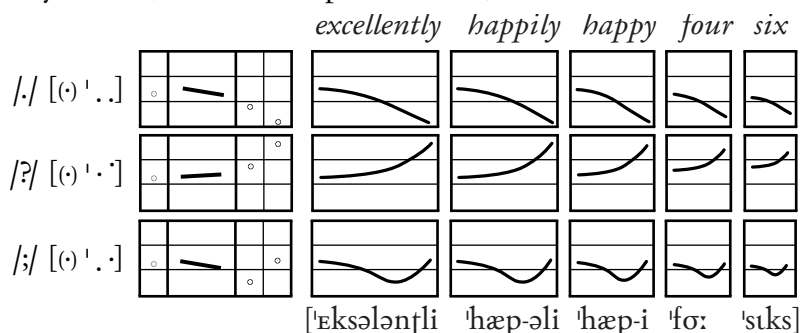
It is only a way to make enunciation a little more effective and natural: [lɒk.wɪ. ði'ɪmpɪnts. əvə'beɪ.]. (and variations).

By considering an example like *You must read further books on this particular subject*, again, one can easily see that there are several ways of saying it.

Apart from a quite flat realization in a single tuning, as: [jʊməs'ɪrɪdʒ 'fɜːðə 'bʊks ɒn,ðɪspə'fɪkʃələ 'sɛbdʒɪkʃ.], one can have: [jʊməs'ɪrɪdʒ 'fɜːðə 'bʊks. ɒn,ðɪspə'fɪkʃ-ʃələ 'sɛbdʒɪkʃ.], or: [jʊməs'ɪrɪdʒ. 'fɜːðə 'bʊks. ɒn,ðɪspə'fɪkʃ-ʃələ 'sɛbdʒɪkʃ.], or else: [jʊməs'ɪrɪdʒ. 'fɜːðə 'bʊks. ɒn,ðɪspə'fɪkʃ-ʃələ 'sɛbdʒɪkʃ.].

One could also have [jʊ-], or even: [jʊ- məs'ɪrɪdʒ. 'fɜːðə 'bʊks. ɒn,ðɪspə'fɪkʃ-ʃələ 'sɛbdʒɪkʃ.]. (with more and more numerous nuances and implications).

fig 10.5.1. The four tunes of neutral British English, expanded or contracted according to the number of their syllables (here with no pretonic one).



10.8. A *conclusive* tune is necessarily used whenever a given concept is completed in the speaker's mind. Thus, in addition to the words which form the sentences, it concerns communicative functions as well, as if, in saying *It's raining cats and dogs*, you added 'I am stating' – so: *It's raining cats and dogs* [ɪts'ɹeɪnɪŋ 'kætɪz ən'dɒɡz.].

Each tune has a specific function: the *interrogative* communicates 'I am asking': *Is it raining cats and dogs?* [ɪzɪt'ɹeɪnɪŋ 'kætɪz ən'dɒɡz.]; the *suspensive* one communicates 'I am underlining': *If it's raining cats and dogs... (it's a calamity!)* [ɪfɪts'ɹeɪnɪŋ 'kætɪz ən'dɒɡz. | (ɪts'ɹeɪnɪŋ 'kætɪz ən'dɒɡz. |)].

The *continuative* tune, instead, simply communicates 'I'm not finished': *It's raining cats and dogs (but I don't care)* [ɪts'ɹeɪnɪŋ 'kætɪz ən'dɒɡz. (bət'aedʒəŋ'kheɪ.ɪ)].

It is possible to have a series of conclusive tunes: *Yesterday it rained. Today it's raining. Tomorrow it'll pour. I'm sick and tired. I'll go away!* [ˌjɛstə'deɪ ɪt'ɹeɪnd. | ˌtəʊdeɪ ɪt'ɹeɪnɪŋ. | ˌtɒmɒrəʊ ɪt'ɹeɪnɪŋ. | ɪ'm sɪk ən'deɪtɪd. | ɪ'l ɡoʊ əweɪ.]. However, a suspensive tune is very likely for *Tomorrow it'll pour* [ˌtɒmɒrəʊ ɪt'ɹeɪnɪŋ.].

