

Natural Phonetics & Tonetics

Articulatory, Auditory & Functional

Luciano Canepari

University of Venice, Italy

*Nihil nihilo –quippe– plenius,
nihil numinibus –hui– vanius,
nihil hominibus –heu– deterius.*

2007 (updated Dec. 2008)
LINCOM

Contents

p. VIII

Foreword

Natural Phonetics & Tonetics *Articulatory, auditory, & functional*

1	1. Prelude
2	Transcriptions
6	The contents of <i>NPT/HPb</i> (& of <i>HPr</i>)
12	Observations on phonetic terminology
17	2. Doing phonetics
22	Guide to the figures
31	Guide to different types of transcriptions
34	Transcribing by hand
35	3. Pronunciation & phonetics
38	The phonetic method
46	4. The phono-articulatory apparatus
51	The vocal folds
55	Resonators (5 phono-articulatory cavities)
58	The lips
59	5. The classification of sounds
63	6. A gradual approach
63	Vowels
65	Voicing
66	Consonants
67	Places of articulation
70	Manners of articulation
74	Prosodic elements
74	Stress
75	Sentence stress
76	Tones
77	Intonation
81	7. The official <i>IPA</i> & other notations
81	Official <i>IPA</i>
81	Consonants
87	Vowels
89	Prosodic indications & other diacritics
90	How come the <i>IPA</i> is not used by everyone?
93	Quick comparison between <i>offIPA</i> & <i>canIPA</i>

p. 96	The official revision of the <i>IPA</i> (1989-96): A missed reform
98	Official diacritics
99	Segmental diacritics
100	Suprasegmental diacritics
100	Official tones & word accents
101	About non- <i>IPA</i> alphabets
103	Comparison with the main non- <i>IPA</i> symbols used in Romance studies
105	From a couple of <i>IPA</i> to many different non- <i>IPA</i> 's
104	The phonetic alphabet of the <i>ALI</i> : Another example not to follow
107	Observations on the (non) ‹respect› of symbols
109	Hypostatization & ‹ <i>ip</i> astatization›
111	8. Vowels & vocoids
116	Other, less useful classifications
119	More about vocoids
128	<i>canIPA</i> vocoids
134	Articulatory practice
137	Diphthongs: one phoneme or two?
140	<i>canIPA</i> vocoids & correspondent <i>offIPA</i> symbols
142	9. Consonants & contoids (1)
147	Nasals
148	Stops
149	Constrictives (‹fricatives›)
152	Stop-strictives (‹affricates›)
155	Approximants
157	Trills, taps & flaps
158	Laterals
159	Memorizing
163	Articulatory practice
165	10. Consonants & contoids (2)
165	Table of the main <i>canIPA</i> contoids
165	<i>canIPA</i> contoids (displayed according to articulation manners)
169	Nasals
172	Stops
175	Stop-strictives (‹affricates›)
180	Constrictives (‹fricatives›)
183	Approximants
187	Trills, taps & flaps
189	Laterals
192	Comparisons between similar contoids
196	11. Phonic peculiarities
196	Intense (‹syllabic›) contoids
196	Coarticulation
197	Modifications
199	Variations
200	Contoids with particular offsets
202	Prenasalization
203	‹Aspiration›
205	Non-pulmonic consonants
205	Ejective consonants
206	Injective consonants
207	Dejective consonants (clicks)
210	Nasalization of vocoids

p. 212	Devoicing vocoids
212	Vocoids in unstressed syllables
213	Vocoids in singing
215	Abolition of the term (and concept of) ‹retroflexion›
215	Generic symbols (for phonic categories)
218	12. Microstructures
218	Syllables
219	Scale of syllabicity
222	Syllabification
224	Syllables & the speech chain
225	Length
227	Stress
231	Pitch & tones
233	Tonic practice
238	13. Macrostructures
238	Prominence
238	Rhythm & rhythm groups
239	Pauses
240	Pitch & intonation groups
240	Paragraph & text
241	Rate
241	Intonation
244	Intonation groups
244	Preintonemes
245	Intonemes
247	Questions
250	Intonemes modifications
254	Parentheses & quotations
255	Considerations on communicative ‹roles›
258	Considerations on intonation
259	Structures & generalizations
262	14. Superstructures
262	Paraphonics
263	Pitch
264	Other paraphonic elements
266	15. Phonosyntheses
272	16. Italy
307	17. Europe
359	18. Africa
378	19. Asia
414	20. Oceania
419	21. America
439	22. Dead languages
480	23. ‹Extraterrestrial›
482	<i>Utilizable bibliography</i>
490	<i>Index</i>
496	<i>Language index</i>

o. Foreword

o.1. Lay people are convinced that the pronunciation of a language and its official spelling are one and the same thing (in spite of some lamented incongruity). This impression derives from the fact that school –to be true– generally only worries about writing, and completely neglects and ignores pronunciation. Society does no better on the whole! Thus, people are irrationally terrified by spelling errors, to which they will be passively subjected without asking themselves any questions.

The belief of the (almost) ‹natural› correspondence between spelling and pronunciation is merely illusory. But, if spelling can be fairly homogeneous in every nation (at least as far as middle and upper education levels are concerned), this instead is not the case for pronunciation, which is generally more or less regionally marked.

o.2. Too often, people think that they cannot improve their pronunciation (or their verbal communication). At first, it is not always easy, nor is it evident –although, *afterwards*, it is quite obvious– that each person has a different pronunciation from other people’s, not so much –or not only– because of individual peculiarities of voice, due to personal *timbres* (which are determined by the somatic and temperamental characteristics of each person); but, most of all, because of the regional and social characteristics which every person ‹spontaneously› has (unless we have freed ourselves from them, by applying the *phonetic method*), since they have been acquired together with the language, as a part of language itself.

o.3. In general, this English version of the *Handbook* (originally written in Italian, and already published by Lincom Europa) maintains the original structuring, so that a wider readership may discover the *phonetic method*, although examples have been adapted to English, whenever possible.

However, we are preparing the book *English Pronunciations*, which will be structured from an English point of view. On the other hand, those who can fully master the phonetic method will surely take advantage of it, independently from their mother tongue.

o.4. Phonetics must not be ‹studied› unwillingly, or mnemonically, as if it were just a useless and thankless toil. On the contrary, it has to be ‹discovered›, while enjoying playing with sounds (and with words, sentences, and texts).

Although we are hardly aware of it, phonetics is always with us: it is *in us*. In fact, just as chemistry or physics already existed, independently from Man’s awareness or will; so is phonetics naturally inevitable, when we speak. Once the laws and principles of chemistry and physics have been discovered, these sciences can

be applied in many useful ways. Equally, if we learn to use the categories and principles of phonetics, we can succeed –with spontaneous naturalness– in *recognizing* the different sounds of our own language, including nuances. Later on, we can also manage to *recognize* the sounds of other languages, dialects, and accents. And this will become easier if we are able to apply correctly what phonetics offers freely, with no need of expensive equipments, and with no particular talent: it is sufficient (but necessary) to start to really <listen> to the sounds, it is not enough to simply <hear> them.

A very effective help, with this analysis of sounds, is provided by the use of phonetic symbols, which allow us to <see> the sounds and then to compare them, by reflecting on their similarities and differences.

Another way to <see> the sounds is using appropriate articulatory figures which, besides facilitating reciprocal comparisons, also activate our reflection of the movements that occur inside our mouth, when we just produce a given sound. Once we have started, it is surprising to notice that certain sound nuances correspond to given movements (although small). We really wonder why we did not realize such a simple and natural thing earlier. However, it is paramount that we manage to do that, albeit a little late.

0.5. Schools should endeavor to make this <miracle> come true. And, as a result, phonetics will certainly continue to give its useful benefits, in a spontaneous and natural way, whenever it is channeled appropriately, for studying foreign languages, as well as improving one's own national language. It would be sufficient to introduce the basic elements of phonetics, as a game, in the first three months of the first year of primary school, by means of a videotape, or a CD-ROM, or a multimedia interactive CD, appropriately organized, and with wall boards which show some figures (cf § 0.7) and the most important phonetic symbols, which should be rigorously written in red (and perhaps in [], or / /), in order to always distinguish them from <normal> spelling, following the *bi-alphabetical principle*, which is part of the phonetic method (of *Natural Phonetics*).

These materials would also be useful to compensate for the inevitable want of teachers (who have no clear notion at all about phonetics, since –certainly– they have not been prepared for this). But, children themselves would mostly obtain the best results, by amusing themselves while acquiring –without even being aware– the fundamental and necessary key to succeed in usefully separating the phonic level from the written level (as we have already seen during years of experimentations in primary schools). These <discoveries> would remain with the children, naturally interiorized, also when they grow up, being able to avoid making the same mistakes as their predecessors did (including teachers).

As a matter of fact, pronunciation is not a perfidious invention of some phonetics-mad phoneticians, but –indeed– it is the first manifestation of any language, which is certainly not only spelling, grammar, lexicon...

0.6. The real hump to get over is exactly *non-knowledge*. In fact, what we do not know is considered as completely <useless and impossible>. In this connection, the

absurd belief comes to mind that certain sounds of given languages are impossible to indicate or describe, since they could only be learned ‹by directly hearing them from a teacher's mouth› (as we still happen to read in quite a few grammars!).

Let us now see a simple type of transcription, which puts words (and sentences, as well) between slashes. *Phonemic* transcriptions show *phonemes* (or functional sounds) in relation to spelling. We can notice at once that stressed syllables are clearly indicated by the sign /'/, which precedes the consonantal and vocalic segments which form them.

In a word such as *concise*, the two *c*'s have two different phonic values: /kən'saɛs/; whereas the *c* and *s* in the last syllable have exactly the same value. Thus, it is obvious that a transcription is definitely more precise and less ambiguous than traditional spelling, where a *c* generally corresponds to /k/ when it is *not* followed by *i*, *e*, *y*, that is before *a*, *o*, *u*, or before a consonant; but *s* may be either /s/ or /z/ in different positions: *seas* /'si:z/, *cease* /'si:s/. Let us also consider *g*, which can either correspond to /g/ or /dʒ/: *get* /'gɛt/, *gem* /'dʒɛm/; *porgy* /'pɔ:ɹgi/, *orgy* /'ɔ:ɹdʒi/; while *.gif* can have either phoneme: /'gɪf, 'dʒɪf/.

Another difference, in comparison with current spelling, will be quite obvious when we consider the following examples, as well: *to live* /'lɪv/, *a live* /'læv/ *concert*, *to tie a bow* /'bɔʊ/, *to bow* /'bɔʊ/ *and scrape*, *a minute* /'mɪnɪt/, *in minute* /ɪn'mɪnɪt/ *detail*, *of no import* /'ɪmpɔ:ɹt/, *to import* /ɪm'pɔ:ɹt/.

0.7. This handbook of *Natural Phonetics & Tonetics* (NPT, which is the fully revised, amended, and updated edition of *A Handbook of Phonetics*, HPh, with additions and substitutions, henceforth indicated as NPT/HPh) provides everything which is needed for doing really useful phonetics. This means *natural* phonetics—that is *articulatory*, *auditory*, and *functional*—with all necessary explanations through articulatory figures for consonants: *orograms* (cf § 3.2.1) of all articulations (and some *palatograms* and *linguograms*); for vowels: *vocograms* (or vowel quadrilaterals) and *labiograms* (some of the last ones are useful for consonants as well). In addition, auditory figures are used for intonation and tones, by means of *tonograms*. When it is necessary or useful, some other types of figures are used, as we will see below.

We are speaking of *natural phonetics*, since—to put it into practice—nothing is needed except one's own personal abilities and interest for the sounds of *the world's languages*: both languages and dialects, and both accents and variants, as well. The only ‹external›—technological—instrument which is highly useful is a good tape recorder with good earphones, and with a quick, sharp pause button, that is mechanical. This button allows one to interrupt the listening of a recording (and to start it again at once), being able—with practice—to segment each sound, even a short one, into two or three parts, avoiding the running problems of so many electronic pause buttons, which distort and hide whole syllables. Such a tape recorder—for these purposes—is definitely superior to both a video recorder and a CD player, also (and mostly) when they are connected to a computer (with its slow and complex electronic commands).

On the other hand, all of us have learned our own *mother tongue* perfectly (ie a

dialect or a regional variant), even with no need of a tape recorder, since we were of the ideal age and had the necessary motivations to build up a language, in order to be able to communicate with those who were around us. If we did not have a language to use, we would have terrible practical, social, behavioral, and conceptual limitations. After baby-hood, it is more complicated to learn *another language*, or *more languages*, also because the new structures rarely coincide with those of our first language, and every kind of interference is always in wait, even between similar languages.

In this handbook, we will deal with about 1000 fundamental segmental sounds (ie vowels and consonants). Surely, they all exist, since the speakers of the different languages and dialects of the world learn and use them perfectly, although few specialists manage to perceive (and pronounce) them all adequately. But it is more than likely that some more are to be found, when further languages can be analyzed.

o.8. To produce transcriptions easily using *phonetic fonts*, *Macintosh* computers are the most suitable. With our own *Mac* and *Word 5.1*, and our own Italian extended keyboard (= *qzerty* and, above all, *jklmù!*, by carefully avoiding the *Italy Pro* soft keyboard) we can obtain at least four symbols for each key. For instance, using our own *Simon(ani* font, we immediately obtain [a, ɑ, ɐ, ɶ] (by simply typing on the *lower-* and *upper-case* keys, with the addition of the *option*, *option+shift*, and *option+caps* keys), for 221 total signs for each font element/style (ie roman, *italics*, **bold**, and **bold italics**). This is done without having to perform complex acrobatics to *insert symbols* with many clicks of the mouse (but non-Mac users will hardly be able to realize what this statement actually means, since it seems impossible –or even unimaginable– for those who are bound by the absurd limitations of common PC's).

Unfortunately, the *Windows* operation system (which has cribbed many good things from the Mac) still allows no more than about 100 signs directly from the keyboard, whereas a Mac can give up to 884 signs (for each complete font).

o.9. Obviously, many people have facilitated the realization of both *NPT/HPb* and *HPr*. As it is impossible to mention all of them, we want to thank those who have generously provided advice and suggestions, by reading and commenting one or more chapters of the 36 which form the *two* handbooks (whereas at the beginning only one book was expected to be sufficient). Some people have also indicated some bibliographical materials or provided some of the precious recordings we have used for direct analyses, either for some sections of our descriptions, or for some or many of our phonosyntheses. All advice and suggestions have been considered, obviously by fusing and uniforming them according to our theoretical and practical positions.

Therefore, we would particularly like to thank Filippo Tassetto, as well as Samuele Dovico, and Emanuele Saú; besides, Paola Barberis, Lidia Costamagna, María Magdalena De Raedemaeker, Floréal Molina, Renato Seibezzi, Maria A. Simionato, Alberto Venturi, and Daniele Vitali. In addition: Lulzim Ajasi, Piera

Buono, Alessandra Di Bonaventura, Alessandro Duranti, Elisa Fiorio, Fernando Grasso, Jaime Magos Guerrero, Nikodimos Idris, Vito Matranga, Renzo Miotti, Riccardo Mura, María Emilia Pandolfi, Renato Pinturo, Paolo Piras, Salvatore Riolo, Suzanne Ruelland, Hans Schiessler, Walter Schweikert, Patrizia Siniscalchi, Roberto Taddio, Alfio Torrisi, Geneviève, Yans. We do *not* thank, instead (but without naming names), those who did not keep their alluring promises to provide good recordings...