Too often, current writing (which is not at all sophisticated) uses only commas:

*Yesterday it rained, today it is raining, tomorrow it'll pour, I am sick and tired, I'll go away.*

Thus, with the guilty complicity of schools, one is led to a kind of 'child-like' reading, which makes people utter things like: [°jɛsɪədeɪ ʊtʃeɪndɪ..° fə'deɪ ɪtsʃeɪnɪŋ..° fə'mɔɪzɔ ʊtʃphɔ:..°] æm'sɪk ən'fhaeədɪ..° æft'gɜ'w ɪ'weɪt..°].

The small rings show the additional pitch movement which is typical of 'bookish intonation', which must be kept well apart from normal (ie conversational) intonation, and also from the typical intonation of text exposition (even if simply done mentally).

A further example to show that, normally, writing and punctuation are just miserable devices with syntactical functions, and not at all helpful for reading: *I'm terribly busy: I can't come; I'll let you know; don't be cross* [æm'tʃheɪəbli 'bɪzi.. æ'kʰɑŋ'kʰɛm:..] æft'letʃ-ɪ 'nɜ'w.. ɪ'dʒɔm bɪ'kʰɪps:..].

Also the example *I've been looking for this for ages* [ɪəvbɪn'lɔk-ɪŋ fə'dɪs.. fʃeɪdʒɪz:..] shows this characteristic.

Contrary to what grammars keep on repeating, a *comma* does not necessarily indicate a short pause, as a *semicolon* does not indicate a pause which is half-way between the 'short' one of commas and the 'long' one of *full stops* (as it is absurdly 'prescribed').

However, these are the results achieved by schools, ie sadly rigorous and monotonous pauses, which are not able to convey appropriate meaning to sentences (especially when they are read).

And all those who today abuse punctuation, by omitting it almost completely, will they ever pause?

fig 10.5.2. Difference between *total* questions (1) and *partial* questions (2).

1	(Will they) see you on Saturday?
2	(Why won't they) see you on Saturday?

## Parentheses & quotations

10.9. Lastly, fig 10.6 shows the diagram of *parenthetical phrases*, or simply *parentheses* (either *low*, [ ɪ ]]; or *mid*, [ ɪ ɪ ]), and of *quotations*, [ ɪ ɪ ] (*high*). In the following sentence, each of them occurs once:

*First of all –he said– let's consider 'Natural' Phonetics, as it's properly called.*

[fɜ:st əv'ɔ:t' hɪ'seɪd' lets kʌn'sɪd-ə 'nætʃ-ɪtʃ:.. fə'neɪ-ɪks:.. ɪzɪts'phɪtɒp-əli 'kʰɔ:tʃd:..].

*Parentheses* typically feature an overall reduction of their stress and an increase in the rate of speech, while the pitch is compressed in the low –or mid– range of the tonogram.

*Quotations*, instead, are quite the opposite, as their role is –precisely– to put one or more words in full evidence, by means of a slightly louder and distinct enunciation: thus, their stress is increased, their speech rate is reduced, and their pitch is raised (*without* compression).

Marking such prosodic subtleties in phonetic –let alone *phonemic*– detailed transcriptions is neither necessary nor recommendable. The symbols [ ˩ ], [ ˨˨ ], [ ˨˨˨ ] are more than sufficient to bear in mind all these differences, with respect to ‘normal’ utterances.

Thus, even with regard to stress (and taxotones, as far as tone languages are concerned) there is no need to deviate from the usual notation.

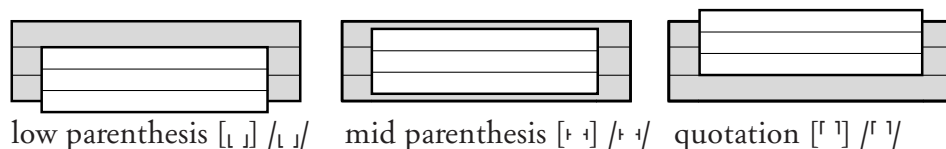
Quotations must not be confused with *direct speech*. Let us go back to *First of all –he said– let’s consider ‘Natural’ Phonetics, as it’s properly called*.