Lastly, we «conditionally» thank those who continue mentioning the... first editions of our books, although more up-to-date editions exist. For instance, the third edition of *Italiano standard e pronunce regionali* (1986) has 43% more material than its first edition (of 1980). Whereas the first edition of the *Manuale di pronuncia italiana* (1992) had little more than 400 pages, the second edition (of 1999) contains 282% more divided into two volumes of about 1200 total pages: the new *Manuale di pronuncia italiana* (proper, *M^aPI*, which contains a thorough and up-to-date description of regional pronunciations as well, although they are not explicitly mentioned, not even in the amended version of 2004). The second volume is the *Dizionario di pronuncia italiana* (*DⁱPI*).

To avoid further ambiguities, we would like everyone to know that the 60,000 forms given in the *Dizionario* are all chosen and listed for pronunciation purposes, having excluded those which cannot raise doubts, such as *cane, gatto, rifare, irrigidirsi, opportunità*... Equally, we have excluded any inflected forms which pose no problems, and also those which can be derived by the combination of *endings*, although with orthoepic relevance (but they are all dealt with in § 4, for a better and pondered usage and to save space).

Therefore, although it is difficult to calculate precisely, we can say that the «real» forms provided in the *DⁱPI* are certainly not less than 180,000...

Still more conditionally, we «thank» those who mention our works to demonstrate different (when not opposite) things, compared to what we have actually written or transcribed...

From a graphic point of view, *NPT/HP^h* is rather complex, both for the very many figures and for its transcriptions. *HP^r* has very many transcriptions (including about 75 of Aesop's passage *The North Wind and the Sun*), in addition to many figures. We have done our best to avoid mistakes and misprints...

o.10. It is important to remember that when we criticize pure phonologists or pure acoustic phoneticians, we do not mean that their work is of no use at all. Of course, they are free to do what they want (or what they can). We would just like young and unprepared readers not to be made to believe that doing descriptive phonetics is something out-of-date, useless, or unscientific.

On the contrary, it is often fundamental –and proper– to rebel against what some «cultural lobbies» want to make people think about many things, trying to eliminate any currents which happen to be different from theirs. This is true freedom indeed.

Finally, it is a must to fiercely rebel against those who do descriptive phonetics in a very generic and worthless way, by using approximate and ambiguous terminology and symbols, as well. However, those who only do phonology, or acoustic phonetics, limit themselves to simply *thinking* of sounds, or *seeing* them. In fact,

they prevent themselves from really *listen* to actual sounds. And, after all, sounds are *sounds*!

o.11. Lastly, we would like to end with a *call* –which is issued to all those who are interested in the five lines of phonetic research (that we will mention shortly)– to contact the present writer, in order to see whether we can carry out some kind of fruitful collaboration.

The five lines of research are:

(A) the description of *neutral* and *socio-regional accents* (of native speakers all over the world) for English, French, German, Spanish, Portuguese, Netherlandic [ie Dutch and Flemish], Greek (and, perhaps, other languages, such as Russian, Turkish, Arabic, Hindi, Chinese, Japanese). For Spanish alone, all the nations of Central and Southern America have to be thoroughly analyzed...

(B) the description of *foreign accents* of English, Italian, French, German, Spanish, Portuguese (in particular those of immigration, from eastern Europe; northern, western, and eastern Africa; southern Asia, from the West to the East);

(C) the *expansion* of the *phonosyntheses* to transform them into real descriptions, with all the necessary integrations and examples, as we have done in the 12 chapters of *HPr*;

(D) the *increase* of the number of *phonosyntheses* by adding other languages, especially from the most distant and still lesser known areas of the world;

(E) the preparation of *pronouncing dictionaries* using *IPA* (ie phonemic transcriptions, but with an appropriate phonetic introduction as well, using *canIPA*, with vocograms, orograms, tonograms, &c) to start with Spanish, Portuguese, Greek, Russian, Turkish, Arabic, Hindi, Chinese, Japanese, Vietnamese, Indonesian, &c. It would be sufficient for some native speakers, interested in the pronunciation of their own languages, to start the <enterprise>, guided by us, in order to continue alone. The most important thing is for people to begin to try hard. However, useless and confused <orthographic transcriptions> are to be avoided, although integrated by some diacritics or special signs, because they disguise any phonetic reality and are only misleading even for native speakers...

Lu@

University of Venice, Italy
Dept. of Speech Sciences
Phonetics and Phonology
19 January 2007

Natural phonetics is dealt with in our website, for updates, reflections, anticipations, and to spread the importance of *articulatory*, *auditory*, and *functional* phonetics, that is *natural phonetics*. The site is also meant to <talk> with interested readers, through e-mail, for the *call* in § o.11, too (even for curiosities or errata, and other things): <http://venus.unive.it/canipa/> – for *canIPA*.

1. Prelude

1.1. Unfortunately, it is a well-known fact that the traditional orthographies of the languages of the world are, to a greater or lesser degree, deficient in representing the pronunciation of these languages clearly and reliably. As a matter of fact, orthographic symbols do not correspond in a one-to-one manner with distinctive sound units or, at any rate, they do not correspond well now. Even languages like English and French, which have a very high level of *non-correspondence* between writing and pronunciation (especially, but not exclusively, in regard to the vowels), had, centuries ago, a much greater degree of correspondence. Their situation was roughly similar to that of modern Italian, or Spanish, which are in any case far from perfection (which would entail rendering the pronunciation, including the position of stress, completely deducible from the orthography).

The lack of correspondence is primarily due to the fact that the pronunciation of languages changes with the passage of time (and in the case of English and French, it has changed quite a bit in the last half millennium). On the other hand, writing systems have always been extremely *conservative*, and in the case of the European national languages have remained more or less fixed since the time of the first great literary works in those languages.

1.2. For example, in the time of Chaucer (14th century), the English words *seed*, *name*, and *night*, which today are pronounced /'si:ɪd, 'neɪm, 'naɪt/ ['sri:dʒ, 'neɪm, 'naɪt], were then pronounced /'se:ɪd, 'na:mə, 'niçt/. English has reached extreme cases of non-correspondence, including ones where the same phoneme is represented in many different ways. For example, /i:/ in: *green*, *eve*, *mean*, *field*, *seize*, *key*, *police*, *people*, *aeon*, *quay*. Conversely, a single letter combination can represent various pronunciations: for example, *gh* in: *ghost* /g/, *hough* /k/, *hiccough* /p/, *enough* /f/, *Edinburgh* /ə/, *though* /θ/ (‹zero›).

As was mentioned above, the Italian writing system does not arrive at excesses like those exemplified above. However, it does not give much information on the pronunciation of *e*, *o* (/e, ε; o, ɔ/), nor on that of *s* (/s, z/), or *z* (/ts, dz/). It is also deficient in showing where words are stressed, and this confusion is the cause of notable oscillations and ‹idiosyncratic› stress positions. In Italian (as in English and French), there are, moreover, inconvenient cases of etymological ‹conservation› of ‹silent letters› (often inserted by purists even where they had never been pronounced, such as English *doubt* /'daʊt/, once written *dout*, but with the *b* added later because it had been present in Latin – *dubitare*). English *science* (from Latin *scientia*[m]) is /'saɪəns/, while in Italian we have *scienza* /'ʃɛntsa/.

Transcriptions

1.3. In order to analyze the pronunciation of a language, for learning and teaching purposes, it is necessary to use two different types of transcriptions: *phonetic* and *phonemic* (or *phonological*).

It is important for both of these to be derived more from *interlinguistic* strategies (considering several languages) than from *intralinguistic* considerations (working within just one language). As a matter of fact, an INTRALINGUISTIC transcription is designed only for the needs of a particular language, and usually, only for native speakers of that language. The result is a transcription which is merely phonemic and not convenient for comparisons with other languages. This type of transcription is, however, completely legitimate when there is no intention of relating the transcriptions to the pronunciation of other languages. In these cases, it is even possible to use extremely vague symbols as long as each phoneme is distinguished from all the others.

Therefore, in theory, it could be acceptable to represent Italian diphthongs such as [ai, au] by </aj, aw/>, while representing the sequences [ja, wa] with </iá, uá/>. In English, it would be possible to transcribe the diphthongs [ɪi, eɪ, əə, ɔə, aɔ, ɜɔ/σɔ, μu/υu] as </i:, e:, ai, oi, au, ɔ: u:/> (or even </ij, ej, aj, oj, aw, ow, uw/>, as not a few do). According to the <logic> of this last case, even the German diphthongs ([ae, ao, ɔʏ]) could be represented by </aj, aw, oj/> without problems.

However, it immediately becomes clear that for descriptive and teaching purposes, transcriptions such as </aj, aw/> are definitely unfaithful and also rather misleading, respect to different realizations such as the Spanish or Italian [ai, au] (with the second vocalic elements in each diphthong quite high), the English [əə, aɔ] (with the second elements much lower and centralized), and the German [ae, ao] (with the second elements lower as well, but not centralized).

In the case of the German </oj/>, there would be the additional problem of ignoring or hiding the fact that even the second element, in neutral pronunciation, is rounded (and also not completely high nor front): [ɔʏ]. In English, we should hope that one day not only [eɪ, ɜɔ/σɔ], but also [ɪi, μu/υu], are diphthongs will once and for all be evident to everyone (the variants after the slash are American).

INTERPHONEMIC transcriptions, instead, seek to show adequately the characteristics of the single languages, although within a restricted inventory of symbols, which, however, takes into account the similarities and differences of the particular languages. Thus, in Spanish or Italian, we could certainly have /ai, au/, but in English, /aɛ, aɔ/, and in German, /ae, ao/ (a careful analysis of the vocograms in ¶ 2-3 & 5 of this manual's twin volume, [*ie HPr – A Handbook of Pronunciation*, in the bibliography] will help to explain the difference between the symbols used in English and in German). Therefore, completing the series mentioned above, we have: English, /ɪi, eɪ, ɔɛ, ɔʊ, υu/; German, /ɔʏ/.

1.4. The most useful *phonetic transcriptions* are TAXOPHONIC – ie, which use all available symbols for phones (vocoids and contoids) and for prosodic ele-

ments, in an effort to show clearly every necessary nuance. Only in this way is it possible to make truly useful comparisons between different pronunciations (of different languages, or different pronunciations of the same language depending upon geographical or social differences). Otherwise, everything becomes vague and consequently less productive – one can believe the result is useful work, but more often this belief is only a mirage.

For teaching and learning pronunciation well, simplifying matters excessively (in order to put things on the students' level) is also detrimental. For example, when English phoneticians continue to use [ə] even in more precise transcriptions (often with extensive use of diacritics) for every case of /ə/ (and ⟨/əʊ/⟩), they lose the opportunity to show *reality*. Thus the neutral British pronunciation of *to go*, *the man*, *further* is often given as ⟨[tə'gəʊ, ðə'mæn, 'fɜːðə] (or ['fəːðə])⟩, instead of the actual [tʰu'gɜːʊ, ðə'mæ'n, 'fɜːðɐ] (cf § 2 in *HPv*). The same occurs with ⟨[i, uː]⟩ instead of [i, ʊ/ʊ] (mentioned in various parts of the book).

Therefore, particularly for those who deal with several languages at once, but also for those who use a single foreign language, the best choice is *interphonemic* transcriptions, since these ones are the least arbitrary in their use of symbols. These transcriptions do not level out phonetic reality, rather, they emphasize similarities and differences, thereby representing much more faithfully the relationships between different languages and pronunciations. For example, in Japanese it is definitely better to use /ɯ/ rather than ⟨/u/⟩, even though there is no chance of confusing them (since Japanese has no /u/) – the use of /ɯ/ makes the differences between Japanese and other languages quite clear.

DIAPHONEMIC transcriptions are also important. They are based on an *interphonemic* foundation (among different languages), although it is also possible to work with an *intrapphonemic* foundation (within one single language), with, however, the defects implicit therein. DIAPHONEMES (from *dia-* ⟨distinction⟩) are a fundamental tool for transcribing, simultaneously, partially different accents of the same language, since they remind the reader of differences of pronunciation, while making it possible to show, in a single transcription type, systematic variations (without having to repeat the same words and change the symbols in the parts which differ from a phonetic point of view). In this way, for example, /ʊ, ɔʊ/ can be used to represent at the same time [ʊ, ɜʊ] (British), and [ʊ, ɔʊ] (American): *two* /'tʊʊ/ [tʰɯʊ, tʰɔʊ], *go* /'gɔʊ/ [gɜʊ, gɔʊ]. Similarly, /æ/ for [ɑː] (Br.) and [æ] (Am.): *last* /'læst/ [lɑːst, læst], or /ɒ/ for [ɒ] (Br.) and [ɔː] (Am.): *lost* /'lɒst/ [lɒst, lɔːst]; or, /ɿ, ɿ/ in /'lɛɿɿ/ for [lɛɿɿ] (Br.) or [lɛɿɿ] (Am.).

1.5. No matter what type of phonemic transcription is used, in order to be able to reach a phonetic transcription (and thus an adequate pronunciation), it is necessary to consider each distinctive element – each phoneme – as *one* out of a certain number of fixed points in the *phonemic space* (whether vocalic or consonantal) of a given language. Each of these points, or elements, is necessarily different from all the others, and each is also different from the ⟨zero occurrence⟩. So /a/ is different from /i/, or from /o/, &c – but it is also different from /∅/, ⟨zero⟩).

Even the opposite process, namely the decodification of a spoken message, proceeds according to the following principles. Every *phone* of a given utterance –also depending on both the global meaning and the speaker's phonic system– is assigned to a certain *phoneme* (as its actualization), which is then located in a fixed *phonemic space* (within the phonemic system of each language).

Even if one is unfamiliar with the speaker, when one hears the utterance *this is my house* /ðɪsɪz 'maɪ 'haʊs/ beginning with *this is* pronounced as [ˈvɪsɪz], or [ˈdɪs(ɪz)], or [ˈðɪsɪz], or [ˈðɪsɛz], or [ˈðɛsɛz], or [ˈðɪsɪz], it is already possible to anticipate that *my house* will be pronounced as [ˈmɒɹ 'ɛɪs] (Cockney), or [ˈmɑːɑ 'haʊs] (Black American), or [ˈmɑːə 'hæʊs] (New Zealand), or [ˈmɑːɪ 'hæʊs] (Australian), or [ˈmʌə 'hæʊs] (mediatic American), or [ˈmɑːə 'haʊs] (neutral British). For now, the use of different symbols (which will be explained systematically later) should be sufficient to make clear that there are notable differences among the various pronunciations, which should definitely not be ignored. Thus, it is important to emphasize the difference between *phonemes*, *phones*, and *sounds*.

PHONEMES have a *distinctive value*, within a given language, because they are capable of changing the meaning of words, as in English: *bit* /bɪt/ and *bet* /bɛt/, or *beat* /bi:t/ and *boot* /bu:t/, or *thin* /θɪn/ and *thing* /θɪŋ/. A *phoneme*, therefore, has the function of distinguishing itself from all the other phonemes of a given language, to be different from the others. Or in other words, its function is *not* to be what the others are: it is purely *form*.

PHONES, on the other hand, have an *identification value*, because they help to characterize the pronunciation of a language, or of several languages. They perform this task through segments which are more or less typical and recognizable, beyond simple phonemic representations. Different accents of English, of course, realize the phoneme /ʌ/, as in *shut*, in different ways – neutral British English [ɐ], neutral American English [ʌ], mediatic British English [a], mediatic American English [ɚ], Northern British English [ɜ] (when not exactly like /ʊ/ in *put*), ⟨general⟩ and ⟨broad⟩ Australian English [A].

The function of a *phone* is to maintain consistency between the elements of a given pronunciation: it is *substance*.

Finally, the value of SOUNDS lies exclusively in *transmission*, their function being to enable human communication through sound waves. Therefore, a *sound* is a single emission, in practice unrepeatable even by the same person. It can oscillate quite a bit, often producing quite different realizations: it is *matter*. It should always be kept in mind that at different moments, both phonic production and perception can vary to a greater or lesser extent.

As will be seen in § 2.4, one way to allude to the fact that actual sounds are always a bit different would be to represent them with different fonts: o, o, o, o...

In conclusion, *many* different –yet similar– sounds constitute a single *phone*. Then, in the context of a particular language, several phones, not wholly identical (but with the alternation governed by fixed and systematic rules, which can and should be discovered and then explained simply and completely), constitute a *phoneme*. The phones referred to by a given phoneme are called *taxophones* (or combinatory phones, or ⟨allophones⟩ [a more ambiguous and less advisable

term, since it implies modifications not necessarily due to combination, but simply any sort of difference, for any reason, whether general or random]).

Sometimes, in the course of listening to recorded materials, certain sounds still need to be disregarded in the process of constructing the phonetic inventory of that language. This is because it is possible for single speakers to occasionally produce sounds presenting abnormal deviations, whereas it is essential to consider mainly what is more typical and frequent.

Therefore, actual sounds are practically infinite, and phonetics and phonology would be decidedly complex if it were not possible to rely on the *systematicity* of phones. As a matter of fact, the set of phones is the result of a past classification and structuring of sounds (abstracting from the unrepeatability of the same sound). This set allows us to reach recognizable types, which can in turn be represented through precise phonetic symbols, regardless of any particular language.

Thus, what is similar in different languages and dialects is realized by phones and their symbols, which make it possible to compare different languages (an essential point in order to be able to learn and describe those languages). Therefore, the phonemic representation of single languages necessarily uses a selection of (phonemic) symbols, with distinctive functions and purposes, even if, of course, some symbols may be the same but with rather different phonetic values.

Thus plain phonemic symbols (generally chosen among the most common, as the official *IPA* ones, *ie International Phonetic Alphabet*, or *offIPA*), do not represent pronunciation exactly, but rather the *relationships* between the phonemes of a given language. They are useful for the specific purpose of keeping the current *writing* system distinct from the *phonic level*. Therefore, phonemic transcriptions make it possible to avoid interference deriving from *not knowing* orthographic *rules*, or from the *inconsistency* of these rules.

However the exact pronunciation, for learning and teaching purposes, can only be shown with the careful, consistent, and systematic (because *normalized*) transcriptions of a phonetic alphabet such as *canIPA*.

1.6. It will be useful to reflect further on these matters and to give some examples, so that the purposes and importance of each kind of transcription will be more apparent. As a matter of fact, we can use *phonemic*, *tonemic*, or also *phonotonemic* transcriptions (*ie emic* transcriptions), giving only functional elements), as well as *phonetic*, *tonetic*, or *phonotonetic* (*ie etic* transcriptions), giving also contextual, or taxophonic, variants, representing actual pronunciation).

The examples given in the preceding section help to illustrate the difference between *phonemic* transcription: /ðɪsɪz 'maɛ 'haʊs/, showing only phonemes, and *phonotonetic* transcription [ˈðɪsɪz 'maɛ 'haʊs.] (neutral British), which shows taxophones and intonation as well. Thus a phonotonetic transcription is at the greatest distance from a simple phonotonemic one, since it gives all of the details necessary for a decent rendering of the pronunciation.

There is still another kind of transcription, which adds the possibility of showing *PARAPHONIC* elements. These are indicated at the beginning of a string of text enclosed within angle brackets. Thus, the example [⟨ˈzʰ⟩ ɪˈðɪsɪz 'maɛ 'haʊs.] is

uttered with *insistence*, resulting from an enunciation which is slower ⟨⟩, more rhythmic ⟨≈⟩, and more energetic or intense ⟨"⟩ than normal. However, these particulars are ⟨refinements⟩ (although they are hardly superfluous), which can be added to complete the description of an enunciation, by showing attitudes and states of mind.

1.7. The *phonotonemic* transcription, with neutral pronunciation, is: /ðɪsɪz 'maɛ 'haɔs./, and one important point to keep in mind from the beginning is that a stress mark, in *phonemic* transcription (/'/), and in *phonetic* transcription ([']), indicates only a greater prominence on that syllable, with respect to other syllables nearby. It has no relevance to the problem of denoting the pitch of a syllable, which should be determined from the part of a text describing that particular language's intonation.

In the books which do not completely ignore it, *intonation* is usually treated after vowels, consonants, and stress (and other prosodic characteristics, such as length) – this is due to the greater difficulties involved in describing intonation. However, intonation should not be ignored, or relegated to the end in *teaching*, since it is inseparable –in actual language– from the other elements.

Therefore, phonetic transcriptions, in the strict sense of the word, force reality a bit by indicating it only partially; in fact, it is as if they indicated –in any case– a continuative intoneme, as in [ðɪsɪz-], ['maɛ-], ['haɔs-]. (The *dot* at mid height indicates the presence of intonation, even though of an unmarked type.)

More or less the same thing is done to indicate tones (as in languages such as Chinese, for instance); but everything necessary to distinguish the tones in question is shown, giving them in their ⟨citation forms⟩, the same forms which appear on tapes for learners. Demonstration forms have a pitch movement very similar to the continuative intoneme, but without the compression commonly associated to this intoneme (cf § 13.33 & f 13.9). For example, in Mandarin Chinese, we have: *gū* /'ku/ ['ku].

However, in *tonetic* and *phonotonetic* transcriptions, the notation ['] represents not only stress but *also* a mid-range pitch (both in the context of tones and in that of intonation) – it is thus in opposition with other signs such as [ˉ], [ˊ], [ˋ], [ˌ], &c. In tone languages, also *phonotonemic* transcriptions indicate pitch, of course. For example, in Yoruba (an African language), we find: *kò* /'kɔ/ ['kɔ] ⟨to write/to sing⟩ with mid tone, in opposition to: *kò* /_kɔ/ [_kɔ] ⟨to refuse⟩, *kó* /ˊkɔ/ [ˊkɔ] ⟨to teach/to learn⟩, with low or high ton(em)es respectively.

The contents of *NPT/HPb* (& of *HPr*)

1.8. *NPT/HPb* is composed of two different parts, each of which completes the other, so that together they work towards a general understanding of the subject. Adding *HPr*, as well, which was meant to be the central part of the original *NPT/HPb*, this knowledge will be even more complete, particularly by helping to see how the pronunciation of languages can be addressed practically, descrip-

tively, and profoundly as well.

The *first part* constitutes the necessary *introduction*, both for *practical* and *theoretical* aspects. It has been organized *progressively*, beginning with what is simpler (ie the indispensable basic knowledge), and moving towards what is more complex (ie the details necessary for specialization). The treatment will be consistently general, using individual languages to provide examples.

The phonosyntheses (G 15-21) and the descriptions in *HPr* could also function as a more practical introduction, by placing the reader in contact with languages and dialects which are already familiar. Here, in fact, it is possible to revisit and gain better acquaintance with the vowels, consonants, and intonations (not to mention tonemes, when these occur, which are realized by tones and taxotones, ie context-dependent variants of tones) of languages and dialects which we already know.

1.9. The *second part* (G 15-23) is a store of *phonotonic* (and *phonotonic*) *information*, which will not fail to suscite the attention and curiosity of those who are very fond of and interested in the *phonetics science*. This science is not just a part of phonology (as it is still considered even by some linguists), but rather comprehends everything related to the sounds of language, including the phonological aspect. Phonology could therefore legitimately be called *functional phonetics*, whereas it would not make sense to speak of <articulatory>, or <auditory>, or <intonational> phonology, since the necessary conditions for constituting a subject of study and analysis would be lacking.

Indeed, *phonemes* are merely functional abstractions (ie mere structured *form*), and they would be completely intangible and unusable without *phones*, which realize them (as *substance*, also structured), through the concrete *sounds* of each particular language (as mere *matter* – amorphous, fluid, oscillating, and variable). Thus we repeat that phonology is only a part of phonetics, and not vice versa. On the other hand, phonology without phonetics makes little sense, save only perhaps on an abstract or theoretical level, without much or any practical usefulness. There is also little sense in carrying out pure acoustic phonetics, without any link to phonemes.

Therefore the phonosyntheses offer rapid summary indications on the *V* (ie vowels), *C* (ie consonants), and *T* (ie tonetics: intonation and ton[em]es when these last exist), obtained directly from audio recordings – *not* simply copied from second-hand sources, although we have naturally consulted the scientific (and not so scientific) works of others. The recordings have been analyzed by the same person who has produced *NPT/HPb* (and *HPr*). This constitutes a strong indication of consistency and overall view, even though –due to restrictions of space and time– the information is given briefly and concisely.

The function of the phonosyntheses is not limited to this, but includes two fundamental purposes: to provide information and phonotonic (and phonotonic) tools, both on practical and theoretical levels.

It is definitely useful to reflect upon the structures given, to make interesting comparisons between different languages, and to explore the richness of various

phonic systems as well. These thoughts can be profitably directed towards descriptive, comparative, contrastive, and teaching purposes. Merely looking over the phonosynthesis of a particular language can enable the reader to make important predictions concerning the problems that will most likely be encountered both by people trying to learn that language, as well as by native speakers of that language who seek to learn others.

1.10. The phonosyntheses are also useful because they make available –for researchers and all others interested in the subject– reliable frameworks for describing the pronunciation of 350 languages. These languages have up till now been described vaguely, superficially, or incompletely, leaving aside cases where the descriptions have been simply incorrect. Among these 350 languages (counting the 12 languages of *HPr*, together with 30 variants [also from different continents]), we have 63 *Italian* dialects (not all of which are Romance tongues), 79 *European* languages (including several dialects), 25 *African* languages, 58 *Asian* languages, 6 languages of *Oceania*, and 31 *American* languages. There are moreover 72 *dead* languages, and at the end we have included the description of a possible <inter-linguistic> (and <panchronic>, cf \mathfrak{G} 23) *extraterrestrial* language... (We are working actively on other languages and variants.)

All of this could be helpful in preparing more detailed descriptions such as those in *HPr*, or even full pronunciation handbooks and pronouncing dictionaries (such as *Manuale di pronuncia italiana* <A Handbook of Italian Pronunciation> and *e Dizionario di pronuncia italiana* <An Italian Pronouncing Dictionary> of the present writer).

It would also be possible to use these phonosyntheses to produce brief, but accurate, descriptions of a given language's pronunciation, to be inserted at the beginning of a grammar or a dictionary. The purpose would be to substitute, at long last, the confusing, disappointing, and often misleading <pronunciation guides>, which typically give only the presumed phonic correspondents of the graphemes used in a language or dialect (giving the disastrous results with which we are all familiar).

We should reserve a few words for the special case of the 72 dead languages included in \mathfrak{G} 22, since obviously it has not been possible to <listen> to them. We have therefore been limited to reconstructions based upon the works of experts in those fields, aided, however, by the *direct* –practical and theoretical– *experience* provided by the study of all the other phonic systems treated in \mathfrak{G} 16-21 and in *HPr*. In this manner, it has been possible to perceive, as if from the inside, the dynamics and the mechanisms of these languages as well, almost to the point of seeming to <hear> them.

1.11. It seems useful to anticipate more explicitly and systematically the contents of *HPr*, which is about *applied phonetics*, since it gives accurate and fairly extended descriptions of 12 *languages* (ie English, Italian, French, German, Spanish, Portuguese, Russian, Arabic, Hindi, Chinese, Japanese, and Esperanto), putting into practice what is explained here. We thus include the crucial *phonemic*

component –constituted by *functional phonetics*– which is realized with *phonemic transcriptions* given together with *phonetic transcriptions*, so as to show the relationship between *phonemes* and *phones*.

Together with these two types of notation, the *official spelling* of each language is provided (or a *transliteration* in the case of the four Asian languages). Spelling is given last, so that it will interfere as little as possible with the efficacy of the *phonetic method* and the associated transcriptions – thus the reader will come upon the spelling only after having fixed in mind the phonic structures. As will be seen, pronunciation variants are also given, and these will be worth studying.

Natural phonetics is an *artistic science*, and the best way to extend its rich potential concretely (after the necessary fundamental premises) is through applying its techniques and knowledge to languages which are in great demand for *teaching–learning*. In fact, these languages can be taught to others, or they can be self-taught (while remembering that, in order to teach adequately, it is necessary to have learned well first).

Therefore, it is best to begin the course of phonetic analysis with one's own language. The language one speaks is frequently a regional variant, more or less strongly marked. Consequently, it is important to understand how *neutral pronunciation* operates and how it differs from *one's own pronunciation*. This method helps to build consciousness of one's own speech patterns. These patterns can then be compared with those of the neutral pronunciation, which exists in every language of culture, even though speakers and society rarely have a consistent and precise idea of what it is.

The reader should note that we use the term *neutral* to mean something different from ⟨standard⟩ – in fact, at this point in time, ⟨standard⟩ is used with a meaning quite different from what we mean by *neutral* pronunciation. As a matter of fact, ⟨standard⟩ is very often used to mean ⟨almost standard⟩, in the sense of lacking obvious regional and social characteristics. The pronunciation referred to might actually be a mediatic one (typically spread by television), rather than the true neutral one, which is clearly defined and represents a conscious (and normally, professional) acquisition. We also avoid the term ⟨standard⟩ because too often it is used excessively, for example in speaking of ⟨standard London English⟩ or ⟨standard New York English⟩!

1.12. In every society, at least when there is a written language and literary production, a *neutral pronunciation* exists, which is the pronunciation used by *professionals in public speaking* (particularly actors, presenters, and announcers of high quality). The Danish linguist Otto Jespersen declared that *the ⟨best⟩ language* is spoken when the social and regional origin of the speaker are least obvious. Neutral pronunciation brings this ⟨miracle⟩ to pass, sometimes even to the point of making the identification of social and regional origin completely impossible.