In that sentence, only *he said* should be excluded, because all the rest –and what may follow– *is* direct speech, indeed.

Japanese *parentheses* are low, [ ˩ ], but mid, [ ˨˨ ], after non-conclusive tunes; *quotations*, as said, are high, [ ˨˨˨ ].

Such intonational superstructures are quite useful in tone languages, thanks to small variations in pitch (small, yet perceptible by native speakers, but not so much –at the very beginning– by foreign learners, who have already a hard time struggling with tones and taxotones in ‘normal’ utterances), or in the Japanese tone-patterns.

fig 10.6. Tonograms of parentheses and quotations.



## Japanese intonation

10.10. As far as intonation is concerned, fig 10.7 shows the four *protunes* and *tunes* of Japanese. Any combination of two of them is called a *tuning*, as we know.

Of course, their ‘stritch’ patterns (*ie* the *stress & pitch* patterns) of the rhythm groups, that form the phrases and sentences of Japanese texts, are the basic material on which intonation is added.

Thus, there is a complete interaction between these two partially different prosodic elements, in order to produce the natural results that we can hear.

Foreign learners are advised to listen to high-quality recordings of reliably ‘neutral speakers’, in particular the professionally-trained ones.

Therefore, we cannot help recommending our readers, especially those whose native language works differently from English, to study Japanese intonation just like beginners learn how to play the piano.

That is to say, by starting from practicing single keys, even if they are already proficient at another keyboard instrument, and then move on to more complicated sequences.

Aural imitation –but not prone aping!– plays an important role in acquiring linguistic intonation (and even more so, paraphonic intonation).

So, we are confident that a good number of examples accurately annotated with our intonational symbols will greatly help learners get the most out of their listening practice.

fig 10.7. The four protunes & tunes of Japanese.

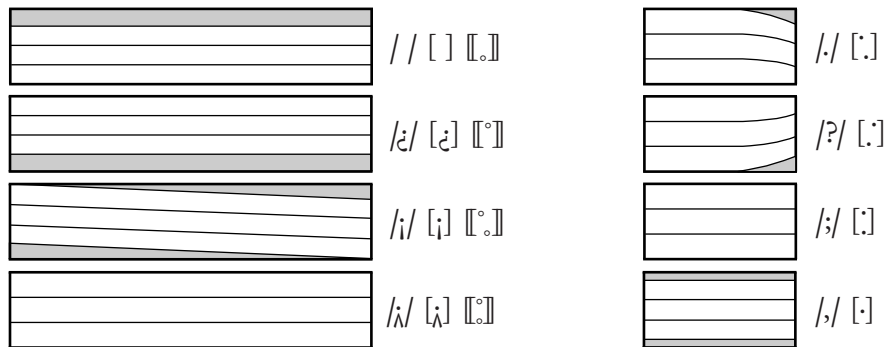
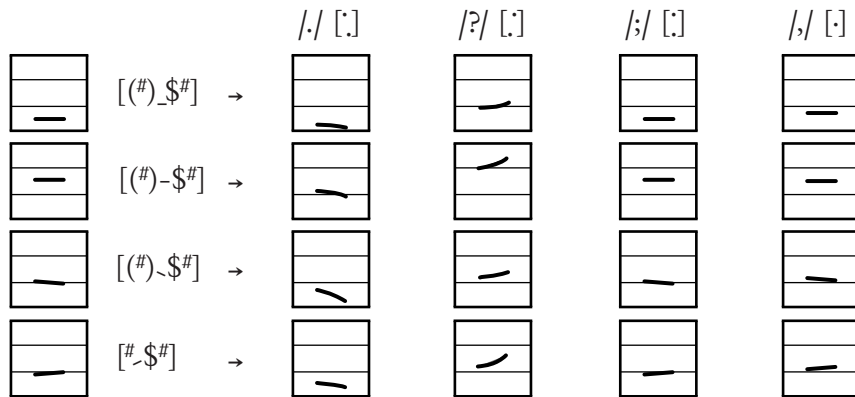


fig 10.8. The pitch patterns of the four tunes of Japanese combined with a tonic syllable.