The ⟨supporters⟩ of regional pronunciations, who encourage individuals to maintain their ⟨genuine⟩ pronunciation (even when they carry out ⟨public⟩ professions), deceive –themselves as well– because of their lack of success in better-

ing their own pronunciation (a task which would [have] require[d] commitment and persistence).

In other cases, they fail to understand the importance of being able to free themselves from a heavy burden (often extremely distasteful to the individual in question, at least on an unconscious level), all the while without renouncing their origins or identity. In fact, regional origins can be exhibited more effectively by alternating between one's dialect and the national language (competently, and above all by choice). The least preferable approach is to remain prisoners of an uncontrollable hybrid, which is neither the official language nor a local dialect. In Italy and the German-speaking countries, dialects differ largely not only in terms of pronunciation but also of grammar and vocabulary too; in fact *dialect* –rigorously speaking– refers (or should refer) to these other differences, while *accent* predominantly involves differences in pronunciation.

In the other 12 chapters of *HPr*, therefore, we will apply the *phonetic method*, which consists of a careful and rigorous comparison of the *phonemes*, *phones*, and *intonation* of the languages described, by listening to good audio recordings, using the two types of transcription (*etic* and *emic*), analyzing the *phonetic figures*, and showing *commitment*. The last quality is clearly indispensable: it is not enough simply to *want to know* – a *sustained effort to learn* according to an effective method is also required.

1.13. The approach we will use is a middle road between two irreconcilable extremes, that of abstraction and that of empiricism. Excessive abstraction would lead to a thoroughly theoretical formalization, entirely lacking in practical applications and connections with other fields. On the other hand, excessive empiricism would make everything so specific and particular as to render vain any possibility of generalization or normalization.

It is important not to slide into fascination with pure theory, just as it is important not to rely exclusively on the results given by computers. In the first case, one would end up by submitting to concepts like ⟨ATR⟩, ⟨VOT⟩, and ⟨downdrift⟩, in the fields of vowels, consonants, and intonation respectively.

As a matter of fact, ATR (ie [an ⟨articulatory setting⟩ with] *advanced tongue root*) is better explained by looking carefully at the relative vocograms. In African languages such as Igbo and Somali, in word formation, there are two series of vowels (as can be seen in the phonosyntheses in § 18.7 & § 18.22). The first series has dorsal positions which are conspicuously *fronted and raised*, thus contrasting with the *backed and lowered* positions of the second series (cf § 1.14).

As for VOT (ie *voice onset time*, the time taken for voicing [ie vibration of the vocal folds] to begin after a consonant), here, too, there is no need for excessive formalization. Actually, sequences such as [kha, ga] are clearly constituted by [k]+[h]+[a], and [g]+[a]; the first contains a true segment (ie a laryngeal approximant, [h], inevitably colored by varied and complex coarticulation phenomena with the two surrounding segments). The essential point is to see how much this segment is evident (as ⟨aspiration⟩), in order to decide whether to represent it in a transcription, or not to, in the cases where it is rather short and elusive. In any

case, ⟨voiced aspiration⟩ exists too (or, better, sequences of [C]+[h]), as in Hindi: *daan* /'daan/ ['daan] and *dhaan* /'dhaan/ ['dhaan], which would be quite difficult to reintegrate into the theory of VOT. In the case at hand, the supporters of the theory have often been forced to ⟨invent⟩ transcriptions such as ⟨['ḍaan]⟩ (for ['dʰi-]), together with ['th-], obviously for /th, dh/. The same is true for the possible devoicing of [ḡ] (referring to [g]+[a] above): if it is perceivable, we have to mark it. See the observation in § 10.13 as well, and the possibility of choosing between ten symbols, including ⟨zero⟩, for ten different realities.

1.14. Similar phenomena, passing from one phone to another, are hardly exceptional – in fact, they are actually the norm, given the extremely different articulatory settings of the phones which can come together in each language. It is, of course, true that instruments notice ⟨inconvenient cases⟩ of this type in great abundance. However, in so doing, they only reveal their inferiority to the ear of the native speaker and of the professional phonetician. As a matter of fact, while the instruments do not normally filter and separate the data, a trained ear manages to compensate well for unimportant variations, thus succeeding in recognizing and classifying the important functional attributes, without misleading complications.

The distinction between ⟨stress-timing⟩ and ⟨syllable-timing⟩ is an extreme occurring only in pure theory, since no language effectively belongs completely to one category or to the other. The term ⟨isochrony⟩ hints at the idea that languages contain rhythmic sequences of constant length, depending upon stresses or syllables.

However, not even English is systematically stress-timed, just as Spanish or Italian is not purely syllable-timed. A good phonotonic transcription succeeds in describing languages precisely and in showing peculiarities concretely much better than other overly rigid or overly theoretical attempts.

We will not treat explicitly *morphological* phenomena, which are generally shown in the orthography and are regularly described in grammars. Such cases include ⟨vowel harmony⟩, which means that, in certain languages (such as: Hungarian, Finnish, Turkish, Igbo, Somali, Telugu), within genuine native lexical forms (while, in loanwords, of course, rules are different) only V from particular categories or definite groups may occur. Among these we find: front *vs* back V, or unrounded *vs* rounded V, or high *vs* low V (or even raised [and] fronted *vs* lowered [and] backed V, as in Igbo and Somali, cf § 1.13). At times, the same vowel belongs to more than one category, and this can make things more complicated.

More rarely, there can also be ⟨consonant harmony⟩, as in Basque, where whole words may contain, for instance, only apical *vs* laminal C.

1.15. It is very important to attain greater precision in identifying and representing the true realizations of each phoneme. This task therefore implies the use of phonetic *transcriptions* with the aid of a symbol inventory endowed with a sufficient number of taxophones, so as to be able to represent the actual phonet-

ic reality, and not merely what this reality is assumed to be.

In this way, the ‹articulatory basis›, ie the collective phono-tonetic habits inherent in each language, is implicitly made evident, without the need for specific and complicated efforts (more targeted and burdensome, but –normally– leading to less satisfactory results). Thus it will not be necessary to transcribe the partial nasalization each time in cases such as *thanking* [ˈθæŋkɪŋ] ([ˈθæŋkɪŋ]), unless it should become more noteworthy, in which case it could be transcribed as [ˈθæŋkɪŋ̃]; the same holds true of words like *none* [ˈnɒn:, ˈnʌn:] ([ˈnɒn:, ˈnʌn:]), if pronounced differently from [ˈnɛn:, ˈnʌn:]. It will be more than sufficient to observe, once and for all, that a light nasalization in such cases is practically inevitable, in contact with following [N] segments.

In the case of *intonation*, mechanical instruments are capable of extracting and measuring single characteristics (generally through a series of separate processes), but they end up confusedly mixing up the various components. Expert human perception, on the other hand, can sift out superfluous elements and concentrate on what is truly important.

For these reasons, it is impossible to accept without reserves the raw melodic curves produced by acoustic methods. It is necessary instead to filter them in imitation of the process employed by the human ear, which manages to compensate in an extremely effective way for the many irregularities, whether objective or incidental, which can occur. The results are then standardized in tonograms, by taking a sort of average of large numbers of utterances.

The machines, instead, merely provide a graph for each event, comprising inevitably every divergence which is present. These divergences should instead be rationalized, interpreted, and normalized – not to a prefabricated standard, but one which is derived from careful study. The ear succeeds in correctly adjusting for overly high and low peaks, as well as for overly abrupt movements. These include the movement from high to low tones, with the result that the theory of *downdrift*, postulated for certain tone languages, becomes unnecessary. In fact, the movements in those language are not aberrant from the point of view of native or expert perception, and can thus be reintegrated in the norm without creating special difficulties.

Observations on phonetic terminology

1.16. In the field of *terminology*, as well, scientific rigor is a great help, while the vagueness and lack of clarity of certain obsolete traditions with little scientific foundation is decidedly negative. On every page of *NPT/HP* (& *HPr*, too) this requirement is addressed continually. Indeed, everything becomes simpler and easier to understand when the technical terms are clear and intuitive (even more so than in normal speech). For example, *tonic* should refer only to *tone* –pitch– and not to *stress*. A *diphthong* should contain only vocalic elements, such as in [ai], not vocalic and consonantal elements together, as in [ja]. Were it otherwise, it would follow that [la, ˈma, ˈsa] would be diphthongs as well – in fact,

while [a, i, u] are vocalic elements, [j, l, m, s, t, r, h] are consonants. Thus, contrary to the opinion broadly promulgated by grammarians, Spanish or Italian [‘ia] –for instance– is a true diphthong as well (cf § 5.2-3).

The concept of a phonetic *syllable* also remains overly subject to the influence of writing systems and of grammatical and metrical traditions. It is natural that books on linguistics and dialectology should speak about the phonic side of things as well, using transcriptions. However, a minimum of rigor would be of great benefit, since otherwise there is the risk of spreading and reinforcing unmitigated errors, which compromise and discourage effective learning.

For example, with the third millennium already here, books are still published which give the provisory *IPA* table (of 1993), naturally with the mistakes included as well (corrected in 1996), when it would have been a simple matter to download, or refer to, the (currently) definitive table shown on the official website. Obviously, this would only bring us to the level of the official position, which is hardly satisfactory; but at least major errors and embarrassingly naïve missteps could be avoided (within the limits of the incompleteness and ingenuousness of the official table)... There are also those who succeed in producing brilliant achievements like ⟨ \sqrt{tj} ⟩ (ie the monogram with a sort of ⟨umbrella⟩ on top, in the place of the normal /tj/, or the possible ⟨ \sqrt{tj} ⟩).

It pains one to observe that even those who believe they are doing modern and scientific phonetics (by virtue of it being acoustic) do not consider even for a moment the actual substance of objective phonetic reality. As we will continue untiringly to point out, they still speak of ⟨tonic⟩ and ⟨atonic⟩ vowels where they mean *stressed* or *unstressed*, and of ⟨rising⟩ and ⟨falling diphthongs⟩ (as if on roller coasters), instead of sequences of consonant + vowel, in the case of /jɛ, wɔ/, &c, and natural, normal diphthongs, in the case of both /ai, au/ and /ia, ua/.

Instead of demonstrating their scientific rigor, they make it clear that they continue to believe in the fable of hiatus, recounting the tale of how Italian words like *mai* /‘mai/ (ie [‘ma:i] in the intoneme of a sentence, but [‘mai] in the preintoneme), magically, switch back and forth from having diphthongs to having a hiatus. In this idea, they follow metrical tradition (which is at the antipodes with respect to phonetics and phonology). In Italian poetry, therefore, these words are said to have diphthongs except when they are at the end of a line, where they are said to count as two ⟨syllables⟩. It would be much better to say in such cases that the result counts as two *morae*, however, contained within a single syllable.

The actual reality is much simpler –and more honest– since, as we have already noted, there is a diphthong in both cases. In an intoneme, it becomes longer, as happens generally with most citation forms (since these forms *are* in an intoneme). The lengthening does not, however, reach the level of ⟨[‘ma:i]⟩ shown in certain transcriptions. In any case, the result is always an objective and legitimate *diphthong* (unlike the ⟨presumed diphthongs⟩, which consist of /C+/V/). Hiatus is –therefore– an extremely inconvenient grammatical and metrical ⟨invention⟩, which refers, by virtue of tragically considering important the spelling with two vowel letters, to two quite different phonetic realities: true diphthongs, /VV, VV/

[VV, VV, VV], on one hand, which are monosyllables (including cases like /ia, ua, ie, io/), and bisyllabic structures, /VV/ [VV, VV], separated by a primary or secondary stress, and therefore unequivocally part of two separate syllables.

Perhaps, one day, it will be possible to speak of these things without having to guard against ambiguity. In the foreseeable future, it will be necessary to use one of the following terms, in order to be certain of avoiding any misunderstanding. The false diphthongs can be referred to as *pseudo-diphthongs*, a term with no need of explanation. For true diphthongs, the term *strict diphthongs* can be used, without using Greek prefixes. Prefixoids such as *nomo-* (νόμος ⟨*nómos*⟩ ⟨norm, law⟩), *cano-* (κανών ⟨*kanôn*⟩ ⟨norm, rule, canon⟩), *delo-* (δῆλος ⟨*dêlos*⟩ ⟨clear, evident⟩) would probably be less... *evident*. The terms *grapho-diphthong* and *phono-diphthong* could help to avoid the frequent confusion between writing and pronunciation (as we do with *grapho-syllable* and *phono-syllable*), but they would probably not be sufficient to guarantee the clarity of the exact conceptual difference. Clearly, everything said here applies to ⟨triphthongs⟩ as well.

We will briefly mention several radical solutions, as always for the purpose of greater precision. It would be possible to substitute the terms: *bivocals* (or *divocals*, going back to Greek for *di-* in order to avoid *spectacular puns*) and *trivocals*, at a general phonic level; moreover, to distinguish the phonetic and phonemic levels as well, it would be possible to speak of *bivocoids* (or *divocoids*) and *trivocoids*, and *bivowels* (or *divowels*) and *trivowels*; these terms could be finally completed with *monovocals* (going back to Greek again to avoid confusions with *univocal*), *monovocoids*, and *monovowels*. (Of course, *monovocals* would not be at all similar to... vocals not broadcast in stereo.)

It is certainly not merely a question of nomenclature – more appropriate and rigorous terminology and use of symbols would make it possible to understand better, and faster, without useless obstacles and misunderstandings. For example, the symbol ⟨[æ]⟩ should not be used for Arabic [a], nor for the [a, A] used in central-northern Apulia and Abruzzi (in high-southern Italy), nor still for the [æ, εa] of western Tuscany and the [a(ε)] of Genoa. Similarly, ⟨[t, tt]⟩ should not be used for the [tɛ] found in extreme southern (ie low-southern) Italian dialects and regional pronunciations. Such errors are not acceptable in acoustic phonetics either, or others such as ⟨seeing⟩ a new type of lateral consonant in a spectrogram of a London (particularly Cockney) pronunciation of *milk*, ie [ˈmɪɔk, ˈmɪɔʔk] (for the normal –and neutral– pronunciation, namely [ˈmɪlk]). What is present is in fact the simple ⟨vocalization⟩ of *l*, given that the actual local pronunciation is with a real vocalic element, [ɔ], which together with the preceding element, constitutes a normal phono-diphthong, [tɔ]. This diphthong can be found in § 2.6.6 of *HPr*; similar phenomena are also seen in the phonosyntheses of § 17.27-8 (in this volume), treating the dialects of Munich and of Vienna.

With the many websites dedicated to languages and dialects (originating both with individuals and with institutions, such as universities), everyone now feels authorized to propound his or her personal opinion in phonetics and phonology as well. As is generally understood, at least 95% of what is found on the Internet should be completely ignored. It is in any case quite embarrassing to see cer-

tain absurd or antiquated ideas uncritically revived, without any shame and with evident incompetence, including even in numerous < specially provided > high-level courses, for which the unlucky and uninformed public ends up paying large sums... in all kinds of currencies.