10.11. Now, let us provide some sufficient examples to practice the intonation patterns of Japanese. Let us start from the *conclusive* tune, /./ [.] , in the following examples (further examples of it are to be found when dealing with the other tunes):

*Nihoñgo-ga joozù-ni hanàsetara ii-no-ni nà(a).*  
 [ni-hoŋŋo-ga .joozù-ni .hanàsetara .ii-no-ni .nà(a). (-ne')]  
 (I'd like to speak Japanese well)

*Anàta-no iitai kotò-wa wakàru kèdo-ne.*  
 [e-ne.te.no .ii-tei .ko-to-wa .(ko-) .wa-ke.ru .ke.do.ne.]  
 (I know what you mean)



*Hoñtoo-ni ariğatoo.*

[.hʊŋ-tʊʊ-ni ɐ-ɾi.ŋɐ.tʊʊ. (·ɐ)]

(Thank you so much)

10.12. The conclusive tune is used in *partial questions*, as well, but in that case it follows an *interrogative protune*, /ç̣ . / [ç̣ ˙].

Depending on the relevant paraphonic settings used, and the very issue referred to in the conversation, such combination may sound like anything ranging from a neutral question without any particular undertone, down to a cold-hearted, even hostile, police-style interrogation.

In order to avoid misunderstandings, all languages feature a somewhat milder version of asking *partial questions*, in general with a *continuative* tune instead of a conclusive one: /ç̣ , / [ç̣ ˙] (in order to sound gentler, especially talking to strangers).

By keeping the post-tonic coda in the mid tonal band, instead of letting it fall brusquely, the hearer perceives that the question is posed with discretion, almost with a courteous hesitation.

It is not rare, however, that foreigners lacking politeness <sup>or</sup> education use the conclusive tune, often, making them sound rude and impolite...

So, /ç̣ , / should be regarded as *the first and primary* pattern to be chosen, and taught to foreigners, as far as partial questions are considered.

Thus, /ç̣ . / should be reserved to informal and casual conversation, not talking to strangers, or if one *really wants* to convey indifference, impatience, dissatisfaction, suspicion, hostility...

Regardless of how gentle a speaker aims to sound, let us make it quite clear, though, that it is completely wrong to conclude a partial question with a full interrogative pattern, /ç̣ ? /.

However, that is quite typical and frequent on the part of many who passively interpret a question mark at the end of a sentence as a peremptory call for rising their voice pitch, to signal that they are putting a question.

Indeed, this will only produce incredibly annoying results (and possible misunderstanding, as well).

Therefore, let us examine the following examples:

*Kore, dōo omōu?*

[ç̣.kʊ-ŋɐ. ç̣-dʊʊ .ʊ-mʊʊ.]

(What do you think about this one?)

*Kyōo, chooshi dōo?*

[ç̣.ç̣ʊʊ | ç̣-ṭʊʊç̣.ç̣i ç̣-dʊʊ.]

(How are you feeling today?)

*Dōko iku?*

[ç̣-dʊ.kʊ i-kʊ.]

(Where are we going?)

10.13. A question requiring an overall answer, such as *yes* or *no* (or *maybe*, *dunno*, &c), is called a *total question*, and requires an interrogative tune on the word, or group of words, which the question is focused on: [ç :].

That implies that the interrogative tune may not necessarily occur at the end of the sentence, as the graphemic question mark, once more, leads many to think.

Secondly, more than one interrogative tune (and possibly as many interrogative protunes) may occur in a row, in longer and more articulated questions, even though only one ‘total’ answer is expected, anyway.

Let us examine the following examples:

*Nihoñgō shaberèru?*

[ç.ni-hσñ.ησ| ç.çjε·bE-ŋE.ŋɥ.]

(Can you speak Japanese?)

*Otootosàn-wa wakàru?*

[ç.σ-tσσ·tσ.sεñ.ɥε .ɥε-kε.ŋɥ.]

(Does your brother understand it?)

*Kàre, ashità, kùru?*

[ç-kε.ŋE| l.ε·çi-tε| ç-kɥ.ŋɥ.]

(Is he coming, tomorrow? / [Speaking of him,] tomorrow, is he coming?)

10.14. In most languages, *alternative questions*, like the following ones, are said with (/ç ;/ +) /ç ;/ + /ç ./ – that is they end with a conclusive tune, /ç ./, which is preceded by a suspensive one, /ç ;/.

In case of more than two parts, all the others occurring before /ç ;/ + /ç ./ are generally said with a continuative tune, /ç ;/.

But, in Japanese, this kind of questions is conceived differently. In fact, it becomes a series of two (or more) total questions: (/ç ?/ +) /ç ?/ + /ç ?/.

It is true that only the last one is treated as a full total question, while the others can be attenuated, /ç ?/ (cf § 10.19), though not necessarily.

*Bàsude iku? (Sore-ka) arùite iku?*

[ç-bε.sɥ.de i-kɥ.] (sσ-ŋE·kε·) ç.ε-ŋɥi.te i-kɥ.]

(Shall we go by bus, or on foot?)

*Bàsude iku, deñsha-de iku, (sore-tòmo) kuruma-de iku?*

[ç-bε.sɥ.de i-kɥ.] ç-dεñ.çjε·de i-kɥ. (i)| ç(sσ-ŋE·tσ.mσ) kɥ.ŋɥ-mε·de i-kɥ. (i)]

(Are you going by bus, by train, (or) by car?)

10.15. The *suspensive tune* is used to create... ‘suspense’, with very different implications depending on whether the sentence is a question or a statement and, as usual, the nuances provided by paraphonics.

As a general rule, /;/ [.] calls for the hearer’s attention on a part of the sentence, or simply adds vividness to long sentences made of multiple clauses, which would otherwise sound flat and inexpressive, or even hard to parse into meaningful units.

In some languages, like neutral (and regional) Italian, the suspensive tune is tonetically more conspicuous than in others (as, for instance, modern Arabic, where it is only minimally higher than the unmarked *continuative tune*).

In Japanese, it is generally accompanied by a perceptible pause, before completing the sentence. Carefully, let us examine the following examples:

*Anàta-gā doyòobi-ni korarenàkattara, watashìtachi (watàshìtachi) komàru-yo.*

[æ-nɛ.tɛ.ŋɛ .dɔ-jɔɔ.bi.ɲi .kɔ.ŋɛ-ŋɛ.nɛ .kɛt.tɛ.ŋɛ.] .ɥɛ-tɛ.ç̣i.tɛ.ṭç̣i (-ç̣i-) .kɔ-mɛ.ŋɥ.jɔ.]  
(If you can't come on Saturday, we'll be in trouble!)

*Èki-ni tsùita tokì-ni-wa, deñsha-wa shuppatsu shìtetañda.*

[-ɛ.çi.ɲi -tsɥi.tɛ (-tsɥi.i.tɛ) .tɔ-çi.ɲi.ɥɛ.] -dɛ.ŋ̣.ç̣jɛ.ɥɛ .ç̣jɥp-p:ɛ-tsɥ̣ .ç̣i-tɛ-tɛ.ɲ.dɛ.  
(ç̣i-)]

(When I arrived at the station, the train had gone!)

*Ìkko, nìko, sàñko, yòñko, gòko.*

[.ik.kɔ.ɔ̣ .ɲi.kɔ.ɔ̣ .sɛ.ŋ̣.kɔ.ɔ̣ .jɔ.ŋ̣.kɔ.ɔ̣ .gɔ.kɔ.ɔ̣.]

(One, two, three, four, five) – *complete list* (cf § 10.16).

10.16. The *continuative tune*, /,| [·], can be seen as the terminal part of an unmarked protune bearing a full stress, with two functions: to attract less attention than a suspensive tune, or to underline a word, or concept, without resorting to emphasis.

The following examples will be helpful (especially when compared with some of the similar ones given in § 10.14):

*Ìkko, nìko, sàñko, yòñko, gòko...*

[.ik.kɔ.ɔ̣ .ɲi.kɔ.ɔ̣ .sɛ.ŋ̣.kɔ.ɔ̣ .jɔ.ŋ̣.kɔ.ɔ̣ .gɔ.kɔ.ɔ̣.]