1.17. We give again here the discussion in § 1.11 of *M^aPI* (and as a result, there will be some minor repetition concerning terms such as *syllable*, *diphthong*, and *tonic*). It is almost unnecessary to observe that truly practical phonetics uses only rigorous terminology and rejects instead everything which is imprecise, ambiguous, or insubstantial. In this latter category, we have the terms given now, in quotation marks: <liquid> (*lateral* ^{or} *trill* [together with *tap* and *flap*]), <cacuminal/inverted/retroflex> (*postalveolar*[ized]), <guttural> (*velar*, *uvular*, ^{or} *pharyngeal*), <pharyngeal> instead of *laryngeal*, <palatal> (where *postalveopalatal protruded*, *prepalatal*, or true *palatal* may be meant), <spirant> (*constrictive* or *approximant*), <aspirate> (*constrictive* and *approximant*, ^{or} <aspirated> [Ch]). Moreover, <soft> (*voiced* or [*pre*]palatalized C), or *front V*), <hard> (*voiceless* or *velarized C*, or *back V*), <rough> (*voiceless C*), <mute> (*stop*, or *voiceless C*). Still more: <mixed vowel> (*rounded front V*, or *back unrounded V*, or *central V*), <evanescent/indistinct vowel> (*V* which are *non-peripheral* in the vocogram), <closed *a*> (*low back-central*, [a]), <open *a*> (*low front-central*, [A]), <rising diphthong> (*sequence of a consonant followed by a vowel*, § 5.2-3).

Terms such as <semivowel> and <semiconsonant> make it seem you can have your cake and eat it too, by using graphemes or symbols which represent consonants while being based on symbols for vowels (for example, *i*, *u* or *ĩ*, *ũ*). However, the result is only confusion, and therefore the terms should be avoided. At the same time, no one should think to use the term *phoneme* as if it were simply the <refined> way of saying *phone*, or sometimes even *sound* (cf § 1.5). Moreover, in modern treatments of phonetics and phonology, with large sections conspicuously devoted to intonation, and perhaps entire parts exclusively addressed to tone languages, avoiding the use of <tonic> and <atonic> –where *stressed/unstressed* are meant– would definitely be a step forward. (In fact, <tonic> and <atonic> signify etymologically <with tone> and <without tone>, a use which was quite appropriate in Classical Greek, a language with *tonemes*.) The use of <intervocalic> where *postvocalic* is meant (for example in cases such as /p/ in (Spanish) *copia*, *aplauso* /'kopja, a'plauso/, &c) is misleading or incomplete, and hence should be similarly avoided.

1.18. Despite the great diffusion of the term –especially among singers– it would also be positive to avoid the use of <vocal cords> in the place of *vocal folds*, since anatomically they are not cords at all, but instead two membranes.

And it would not be unhelpful to be able to leave behind the *syllable* of long tradition, since inevitably it evokes the notion derived from writing conventions. The solution is to use consistently *phonetic syllable* (or *phonemic syllable*) and *graphemic syllable*, in order to avoid every possible ambiguity. Thus, according to the need of precision, it is possible to alternate between the extended expressions

given here and the generic term *syllable* in different parts of the same work, apart from the very useful terms *phono-syllable* and *grapho-syllable*.

Other possible candidates for the technical terms could be *phoné* and *graphé* (in the classical Greek style), but it seems more productive to reserve them for the ideas of the *ability to speak* and the *ability to write* respectively, both in terms of innate human potentials and in terms of what has already been acquired and developed.

In *NPT/HPb* (and in *HPr*), a certain number terms replace, of course, some more traditional (but imprecise and generic) ones. These substitutions will be motivated when they are presented. They will especially please the more rigorous workers in the field, and we hope that also those who work in phonetics ⟨part-time⟩ or with less intensity will be convinced that these substitutions are real improvements, not merely changes for the sake of change. Part of this program is the natural substitution of more well-known auditory terms (which are therefore less explanatory and, as a result, less useful) with articulatory terms: for example, we replace ⟨fricative⟩ with *constrictive* and ⟨affricate⟩ with *stopconstrictive* (from *stop*-[*con*]stric*tive*).

1.19. It is moreover inevitable (when seeking continually to improve and complete the practical potential of phonetics), that sometimes certain symbols will need to be adjusted. Phonetic symbols, as a matter of fact, must be very precise – otherwise, they are nearly useless. In the course of the progression of knowledge, it is necessary, from time to time, to add new symbols, so that new phones discovered while analyzing new languages can be conveniently represented. Our notational ideal seeks to maintain, as much as possible, an intuitive and natural link between symbols and phones (or prosodic and paraphonic traits, as well). For this reason, new symbols are devised according to phonic similarities within a given group, subject to the limits inherent in making the symbol drawable and visibly distinct from all others. We also strive to maintain relationships with traditional symbols which at this date have solidly demonstrated their worth.

It can happen that a new symbol can be more appropriate for representing something which once could have been satisfactory, but not now any longer. These substitutions occur in order to maintain and render more regular the relationships between the various symbols, and can be desirable even when they cause the phonic value of a certain sign to change with respect to how it was used in earlier publications. An example of this necessity is what has happened with certain symbols in the table of fig 10.12, such as [ɟ, ɟ̥, ɟ̥, ɟ̥, ɟ̥, ɟ̥]. Often in these cases, only the definition has changed, becoming more accurate and precise, as is the case with [ɟ], which is not defined as ⟨palatal rounded⟩ any more (least of all ⟨labial-palatal⟩), but instead, and more correctly, as ⟨postpalatal rounded⟩.

As we have said in § 8.5, we prefer not to avoid the perfectly grammatical use of suffixal comparative and superlative for *front* and *back*, contrary to mostly non-phonetician common usage.

2. Doing phonetics

2.1. What does ‹doing phonetics› really mean? Well, our point of view is decidedly *practical*, but not superficial; *descriptive*, not without a necessary theoretical component; and also *oriented towards teaching*, in the sense of working towards the consciousness of different systems and making it possible to compare these systems through different kinds of transcriptions and through articulatory diagrams.

Therefore, contrary to a broadly held view, doing phonetics does *not* mean *giving a sound to a letter* (of the orthography), or to a combination of letters. This approach is still used by grammar books, and shows that they fail to understand what true phonetics is. In fact, as we have already pointed out, it is absurd to begin with orthographic systems (which in many cases would be better considered ‹cacographic› or even ‹moody› systems, given all of their exceptions and idiosyncrasies) when attempting to rationalize the relationship between letters and phonemes.

Doing phonetics is *not* –of course– *giving a sound to a symbol*, either, although this notion is a step forward – at least it refers to two distinct systems which in some way come together, thereby providing in the end an indication for orienting oneself. Certainly, this point of view is completely insufficient with respect to the problem of producing a passable pronunciation. Naturally, the results are vague and not concrete; as a matter of fact, the normal practice is a series of attempts, oriented only to a limited extent in a fixed direction and guided by incomplete knowledge, which finally arrive at something which sounds acceptable, or more probably, intuitable and... tolerable.

2.2. Instead, *doing phonetics* seriously, or indeed, doing phonetics at all, is completely different – it means *giving a symbol to a sound*. This task is not so trivial as the man on the street and perhaps acoustic phoneticians and theoretical phonologists might imagine. The problem is not to succeed in ‹digging up› a symbol (out of some miserable few dozen), which might allude to the sound in question – the results would be decidedly absurd, as will be seen shortly. However, what is found explained –and perhaps transcribed as well– in books on the subject (not only on dialectology, glottochronicle [ie history of languages], glottodidactics, linguistics, phonology, acoustics, but even on phonetics!) seems often to have been put together without adequate perception of the sounds in question.

Far too often, people believe that they are doing phonetics ‹scientifically› if they merely use a few *IPA* symbols, which they typically consider naively to be like graphemes. For example, in certain books and websites, the vowels are given using the official trapezoid (cf \mathbb{C} 7). However, the vowel elements are not placed in points within the diagram which correspond correctly to real pronunciations, but precisely in the locations of the ‹cardinal vowels›, and they are therefore marked

by black dots placed right on the edges!

In the case of differing languages, such as Basque, Spanish, Greek, Bulgarian, Hebrew, Swahili, Sioux/Lakota, and Japanese, all of which have five vowels represented phonemically by /i, e, a, o, u/, these authors make it seem like all of these languages have exactly the same vowels, and they all coincide with the five <cardinal> ones. In this manner, they conceal other important facts as well, including the nature of Japanese /u/, which is in reality /ɯ/, without lip rounding and more fronted (without considering prosodic differences as well).

In fact, GIVING A SYMBOL TO A SOUND necessarily implies several successive and linked phases, in which hearing, mimicry, kinesthesia, comparison, adjustment, and mnemonic storage are all activated.

First of all, it is necessary to be capable of sufficiently PERCEIVING the sound, so that it can be identified with a precise PHONE which can adequately represent it.

The next step is to be capable of REPRODUCING the sound using the appropriate phone, which is carried out especially through imitation which is immediate, *ie*, occurring promptly after hearing the sound.

The third step, which is indispensable, is to PRODUCE the phone using kinesthesia (*ie* consciousness of the necessary articulatory and phonatory movements), even when the auditory stimulus is not close at hand. However, auditory memory can still be an important guide, whether it is *particular*, referring to a specific sound in a given language, or *general*, involving comparison with similar phones, with the help of experience in listening to and producing the phones of many languages.

In this manner, it is possible to produce a particular phone even days, months, or years after having heard it (and to be reasonably sure of producing the correct phone, even in the case of languages which one has never heard). The secret of good notation is for it to be realistic, and consequently, truly useful.

2.3. As a matter of fact, the fourth point, which is fundamental and decisive, is –as our <definition> suggests– finding a way to SYMBOLIZE the specific phone by choosing the most appropriate symbol out of several hundred (not just a few dozen) elements. If, after careful consideration, none of the available symbols is capable of decently representing a particular phone, it becomes necessary to find its position with regard to all other known phones. In this way, it will become possible to decide whether it is truly a new phone in need of a new symbol. If a new symbol is necessary, it can then be designed by following the general criteria of necessity, distinctness, and availability.

Doing all of these things is not remotely comparable to trying to have the work of analysis done by acoustic apparatuses, or by computer programs. There is quite a difference between what machines (which are incapable of discernment) are capable of doing, and what can be done instead with the human ear and phono-articulatory apparatus, given proper attentiveness, ability, and passion.

Acoustics is unable to recognize the relative importance of single traits. As a result, it ends up putting on the same level what is *essential* (*ie* fundamental and normal), *complementary* (*ie* equally important), and what is instead *incidental* (*ie* purely inconvenient, whether due to trivial homogenizations of the data or, on

the contrary, uncritical overdinctions).

The *phonological competence* of native speakers is particularly oriented towards what is essential. The *phonotonic competence* of the analyst uses what is complementary as well. *Instrumental competence* does not distinguish between the first two categories and what is merely incidental (and, far too often, it only confuses people).

Another great advantage inherent in the knowledgeable use of a large number of accurate symbols, both *segmental* (phones and phonemes) and *suprasegmental* (prosodic: length, stress, intonation, and ton[em]es) is that in this manner, considerable light is shed on what was once called <articulatory basis>. Indeed, using precise symbols, all of the fundamental *phono-tono-articulatory* information is provided, and this information, on its own, leads spontaneously to comparisons with the phonic systems of other languages, provided that these are likewise transcribed reliably. All of the differences present emerge immediately, including differences of intonation, which could not be included using other more theoretical and much less precise methods.

Our techniques run little risk of seeming to be removed from the subject, or of appearing to have been added for no reason (or perhaps just to make things more difficult). On the contrary, they become the normal way of doing phonetics seriously without leaving out wholly essential <nuances>.

In conclusion, it is *not* enough to <think one is doing phonetics> – it is essential to succeed in *actually doing it*, according to the METHOD OF NATURAL PHONETICS, or, more succinctly, the *phonetic method*. Obviously, it is *not* enough to *perceive*, it is necessary to *receive* as well; one should not settle for *skimming* superficially, one should *observe* and *examine* carefully; it is not at all sufficient to *hear* and *see*, it is essential, instead, to *listen* and *look* (at the transcriptions and the numerous diagrams, of course – *sight*, *hearing*, and *kinesthesia* are inseparable)!

2.4. Therefore, doing phonetics means managing to truly enter the phonic system of one or more languages, thanks to a rich array of symbols as well. The symbols of the official IPA are not at all sufficient, and they create the illusion of successfully doing phonetics, while all that has been achieved is at most a bit of phonology. Far too often, these studies proceed without the minimal understanding of what the phonetic structure in the case in question actually is.

Clearly, it is phonology which is a part of phonetics (cf § 1.9), and not the other way round as some believe. In fact, within the rubric of phonetic analysis and description, there is a functional component as well. Therefore, FUNCTIONAL PHONETICS (or *phonemics*) is an indispensable part of phonetics, but only a part. There is little which could be done working with phonology alone, just as there is little which could be done with just acoustics. Instead, what is needed is a global vision: articulatory, auditory, functional, descriptive, and contrastive (while verifying various characteristics acoustically).

On this subject, it is interesting to note that the phonotonic data of NPT/HP*b*, and of HP*r* as well, have been compared with a sizable collection of acoustic data from various sources, or at times, with data from a single source fur-

nished by *different authors*. In practice, the correspondence is dramatic, not only with data from a single source, but also with data from multiple audio recordings analyzed accurately and subsequently *normalized*. When we speak of normalized data, we refer to the practice of ⟨averaging⟩ the results of multiple speakers and many utterances in various contexts, while keeping in mind phonological considerations and excluding inappropriate samples.

The ⟨discoveries⟩ of sociolinguistics also require normalization; otherwise there is a significant risk of irremediably creating confusion, even when ⟨scientific⟩ data is used. This fact has been responsibly demonstrated in several recent works. We will not cite these sources (as ⟨examples to follow⟩), simply because this ought to be the normal way of doing things, not the ⟨alarmist⟩ or ⟨scoopist⟩ manner of far too many publications.

According to this global framework, every phonic system is an autonomous and complete organism. It contains *phonemes*, with their *taxophones*, as well as *prosodemes*, with their particular *realizations* (concerning length, stress, tones, and intonation). To give a simple example, a vowel element of a language, no matter how similar to an element of another language, must be considered in relationship only with the other vocalic (and consonantal and prosodic) elements of its own phonic system, within its own *phonemic space* (cf § 1.5).

Therefore, if it is necessary to *codify* (pronounce, or *transcribe*), and to *decodify* (listen, or *trans-read*) as well, it is essential to constantly refer to only elements from the system of the language in question. The term *trans-read* should be taken quite seriously, since it means ⟨to read a transcription *appropriately*⟩, using the actual phones (as well as tones and intonation) of the language transcribed. It does not mean at all ⟨to *take a shot* at reading a transcription⟩, using merely the phones of one's own personal accent.

Otherwise, the result is an unbelievable and indecent hybrid which we will try to exemplify here. We will use typographic expedients, which may suggest the discordant effect indicated – an effect keeping the learner far away from the (useful and enjoyable) purpose of the phonetic method. Therefore consider, for example, the following words: *fall, dog, play, tomorrow*. Now, we write exactly the same words again: ⟨fall, dog, play, tomorrow⟩. The reader will not fail to note that something is wrong. The four series which follow now have the same internal harmony as the first one (even though there are notable external differences between any two of them): ⟨fall, dog, play, tomorrow⟩, ⟨fall, dog, play, tomorrow⟩, ⟨fall, dog, play, tomorrow⟩, ⟨fall, dog, play, tomorrow⟩.

Therefore, even though the results are different from each other, each contains the same words, and (within each series), the same character type. This sameness produces the consistency which is fundamental within a unitary and homogeneous system.

2.5. Therefore, in the pronunciation of a particular language, it is essential to use only the phones and prosodic elements of that language. Instead, it is almost a rule that people use the phones of their mother tongue, adding a few phones from the other language because these are not present in their own language. Of

course, the rough idea is to complete the inventory when things are missing, but ideally all work should take place within a single system, even with parts containing elements which are similar in the two languages.

In fact, no matter how similar they may seem, the elements of one language are never exactly like those of another – due to the relationships with other elements, if nothing else. For example, the Italian /i/ is similar to the Spanish, or Portuguese (Brazilian or Lusitanian), or French ones. However, the Spanish /i/ is in opposition with only *four* vowel phonemes (/e, a, o, u/), while the Italian or Brazilian /i/ is in opposition with *six* (/e, ε, a, ɔ, o, u/), and the Lusitanian /i/ is in opposition with *eight* (/e, ε, a, ə, ɔ, o, u, ĩ/, which in an exclusively Lusitanian –not diaphonemic– system, would be represented by /e, ε, a, ə, ɔ, o, u, ĩ/, together with /i/, of course).

At first sight, the Brazilian system could seem exactly the same as the Italian one, as both have /i, e, ε, a, ɔ, o, u/. Instead, the two systems are different, particularly since in Brazilian (and also in Lusitanian, but with additional differences), even nasalized taxophones are expected ([ĩ, ẽ/ẽ̃, ɞ̃, õ/õ̃, ũ], followed by [N] too, ie a nasal consonant element); without these nasalized phones the Brazilian pronunciation would not be genuine (while in Italian it would be a regional one if with nasalized vowels).

As for French, /i/ is in opposition with *fourteen* other vowel phonemes (/e, ε, a, ɔ, o, u, y, ø, œ; ẽ, ã, õ, õ̃/ and /œ̃/, ie, the traditional ⟨ə/ə̃⟩), including the four nasalized vowels (/ẽ, ã, õ, õ̃/), which are true phonemes, in French, and not mere taxophones.

Apart from these considerations, the actual phonetic realizations are not exactly the same, even though we use the same phone [i]. The differences can be seen by comparing the vocograms (or vowel *quadrilaterals*) of these languages in *HPr*. The same is true of the other elements which ⟨correspond⟩.

2.6. If we now consider German and English, even without going into too much detail (the details are available, of course, in \mathcal{G} 2 & \mathcal{G} 5 of *HPr*), it is clear that we will have to take into account phonemic vowel duration. Instead, in the Romance languages we have considered, vowel length is –practically– only phonetic (although it varies considerably from language to language; consult the specific chapters, or to see the differences more rapidly, consult the transcriptions at the end of those chapters).

In German and in English, the ⟨/i/⟩ (⟨short *i*⟩) is considerably more open than in the Romance languages, giving [ɪ] in German and [ɪ] in English, respectively. However, even more pertinent to the present discussion of phonic systems is the fact that in the Germanic languages, the opposition of duration is relevant (ie, phonemic, distinctive) as well: German /ɪ, i:/ (*Schiff, schief* /ʃɪf, 'ʃi:f/ [ʃɪf, 'ʃi:f]); English /ɪ, i:/ (*bit, beat* /'bɪt, 'bi:t/ [bɪt, 'bi:t] – as we have observed repeatedly, we prefer a less abstract type of notation than the predominating form, which less usefully continues to give ⟨/i:/⟩, even for English, &c).

What emerges quite clearly from this discussion is the sorry state of those grammar texts and language courses which ⟨describe⟩ the German /ɪ/, or worse still, the

English one, as ⟨a short *i*, as in French *vite*, or in Spanish *listo*, or in Italian *fitto*⟩. When one considers that many Italian regional pronunciations have a vowel in *fitto* which is not at all short, the absurdity is evident! Of course, the same is true when one tries to teach the close and short French (or Spanish, or Italian) /i/ [i] by citing the same old misleading example: *machine*. Often, silence is golden...

2.7. *Foreign and regional accents* result from pronouncing a national language according to the phono-tonetic system of a particular and recognizable area (or of a particular and recognizable social group), especially in Italy, and in German-Spanish- or Portuguese-speaking countries. Therefore, it is important to begin to carefully examine one's own pronunciation, so as to be in a position to work towards the goal desired. This task involves learning to analyze one's own sounds, classifying them into precise phones (and transcribing them with adequate symbols), and then seeing how these phones are in turn part of particular phonemes.

At this point, it is indispensable to have a reliable and careful description available for the system of the language one is seeking to learn. The necessary comparisons can then begin – they should be carried out objectively and systematically. Obviously, it is imperative to follow an extensive series of specifically directed exercises, together with checks. These exercises consist of listening to good recordings and recording oneself to perform ⟨merciless⟩ examinations of how much progress has been made – without cheating, otherwise all of the effort is wasted.

For the pronunciation of Italian, Italians can rely on the seven chapters of *M^aPI* treating regional pronunciations for the initial diagnosis of how regional their pronunciation is. To (attempt to) achieve a neutral Italian pronunciation, Italians (and foreigners as well) can use the rest of *M^aPI* (including the two audiotapes which come with it), together with *DⁱPI*.

Guide to the figures

2.8. The *orograms* of *NPT/HP_h* (and *HP_r*) use symbols according to certain conventions, which make the orograms easier to understand (and tell apart). It is therefore important to become familiar with these conventions, so as to be able to make the best use possible of the rich iconic framework they provide. It is difficult to understand why some ⟨phonetics⟩ texts contain only a small number of illustrations, or even none at all. It is certainly true, however, that it is better to give no illustrations at all rather than to provide imprecise (or erroneous) ones. It would be better still if certain books were not produced at all...

In the *vowel orograms* (cf fig 8.8), it is of great importance to pay careful attention to the location of the *marker* indicating the center of the back of the tongue. It is even more important to take note of the precise location within the white (or transparent) miniature vocogram, placed at the center of the oral cavity (with respect to the still more precise location seen in the normal, larger vocograms), and the *shape* assumed by the entire back of the tongue. In this way, the various vowel orograms can be compared (or a subset of them, such as those relevant for a particular language).

Observing the position of the lips (particularly for the rounded vowels) and perceiving the opening of the jaw (obtainable from the distance between the upper and lower incisors [ie teeth]) are both equally important, given that they are interdependent.

All this effort should lead to a real understanding of the vocoid articulations and of the different movements of which they are composed, so that an active panorama is produced in the mind, and not simply a passive vision of things. In phonetics, merely passive and memorized knowledge has little use – except to confuse and discourage people!

Obviously, the true analysis and description of the vocoids of a given language occurs through the use of the large *vocograms* (vocograms in the strict sense), given that these diagrams succeed in showing nuances very precisely (as can be seen in the phonosyntheses of \mathbb{C} 16-23, and in \mathbb{C} 2-13 of *NPT/HPb*, as well as in *M^aPI* and the various Italian regional pronunciations given therein).

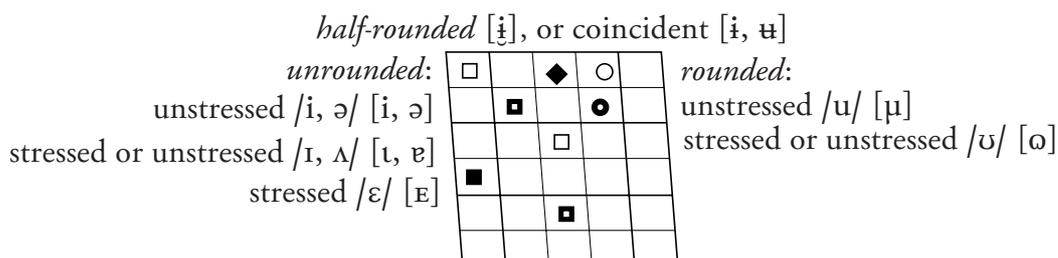
2.9. Therefore, we will now consider what can be found in *vocograms*, which should be observed, analyzed, and scrutinized calmly in all of their particulars. Vocograms are full of details, without which it is impossible to come close to the <spirit> of a language, manifested especially through vocoids, then through pitch, and finally through contoids. Even a single millimeter makes a notable difference on a vocogram (as well as on an orogram or on a tonogram).

This is the <magic> of phonetics. In fact, those who fail to perceive it accuse phonetics of being cold, dry, incomprehensible, difficult, and useless to boot. Instead, it is extremely useful, which is fundamental, and fun as well!

Vocograms are subdivided into 30 boxes, where the appropriate *markers* are placed according to the shape imparted to the lips. *Round* markers denote lip rounding (as in [u, o, ɔ]), and *square* ones denote normal lips (or spread lips; in any case unrounded, for example [i, e, ε, a]), cf fig 8.2 & fig 8.7-9. It is also true that the vowel orograms (in the miniature vocogram part) also contain circular or square markers according to the lip position. However, it is clearly much easier to see the markers in the (large) vocograms, where it is essential to use them properly (cf fig 2.1).

Naturally, there are also vocoids which can occur stressed or unstressed; for these, the symbols are *black-and-white* (ie black with a white center), as in *singing* [ˈsɪŋɪŋ] or in neutral British English *lover* [ˈlʌvə], while we have *lovers* [ˈlʌvəz, -ɜz], or neutral American English *lover* [ˈlʌvɪ] and *lovers* [ˈlʌvɪz, -ɪz].

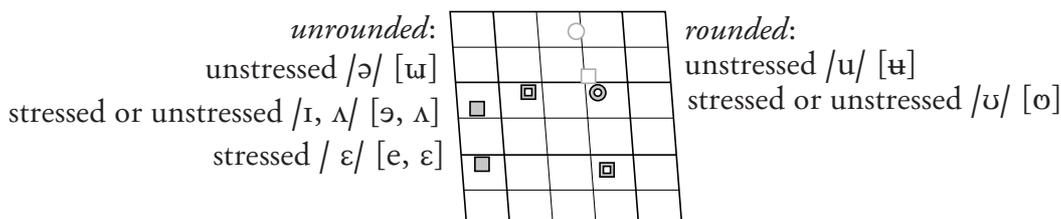
fig 2.1. Different markers for vocoids.



It is possible to use the *diamond* markers as well (squares rotated by 45°: \diamond) to show lip positions which are half-rounded, halfway between round and normal (as in [ɨ], cf § 8.10), or for perfectly coinciding rounded and unrounded V (as [i, ɥ]).

Besides the *shape* of the markers, their *content and shading* are also important. *White* markers (where the normal shape is not filled in by any shading) represent unstressed vocoids, such as English [ə], or for example those represented by ⟨o⟩ in Italian: *poiché*, *grido* [pɔi'ke, 'gridɔ] (or possibly, in certain languages, half-stressed vocoids, but not fully stressed ones). *Solid black* markers represent vocoids which are always –or mostly– stressed, as in *yes*, *but* ['jɛs, 'hɛt].

fig 2.2. Markers for variants.



2.10. Markers can also be filled with *grey* in order to indicate variants (contextual ones – the fundamental *taxophones*, pronounced using special phones; possible ones, such as those used in regional accents – *geophones*; or those related to various social groupings – *sociophones*). In neutral British English, for instance, there are quite a few taxophones, especially including those produced when there is a following [t] – for example in *but*, *bull* /'hʌt, 'hʌt/ ['hɛt, 'hʌt:], or *feet*, *feel* ['fi:t, 'fi:t] (while *feeling* has ['fi:liŋ]). It is impossible to continue to ignore this last group in vocograms and phonetic transcriptions.

A representative example of Italian geophones is given with the varied realizations of the phoneme /a/ [a], which include: [ʌ, ɑ, ɔ, ɛ, ɐ, ɐ, ʌ]. Some of these realizations can constitute sociophones, since they are more commonly found in more marked (instead of less marked) regional accents (these details can be found in *MaPI* in both cases). For English, the reader could see again the examples given in § 1.5-7. Grey markers can also have white centers in cases where they refer to vocoids which can occur unstressed as well.

At times, it can be necessary to improvise a difference in the marker or in its shading in order to represent important realizations which depend upon the position in the word with respect to word boundaries, stress, syllable structure, less common use, or simple occasional variation. In this manner, it becomes possible to avoid the use of supplementary vocograms. The purpose of these special markers will be explained clearly, whether in a text placed close to the vocogram, or in the main treatment. The most common convention is the use of dashed lines, particularly for unstressed vowels which are represented by ordinary unshaded (white) markers.

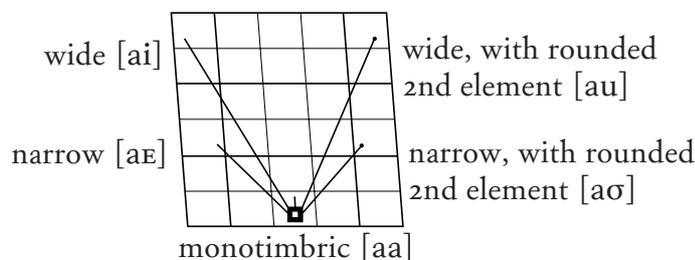
For examples of these rules in practice, the reader is invited to consult a few of the vocograms in \mathcal{G} 16-21, or the vocogram of the ⟨international⟩ French accent (cf fig in § 4.4.1.1 of *HPr*), southern French (cf fig in § 4.4.3.1 of *HPr*), or German

(with its various accents: \mathfrak{G} 5, *HPr*), or Brazilian Portuguese, Russian, or Arabic (cf \mathfrak{G} 7-10 in *HPr*, as well).

But let us now proceed to the conventions concerning *diphthongs* (formed by two tautosyllabic vocoids – *ie*, two vocoids contained in a single syllable, cf § 5.2-3). Diphthongs are denoted with the appropriate marker for the starting element, together with an unbroken *black line* which moves all the way to the precise location of the second element (cf fig 2.3-5). If the endpoint is an unrounded vowel, nothing is added; instead, with rounded vowels, a minuscule *black dot* is placed at the end of the line. If the endpoint is half-rounded, the marker at the end is a *diamond*; this marker is thus analogous to the normal-sized marker that would be used for a half-rounded first element (\diamond).

Meanwhile, the phonemic and phonetic transcriptions completing the information provided are placed around the vocogram.

fig 2.3. Diphthongs (occurring stressed and unstressed).