(One, two, three, four, five...) – *incomplete list*, with all /,| (and no /;| , cf § 10.14)

*Doyòobi-ni korenàkùtemo daijòobu da-yo.*

[dɔ-jɔɔ.bi.ɲi .kɔ.ŋɛ-nɛ.kɥ.tɛ.mɔ .dɛi-ç̣jɔɔ.bɥ .dɛ-jɔ.] (*colloquial* [dɛ-ɔ̣., dɛ-ɔ̣.  
.di-ɔ̣.], *also in other sentences*)

(If you can't come on Saturday, there's no problem) – only /,| (no /;|, cf § 10.14).

10.17. The last example comes at hand to show how *emphasis* can affect the general sense of a sentence (without actually changing its structure) by assigning more prominence to certain words.

Refer to the following examples, which also feature the *emphatic protune* />| [>]:

*Kore-wa totemo bèñri-na jìsho dèsu.*

[kɔ-ŋɛ.ɥɛ .tɔ-tɛ.mɔ .bɛ.ŋ̣.li.nɛ -ḍzi.ç̣jɔ.dɛ.sɥ̣.]

(This is a very useful dictionary)

*Kono jìsho, sùgòku bèñri da-yo.*

[kɔ.nɔ=ḍzi.ç̣ .sɥ.ŋ̣.kɥ (sɥ-ŋ̣.kɥ) .bɛ.ŋ̣.li .dɛ-jɔ.]

(This is a very useful *dictionary* / This *dictionary* is very useful)

*Kore, sùgòku bèñri-na jìsho-yo.*

[kɔ=ŋɛ.] .sɥ.ŋ̣.kɥ (sɥ-ŋ̣.kɥ) .bɛ.ŋ̣.li.nɛ .ḍzi.ç̣jɔ-jɔ.]

(This is a very useful dictionary)

*Sūgōku bēnri da-yo, kono jisho.*

[ʃ.su(ŋ)=ŋσ.ku(ŋ) ʃ.su(ŋ)=ŋσ.ku(ŋ) ˌbɛ̃.ŋ̄.li .dɛ-jσ .kσ-nσ-dʒi.çjσ.]

(This is a *very* useful dictionary / *Very very* useful, this dictionary).

10.18. Finally, we provide a few examples of *parentheses*, [ 1 ] & [ † † ], and *quotations*, [ † † ] (see fig 10.8):

*Īi-yo. Tàroo.*

[ˌiiσ. †-tɛ.ŋσσ. †]

(Of course. Taroo)

*Īi-yo, Tàroo.*

[ˌiiσ. †-tɛ.ŋσσ. †]

(Of course, Taroo)

*Mochìron̄ da-yo, Tàroo. Ashìta purèzeñto aḡeru-yo.*

[mσ-tçìŋŋσ .dɛ-jσ. †-tɛ.ŋσσ. †] ˌpɛ.çì-tɛ ˌpɛ.ŋɛ-zɛŋ-tσ ˌpɛ.ŋɛ-ŋɛ-jσ. (ˌpɛ.ŋɛ-ŋɛ-jσ.) (ˌpɛ.)]

(Of course, Taroo. Tomorrow I'll give you a present)

*Mochìron̄ da-yo, Tàroo, ashìta purèzeñto aḡeru-yo.*

[mσ-tçìŋŋσ .dɛ-jσ. †-tɛ.ŋσσ. †] ˌpɛ.çì-tɛ ˌpɛ.ŋɛ-zɛŋ-tσ ˌpɛ.ŋɛ-ŋɛ-jσ. (ˌpɛ.)]

(Of course, Taroo, tomorrow I'll give you a present)

*Hoñtoo-no kotò iu-to, añmari tàshìkajanaiñda.*

[l.hçŋ-tσσ-nσ .kσ-tσ .i-ɥ-tσ. † (i-)] ˌɛŋ-mɛ-ŋi ˌtɛ.çì.kɛ.dʒjɛ.ŋɛi.ŋɛ.]

(As a matter of fact, it's not so sure)

*Tàroo, señshuu kono èiḡa (|eiḡa) mīta no, obòetenai no?*

[†-tɛ.ŋσσ. †-sɛ̃.ŋ̄.çjɥɥ .kσ-nσ-ɛi.ŋɛ (ˌɛi.ŋɛ) -mī.tɛ.nσ. †] ˌç.σ.bσ-ɛ.tɛ.ŋɛi.nσ.]