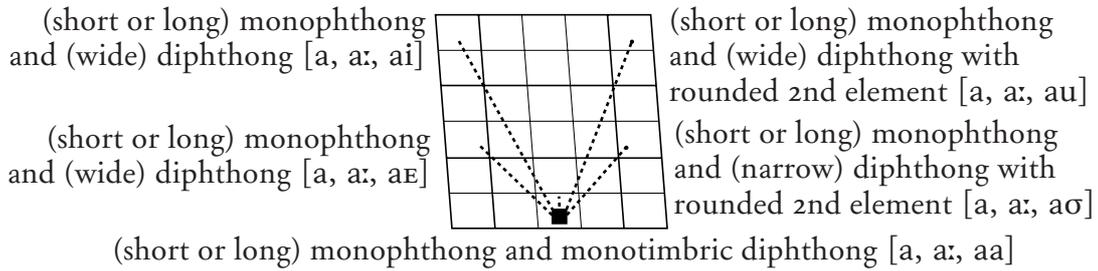
2.11. Diphthongs are considered *wide*, when their figure contains a fairly long line, and *narrow*, when the line is rather short. Besides these *ditimbric* (two-timbre) diphthongs, beginning and ending with distinct vocoids, *monotimbric* (one-timbre) diphthongs also exist, with the second element in the same box as the first element, but at a different point of that box.

This last group consists of quite narrow diphthongs (which therefore are almost like long vowel phonemes), and often the line is extremely short. Because the distance is so short, a dashed line can be reduced especially in such cases to a single short segment, or to the black dot alone for diphthongs with rounded second elements.

Completing the survey, we come to *vowel gemination*, or *vocoid doubling*. These cases involve vocoids which are neither short nor monotimbric diphthongs (as can be seen from the vocograms). However, a given vocoid is repeated within the same phonetic syllable, but without any movement in the vocogram at all: [aa] (cf fig 2.4).

When the first element of a diphthong is the same as that of a monophthong given together in the vocogram, it is possible to show the monophthong and the diphthong together by showing the diphthong with a *dashed line*, instead of an unbroken one (which would indicate the diphthong alone). When variant diphthongs occur, they are denoted by grey markers together with unbroken lines (or, in the case of an unstressed variant, the marker will be an unshaded figure with the edge and line both dashed).

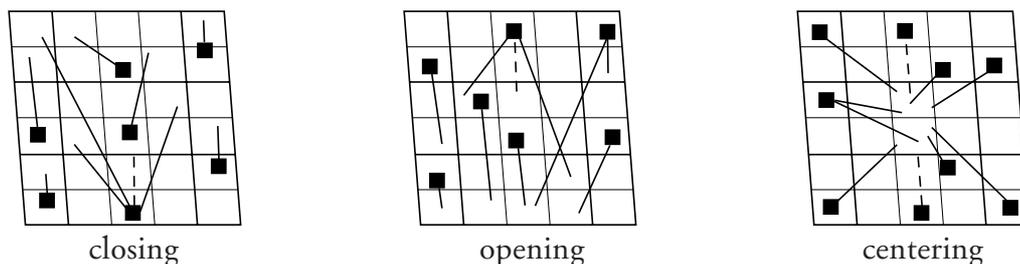
fig 2.4. Short and long monophthongs and diphthongs with the same starting point (here, all stressed).



Moreover, according to the direction of their movement, *diphthongs* can be classified into three categories: *opening* (when the second element is lower), *closing* (when the second element is higher), and *centering* (when the endpoint is [ə], or [ɜ]). In fig 2.5 (where all the symbols given are unrounded, for the sake of simplicity), the diphthongs given with thinner lines in the first and third vocograms ([aɜ]), or in the second and third ([iə]), could be considered closing, opening, or centering, depending upon the phonological interpretation and upon what sorts of similar diphthongs are present (or absent) in that language in other areas of the vocogram.

For example, if [aɜ] comes together with [æɛ, ɒɔ], or [iə] with [ie, uo], then [aɜ] will be considered closing, and [iə] opening. Analogously, diphthongs whose second element is not exactly central, including cases with [ə, ɛ, ɤ, ɛ; ø, ɔ, ɔ, ɔ] (and [ɪ, ɛ, ɯ; ɛ, ɛ, ɛ; ɤ, ɤ, ɤ; ø, ɔ, ɔ, ɔ]), could be profitably considered centering, if structural conditions so indicate. For example, in British English, *beers* /'bɪəɪz/ [ˈbɪvɛz] and *bear(s)* /'bɛəɪ(z)/ [ˈbɛɜ(z)] are in the same group as *beer* /'bɪəɪ/ [ˈbɪvɛ], even in accents which present [Vɛ, Vɛ] for /Vəɪ/. Therefore, the best decisions are normally made considering both phonemic and phonetic transcriptions.

fig 2.5. Closing, opening, and centering diphthongs.



2.12. In the vowel *orograms* (and in the vocogram) supplied in fig 8.1-2, we give the most extreme vocogram positions, with the express purpose of showing the limits of the region of oral space used for vocoids. Instead, the orograms in fig 8.8 give positions which are more commonly found in the languages of the world; these are slightly less peripheral. In fact, certain trapezoids (or even <vowel triangles>) show all of the symbols perfectly aligned along the edges, <threaded> along the lines (rather like pearls, since they are all circular as well), so that they extend outside of the margins. We find such figures decidedly odd, given that the objective reality of vocograms is quite different.

There are still phonetics ~~or~~ dialectology books (and linguistic atlases as well), which continue to present absurd and impossible vowel triangles. Such figures make it seem possible to distinguish all vowel timbres along only a series of lines. In reality, the timbres are distributed throughout the vocogram quadrilateral, hence in the middle columns *as well* (front-central and back-central, with or without lip rounding).

Because of our firm commitment to avoid perpetuating into the third millennium concepts which are so unscientific, we will *not* give an example here of these triangles (even of one reduced to <front> and <back> vocoids). Instead, we will present the <necklace> with its threaded <pearls>, starting with [i] (*i*), down to [a] (*a*), hanging like a pendant, and back up to [u] (*u*), in the following order: *i-i-i-i-e-e-e-e-ä-ä-a-a-ä-a-o-o-o-o-u-u-u-u*.

i i i i e e e e ä ä a a ä a o o o o u u u u

Unfortunately, a rapid examination of these <symbols> (in italics and full of diacritics) shows that there is little chance of being able to convert them automatically into our columns (even by trying to devise rules governing particular alternations): front ([i, ɪ, e, ɛ, æ]), and front-central ([ɪ, ɪ, ə, ɛ, a, A]), then moving back up from the bottom, back-central rounded ([ɔ, ɒ, o, ɔ, ω, μ]), and back rounded ([ɒ, ɔ, σ, o, u, u]). It is nonetheless true that at least the low central element, [a] (*a*), is more realistically located (despite its triangular shape) than in the official IPA...

2.13. With *orograms*, which are fundamental for consonants, we follow several conventions, some of which are more intuitive than others. For example, it is sufficient to indicate nasals by showing the velum lowered, as in the cases of fig 10.2. Nasalized articulations are indicated in the same way, including vocoids (fig 11.9, at the bottom) and contoids (three in fig 10.6.3). There are also *prenasalization* (fig 11.4) and *nasal explosion* (fig 11.3, on the right).

Stops are shown with a raised velum, and as with nasals, there is contact between two or more articulators (fig 10.3).

Constrictive orograms show the articulators close to one another (fig 10.5), together with a useful convention (albeit a bit less obvious and objective) consisting of a *horizontal black line*, immediately above the base of the orograms. This line is intended to allude to the constriction – in this case, the noise of friction produced by the flow of air through the narrowest point of the articulation. If the line is not continuous, but *divided into three* segments (as in the case of [j]), the contoid is *semi-constrictive* (intermediate between a constrictive and an approximant).

In the case of *grooved* constrictives (cf § 9.13), there is a *curved line* placed on the *corona* of the tongue. This curve is meant to be a reminder of the longitudinal groove that characterizes these articulations. The curve (like the horizontal line)

is segmented in the case of semi-constrictives. The same marks naturally occur in the diagrams for stopstrictives as well.

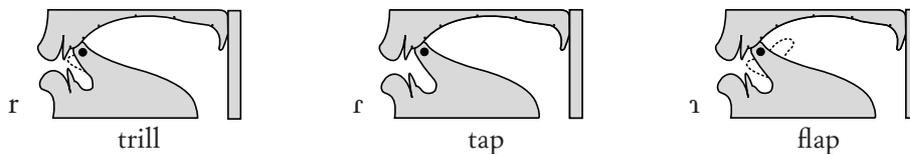
fig 2.6. Slit (or ungrooved) and grooved constrictives: [θ] and [s].



In the case of *approximants*, there is visibly more space between the articulators, and the horizontal line (used with constrictives) is omitted. However, there can be a black *arrow*, slightly smaller than the one used for laterals, and this arrow signifies lateral contraction –*lateralization*– a trait accompanying and characterizing some of the approximants (fig 10.6). Semi-approximants have a *dotted horizontal line*.

Trills, *taps*, and *flaps* are marked with a black *dot* placed upon the articulator which is in motion (whether this is the tongue tip, the uvula, or the lips). Moreover, a *dashed outline* is added to trill orograms, and *two* dashed outlines to flap orograms (cf fig 10.7). In the special cases of *constrictive taps* and *constrictive trills*, the characteristic horizontal line is also present above the base of the orogram.

fig 2.7. Trills, [r], taps, [ɾ], and flaps, [ɹ].



Laterals are identifiable with an *arrow* placed on the most fundamental part of the articulation. If the arrow is black, the contoid is *bilateral*; if white, the contoid is *unilateral*. If these last are also *constrictives*, there is the horizontal black line, as well. If the contoid is instead a *lateral tap*, a white dot is shown. In the particular case of [ɭ], a black dot is shown. The reason is that this phone is actually a diaphone referring to a general oscillation between [ɾ, ɮ, ɭ, ɭ], respectively tap, lateral (or lateralized) tap, tapped lateral, and lateral, in which none of the above predominates over the others in habitual speech patterns (fig 10.8 & fig 10.13).

fig 2.8. Bilaterals, [l], and unilaterals, [ɭ].



2.14. *Stopstrictive* orograms have a small black part, which refers to the stop phase of these contoids. The stop phase (as can be seen in fig 10.4) is homorganic to the place of articulation of the constrictive phase, which immediately follows and constitutes the second part of these phones. The result is a unitary (though

compound) phone, since its full duration corresponds to the length of other stops or constrictives, not to the length of two phones combined in sequence.

Stopstrictives naturally have a horizontal line at the base of the figure. *Grooved* stopstrictives have a curved line as well, representing the groove. Stopstrictive *taps* and *trills* have a black dot, in addition to the dashed outline occurring in the case of stopstrictive trills. Stopstrictive *laterals* are shown with a sort of black oval representing the stop phase, together with a white arrow placed on top indicating the unilateral contraction occurring in the same place of articulation.

Orograms for *non-pulmonic* contoids require several observations. *Dejectives* (or clicks, fig 11.8 – further details can be found in § 11.13-6), whether stops or stopstrictives, present the characteristic backwards movement of the back of the tongue; this is shown by a black *arrow* pointing *rightwards*.

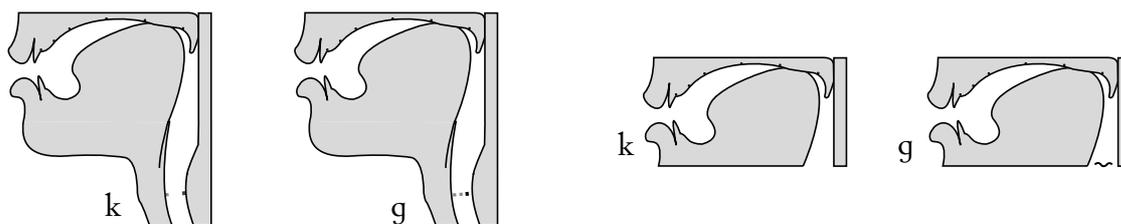
Ejectives and *injectives* have orograms which are necessarily larger, since it is essential to show the movement of the larynx. Its movement is shown with (nearly) *vertical arrows*: pointing upwards in the case of ejectives and downwards with injectives.

2.15. The larynx would also need to be shown in orograms indicating the difference between *voiced* and *voiceless* contoids. For example, in the case of the pair [k, g], it would be necessary to use two different large orograms, as shown in fig 2.9. These orograms would make clear that in [k], the glottis is open and does not vibrate, while in [g], it is closed, but not solidly, so it can vibrate and produce the voicing which distinguishes [g] from [k].

A compromise to save space could be to use two normal-sized orograms and to show, not the glottis, but rather a sort of vibration of the sound wave only with voiced contoids. This vibration could be marked with a wavy line in the place where voicing is first perceptible, as seen in the lower part of last orogram of fig 2.9.

But for ordinary purposes, it is sufficient to use the normal orograms without voicing distinctions. Differences of voicing can be expressed with phonetic symbols, as long as the symbols are sufficiently rigorous. In special circumstances where it could become necessary to emphasize the distinction, such as perhaps a contrastive study between languages with different positions and functions for the larynx, special solutions can be devised. Nevertheless, it seems to us that the careful and precise use of phonetic symbols is generally the best solution.

fig 2.9. Possible methods of showing differences between voiceless ([k]) and voiced ([g]) C.



2.16. *Palatograms* (fig 9.2.2) are used to show contact with the palate throughout the course of an articulation. Palatograms are mainly useful for certain con-

toids. It would be possible to employ them to add detail to the description of vowels, particularly those which are not back. However, vocograms and orograms are more useful in this case for purposes of description and teaching. When parts of the palatogram are shaded black, this convention signifies full contact of the articulators (in the stop phase of stopstrictive contoids). Grey shading indicates instead, naturally, constrictive contact (as in the characteristic constrictive phase of stopstrictives). If the palatograms of the constrictives [θ, s, ʃ] are compared to those of the corresponding stopstrictives, [tθ, ts, tʃ], this difference becomes clear immediately.

Dorsograms (as in fig 9.2.1) give a new perspective, which is not longitudinal but instead transversal. These figures are particularly helpful to illustrate the difference between *flat* tongue position (the UNMARKED position, since it requires fewer phonic features, for *slit* contoids) and *grooved* tongue position or *lateral contraction* (or lateralization) – these are the two *marked* possibilities, in comparison with the unmarked position.

When arrows are used in *labiograms* given in profile (as in fig 8.3, fig 8.7 & fig 9.1), they indicate the direction of the characteristic movements, as produced by particular facial muscles. *Frontal* labiograms require no particular explanations (cf fig 8.9 and those just mentioned) – the vertical gap, which is steadily larger as the jaw opens, is fairly intuitive.

fig 9.2.3, moreover, helps to show the fundamental difference between *trills*, *flaps*, and *taps*, in terms of the type and number of contacts.

Another useful type of diagram is the *laryngogram* (as in fig 4.4), which should be analyzed attentively. These laryngograms are naturally *optical* (as well as being schematic and frozen in a particular instant), representing what can be seen with a laryngoscope, or a throat-specialist's mirror. They are not *acoustic* laryngograms, which measure vocal fold vibration.