(Taroo, you saw this movie last week, don't you remember?)

*Ûuñ, 'watashi-ḡa yaranàkatta-tte', ittèta-yo.*

[ˌɥ.ɥ.ŋ̄. †.ɥɛ-tɛ.çì.ŋɛ .jɛ.ŋɛ-nɛ .kɛ.t.tɛ.tɛ. † (jɛ-); †.it.tɛ.tɛ-jσ. †] (colloquial [†-tɛ-σ. †])

(No, 'I haven't done it', he said)

*Nàñde 'ìi yò' nàñte itta-no, hoñto-wa yokunàiñjanai no?*

[çì-nɛ.ŋ̄.dɛ †.ii-σ. †] ˌnɛ.ŋ̄.tɛ .it-tɛ-nσ. †-hçŋ-tσ-ɥɛ .jσ.kɥ.ŋɛi.ŋ̄.dʒjɛ.ŋɛi-nσ. (ˌjσ-)]

(I wonder, why did you say 'It's ok', when it's not ok?).

### Some general paratonic considerations

10.19. Let us also consider the use of pitch in a partially different situation in comparison with how intonation uses it. Paraphonically, differences can be greater and –what is more– of a fairly different quality, since feelings, moods, and states of mind make use of different –usually more ‘exaggerated’– pitch and length extensions.

Usually, in fact, they do not regard every single syllable of an utterance, as in tunes, where pitch is determined by linguistically codified structures. Although with possible variations, caused by changes in the factors determining their origin (ie moods, feelings, states of mind, attitudes, and physiological, psychological, and pathological conditions), paraphonic stretches mostly regard whole utterances of different lengths, even from minutes to hours.

With respect to the habitual pitch extent of a given speaker (within the three categories of voices, ie male, female, and infant), it will be sufficient to get accustomed to recognizing, first of all, the *raising* of general pitch (indicated by ⟨°⟩, between *angled* paraphonic brackets), visually and mnemonically helped by fig 10.9. There, the two white lines show the relation with the normal (pitch) extent, which is, then, remarkably modified indeed, in comparison with the one considered ‘normal’.

Now, for *raising* ⟨°⟩, it is sufficient to think of a label such as ‘astonishment’ or ‘aggressiveness’ and say something fitting, like: *And, what’s this?* (with astonished wonder) – *De, kore-tte, nàni?* ⟨°⟩[-dɛ| ɕ.ko-ɾɛt.tɛ -nɛ.ni:]. Or else: *Hey, you! Don’t make me angry!* (with arrogant aggressiveness) – *Dàkara, ìraira saseru-na yo, omae!* ⟨°⟩[ɰ-dɛ.kɛ.ɾɛ .i.ɾɛi.ɾɛ .sɛ-sɛ.ɾɛ.ɾɛ -jɔ .o.mɛ-ɛ:]. Let us note that sentences like those dealt with in this section on paraphonic characteristics often have a continuative tune, /,| [·], instead of a more expected conclusive one, /,| [˙].

Often, we simply use raising when we are talking to children, too: *Hello, little boy. What’s your name?* – *Koñnichi wa. Bòku-no namae-wa nàni?* ⟨°⟩[-koŋ.ɲi.tɕi -ɸɛ| ɕ-bɔ.kɸ.nɔ .nɛ-mɛɛ -nɛ.ni:] (with *bòku* as a second-person pronoun towards children).

With other labels, we can identify the opposite movement to raising, ie *lowering* ⟨˘⟩, as can happen for sadness: *Oh, that’s such a pity!* – *Àa, sore-wa zañnèñ da nèè* ⟨˘⟩[-aa| ɰ.sɔ-ɾɛ.ɸɛ .zɛŋ.n-ɸɛŋ .dɛ-nɛɛ:].

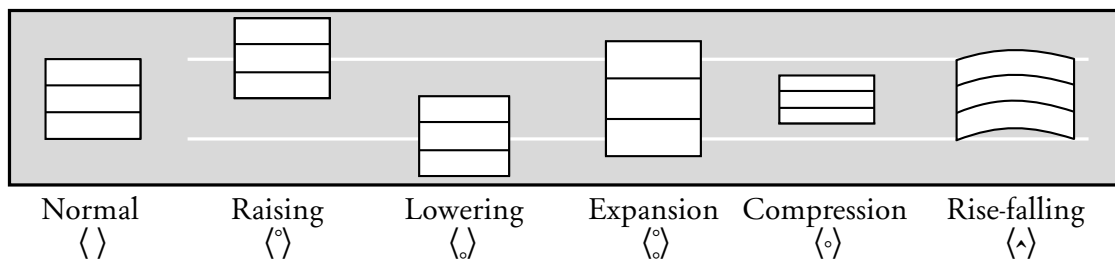
We can have simultaneously raising and lowering, too, which leads to *expansion* (both upwards and downwards), ⟨°˘⟩, by broadening the three pitch bands, as in cheerfulness: *You’re all welcome guests, folks!* – *Sàa minà-sañ, irasshài!* ⟨°˘⟩[-saa .mi.nɛ-sɛ.ŋ| ɰ.i.ɾɛɕ-ɕ:ɟɛi:].

Practically the opposite of expansion is *compression*, ⟨˘°⟩, which is typical of labels like sleepiness: *I’m feeling extremely sleepy* – *Meccha nemui nàa* ⟨˘°⟩[-mɛtɕ.tɕ:ɟɛ .nɛ-mɸi .na:a˙].

Notice that *nàa* can have a further paraphonic peculiarity added, rise-falling ⟨^⟩, which is the last pattern shown in fig 10.9: ⟨^⟩[-na:a˙].

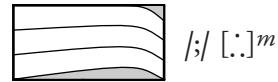
The same pattern is very frequently, and typically, used with exclamations like *Àa!*, *Èe!*, *Hoo!*, *Iya!*, *Yàa!* ⟨^⟩[-aa, -ɛɛ, -hɔɔ, -i-ɛ, -jaa:], also with a more or less evident lengthening, sometimes reaching even [V:V], although, in the conversations of 11, we will just show them as phonotonic elements.

fig 10.9. Paraphonic characteristics of pitch.



It is interesting to note that a similar (intonation) pattern is often used, in *meditative* (and even *colloquial*) pronunciation, for the *suspensive* tune /,| [˙], instead of neutral [˙].

fig 10.10. Suspensive tune variant, frequent in *mediatic* and *colloquial* pronunciation.



### Some kindlier tune usages

10.20. It is well known that Japanese people are very kind and polite towards other people (even if they would prefer not to do so, sometimes).

Well, this attitude can certainly lead to use (in addition to specific paraphonic peculiarities) different tunes than the more normal ones used in colloquial or familiar speech.

So, very often, instead of the conclusive tune, /./ [ː], we hear the continuative one, /,/ [ː] (which sounds less peremptory), or, at least, an attenuated /./ – ie /./ [ːː]. This is the case of the examples in § 10.10-11 (and others, of course).

Also a number of the sentences shown for paraphonic reasons in § 10.18 present this change.

We must know for certain that, for the same reason, partial questions (cf § 10.11), in addition to this change, can frequently present the following one: /ɛː . / → /ɛː ː / → /ɛː ː / → /ɛː ː / → /ɛː ː / → /ɛː ː / . ie the use of some attenuated tunes, including an attenuated inter rogative one.

However, this last can sound too kind, if not falsely kind, and little spontaneous. fig 10.11 shows the general tonogram for attenuated tunes, which are partially modified when forced into that path.

Unfortunately, even in the recordings of teaching courses, too often we happen to find the very use of /ɛː ː / also in *partial* questions (as untrained people generally do, as well, when asked to read some examples).

Let us conclude by saying that *total* questions without *ka?* must necessarily use /ɛː ː /, while –with *ka?*– we have many possibilities: /ɛː ː /, /ɛː ː /, /ɛː ː /, also /ɛː ː /, up to /ɛː . /, most typical of familiar speech.

fig 10.11. Tune attenuation, including possible /:/ [ːː] (cf fig 10.7).