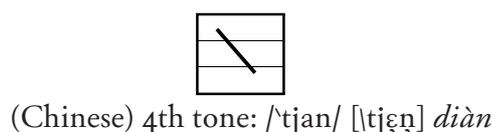
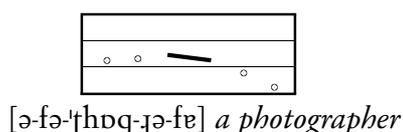
We will also comment somewhat on the use of three special symbols in fig 10.2-8. The *degree symbol* (°) marks articulations which are not included in the table of fig 10.1. The eight-point *asterisk* (*) marks the (few) official articulations that coincide with our own (cf § 10.01). The employment of <&> means that the same symbol is present in an orogram found nearby. This situation occurs with articulations differing only in a single, not readily perceptible nuance. Given that the difference is small, a different symbol is unnecessary in such cases, even though there *is* an articulatory difference (and the orogram should indicate it).

2.17. We will now move on to *tonograms*, which are divided into three juxtaposed bands, of *high*, *mid*, and *low* pitch (where the levels are not absolute, but relative to the voice of each particular speaker). Both in preintonemes and in intonemes (cf § 6.4.5.1-4 & § 13.8-34), as in tones as well (cf § 6.4.4 & § 12.17-18), *lines* (or *dashes*) placed at different heights and with various slopes represent stressed phono-syllables. *Dots*, on the other hand, represent unstressed syllables, while *shorter lines/dashes* represent half-stressed syllables (with secondary stress, but with pitch indicated by the position in the tonogram).

In phonotonic transcriptions, secondary stress is indicated by two dots placed

close together (smaller than a single dot), variously oriented according to the tonetic necessities. Secondary stress on a medium level pitch is denoted with [,\$] in order not to create confusion with the hyphen we use to show syllable boundaries. This use is consistent with marking primary stress with ['\$]. ‹Unstressed› phonosyllables (or better, weakly stressed – *ie*, weaker than half-stressed syllables) with mid level pitch are not marked in any particular way. On the other hand, in tone languages, syllables with mid pitch and weak stress are preceded by a dot placed at medium height, [,\$].

fig 2.10. Examples of intonation and tones.



Guide to different types of transcriptions

2.18. There is no doubt that a great number of symbols is used, in *NPT/HPb*. However, these symbols are not superfluous for those who desire to do phonetics thoroughly (and not merely ‹easily›, and inevitably superficially). A small number of symbols ineluctably leads to mediocrity, whereas a large number of symbols opens the way towards the true understanding and ‹savoring› of pronunciation.

Even if study has been methodical, concentrated, and accompanied by careful exercises, it will naturally remain occasionally necessary to check the value, the nature, and the connections of certain symbols (and concepts as well) which are less frequent. The best way to do so is to look for explanations in the right place, or places. In fact, the necessary answers, verifications, and connections, as well as new perspectives, can be found in the *general index* (contents) and in the *analytical index* (index), by flipping through the chapters and sections, and by paying attention to the *tables*, the lists of *symbols*, and the groups of *figures*.

The major categories are, of course, *vowels*, *consonants*, *intonation*, other *prosodic features* (stress, pitch, duration/length), and *paraphonics*.

It would be complicated –and probably useless– to try to re-explain these matters in a general summary (probably already too compressed and complex). Therefore we will merely suggest following the directions given here, emphasizing only the meaning of the different ‹parentheses› used to enclose the symbols.

Slashes –/|– always denote phonemes, on a theoretical and abstract level; instead, (square) brackets –[]– are used exclusively for phones (and taxophones) – *ie*, the practical and concrete side of things, which nonetheless naturally comprise essential generalizations and normalizations, without which it would be necessary to speak only of single, unrepeatable realizations of particular individuals. For example, *play* /ˈpleɪ/ [ˈphleɪ].

2.19. Doubling the ‹parentheses› indicates a further accentuation of their primitive values. Thus, *double slashes* –//||– refer to a still more abstract or theoretical

level of phonological characteristics, such as what we have with German *wiederhaben* //ˈviːdərhaːbən//, with respect to the ordinary phonemic or phonetic transcription /ˈviːdərhaːbən/ [ˈviːd̪r̪hɑːbm̩].

Double brackets –[[]]– indicate instead more exact symbols, in cases where showing extremely precise nuances is desired. An example of the use of double brackets could occur when showing the articulation with a high tongue tip, [s̺, f̺] (where these symbols need to be drawn carefully so they can be clearly distinguished from [s, f]), instead of the articulation with a low tongue tip, [s, f], which is considered (more) normal.

Another case is given by partial, or even substantial nasalization, as indicated above (§ 1.15): *thanking* [ˈθæŋkɪŋ] ([[ˈθæŋkɪŋ]]). The official *IPA* has no way to show light automatic nasalization, and in fact, it arbitrarily and misleadingly employs our symbol ⟨[æ̃]⟩ for a particular type of phonation, creaky voice (for which we use [æ̰]).

Finally, when *angle brackets* –⟨ ⟩– enclose symbols or diacritics (such as ⟨̣⟩), these refer to *paraphonic* elements. Instead, when they enclose orthographic symbols (such as ⟨a⟩), they refer to *graphemic* elements.

fig 2.11 summarizes rapidly the types of transcription. It excludes, however, the numerous ⟨poorer⟩ transcription systems, found in every country, using extremely varied conventions. These systems are generally based on the graphemes of the language in question (and their symbols are meticulously written in italics – as graphemes), while adding to them all sorts of diacritical markings in every direction. Similar diacritics often denote vowel and consonant timbres which are very different in their articulation, as we will observe critically in § 7. In fact, each of these ⟨phonetic alphabets⟩ tends to attribute to the orthographic symbols used their most familiar values in the language for which the alphabet was devised, without any opening towards other languages at all.

Symbols *not* enclosed in brackets or slashes represent *phones* in treatments of general phonetics; for example, in the diagrams and tables of fig 6.2 & fig 10.1. Instead, in the consonant tables for the phonosyntheses (§ 16-23) and for *HPr*, the symbols *not* enclosed in slashes represent *phonemes*, even though they are represented by fairly specific symbols (this is to avoid sacrificing precision, but without making the visual effect too intimidating – in fact, here the symbols represent phones *as well*).

In the phonemic transcriptions accompanying the phonetic transcriptions, it is also possible to use more generic symbols to indicate phonemes. We have followed this approach in *HPr*, where we have also indicated the corresponding results and presented them in easy reference form in the consonant tables.

The first line of fig 2.11 shows a sentence in conventional spelling, enriched by a helpful notation for the intonation. The succeeding ten lines each apply a different transcription type to the sentence in question. To complete the survey, we have four more transcription types, which we have exemplified with four English words: *merry*, *Mary*, *marry*, *Bill*.

The last transcription is taxophonetic, which could simply be referred to as ⟨phonetic⟩. (Unfortunately, outside of *NPT/HPb* and *HPr*, ⟨phonetic transcription⟩

Transcribing by hand

2.20. Finally, several fundamental observations should be made on the topic of transcriptions carried out by hand. It is important to draw every symbol (and every diacritic) exactly as it is *printed* in this manual (where we employ the elegant *Simoncini Garamond* font, in our personal version *Simon(ani)*). It is necessary to be very careful not to confuse one symbol with another one which is similar to it, but –obviously– not exactly the same and with a different value.

To start out, transcriptions should *not* be carried out in *cursive* (or in *italics*). In fact, contrary to habits in normal handwriting, the symbols should *not* be *connected* to each other, but should instead be kept separate, as in print. Moreover, symbols should not be *simplified*, changing *n* into *u*, or *m* into *uu*; or *a* into *ə* or *ɑ*. Dots should not be omitted, and so *ι* should not be written in place of *i*, nor *j* for *i*. (Here we are using italic symbols, just to draw attention, but they should not be used for transcribing words or sentences – not */fə'netiks/*, but */fə'nɛtɪks/*.)

⟨Stylistic⟩ additions and modifications should be avoided. Thus, *d* must remain different both from *q* and *ð*, just as *g* is different from the grapheme *g*, and *h* from *h̄*. The same goes for cases like *z*, which should be kept distinct from both *z* and *z̄*. Moreover, [*λ*, *r*, *ɾ*, *ʁ*] are, by the same token, quite distinct from [*λ/h̄*, *z*, *ɹ*, *ʁ*], &c. Naturally, *small capitals* should be avoided as well, since the symbols [*A*, *B*, *E*, *I*, *G*, *L*, *N*, *P*, *R*] are different from [*a/a*, *b*, *e/ɛ*, *i/ɪ*, *g*, *l*, *n*, *p*, *r*].

In conclusion, it is necessary to set aside every normal writing habit which could lead to confusion between the symbols. The best strategy to reach this essential goal is to begin observing every symbol with great care. A ⟨typographic eye⟩ should be developed which pays attention to every detail, from the size and orientation of a stroke ([*t*, *ʃ*, *ʒ*, *ʒ̄*], [*β*, *β̄*]), or of a symbol ([*e*, *ə*, *ɐ*], [*ʀ*, *ʁ*, *ʁ̄*, *ʁ̄̄*], [*ɑ*, *ɒ*; *a*, *ɐ*], [*ɛ*, *æ*; *ɛ*, *ɜ*], [*ʌ*, *ɔ*; *v*, *ʌ*; *c*, *ɔ̄*], [*r*, *ɹ*, *ɹ̄*], [*h*, *d*, *ɥ*], [*f*, *ʃ*; *j*, *ʒ*; *ʒ̄*, *ʒ̄̄*; *J*, *J̄*]), to the presence or absence of a sort of *serif*, or the type of serif present: [*i*, *ɪ*, *ɪ̄*; *u*, *μ*, *υ*; *o*, *σ*, *ο*].

In working towards this objective, the observations in Ch 8-9 will help the reader to *look at*, not just *see*, the symbols found throughout the text. In this way, it will prove easier to ⟨accept⟩ the fact that [*g*] always has the value present in the word *get* */'gɛt/*, and not the value in *gem* */'dʒɛm/* (while *GIF* –or ⟨*.gif*⟩– can be pronounced either */'gɪf/* or */'dʒɪf/*), &c.

3. Pronunciation & phonetics

3.1.1. One of the most important way HUMAN BEINGS are different from ANIMALS is that they have developed an extremely elaborate SYSTEM OF COMMUNICATION, in order to exchange information with others of their kind. This system of communication is called ARTICULATED SPEECH, in the sense that it has a DOUBLE ARTICULATION. In effect, the SIGNIFIER, which is constituted primarily by <sounds> of the language is articulated at the first level into units carrying MEANING (ie LEXEMES, or lexical elements, and GRAMMEMES, or grammatical elements, as in: *booklets* /'bʊk-lɪt-s/), which are the SIGNIFIED. The second level of structure is constituted by units which have no intrinsic meaning, but which –in combination– form the meaningful units just now considered. These second level units are the PHONEMES: /'b-ʊ-k-l-ɪ-t-s/, which are our main interest, together with their actual realizations (in the spoken language). This goes for both the study of the pronunciation of one or more foreign languages, as well as for the pronunciation of one's own mother tongue.

Even though the first contact with a foreign language is more often via a visual code, that of WRITING (ie words printed in a book), the fundamental code of a language is, in reality, ORAL. The natural acquisition of language is perfectly reflected in the case of a child, who first learns to understand, reacting to the SOUNDS of the language used by those close by, and afterwards learns to imitate these sounds and use them to communicate.

3.1.2. Thus, the first ability one learns is UNDERSTANDING THE SPOKEN LANGUAGE, then one goes on to use it, to SPEAK it. Much later, if ever, one learns to READ THE WRITTEN LANGUAGE, and finally to WRITE it. The first two stages are fundamental, whereas the last two are complementary – they constitute additions which are not strictly necessary. This point is demonstrated by the fact that many people do not know how to read or write, yet they use their own language or dialect perfectly. What is more, the majority of the languages spoken in the world, aside from the LANGUAGES OF <CULTURE>, do not (yet) have a written form. Even when there is a *written* language, the *spoken* language is generally used much more frequently by the average person.

The spoken language is, therefore, the fundamental form of communication among humans, while the written language is only a practical means for fixing in place the spoken language: indeed, it is well-known that *scripta manent, verba volant* (<writing remains, words fly away>). However, writing is not always (or is no longer) a clear and unequivocal method of representing the phonic form of the language. For example, how are *applicable*, *kilometer* (-tre), or *Caribbean* pronounced? Or *alveolar* and *laryngeal*?

3.1.3. Phonetics –the scientific study of the spoken language– ‹studies› (ie analyzes and describes) the phonic realizations of language, using various approaches. ARTICULATORY (or physiological) phonetics examines and describes the production of linguistic sounds by phono-articulatory organs. In this task, it is aided by AUDITORY (or perceptive) phonetics, which helps to verify which sounds have been uttered, and functions as the fundamental tool for perceiving the sounds to be imitated. ACOUSTIC phonetics studies the physical structure of sounds by using sophisticated devices developed for this purpose, or also computer programs. INSTRUMENTAL (or experimental) phonetics, uses many different instruments, ranging from very complex ones (such as the spectrograph), to simple ones (such as the artificial palate – today, both of these also have computerized versions). HISTORICAL (or diachronic) phonetics is yet another branch of the subject; it studies the development of the sounds of a language in the course of time.

Finally, FUNCTIONAL (or structural) phonetics, also referred to as PHONOLOGY, PHONEMICS, or PHONEMATICS, addresses the value and function of a language's sounds, and their interrelationships. In this way, phonology links the phonic aspect of language to the psychological, cognitive, and social aspects of oral communication, and also to the meaning of words (LINGUISTIC SIGNS), in order to arrive at the meaning of sentences.

The two absolutely necessary and indispensable branches of phonetics are, therefore, ARTICULATORY phonetics, together with auditory feedback, and FUNCTIONAL phonetics, also called phonology. All of the other branches indicated above are inevitably complementary – things which are added successively, providing further details, but without being fundamental. They can be helpful for deepening the understanding of particular aspects of the subject, on levels which are definitely more academic than practical. The transformation of languages and the formation of new languages can be explained through historical phonetics, while acoustic phonetics can help improve communications technology. In phonology, there has been a particular development along exclusively academic lines, which seeks to ‹perceive› the deep internal structure of languages, hoping to arrive at universal conclusions, which could better be defined as GLOTTOSOPHICAL (abstract and philosophical, more concerned with theoretical speculations about language) than GLOTTOGRAPHICAL (concrete and practical, with descriptive and comparative applications, as well as use in teaching contexts).

3.1.4. This *phonetics handbook* leaves out almost everything which we cannot do by ourselves. Consequently, it mainly deals with ARTICULATORY and FUNCTIONAL PHONETICS, starting from zero and working towards very high levels of knowledge and detail. As a matter of fact, one should start any possible subsequent specialization –profitably– only after having really mastered the articulatory possibilities (and not just the rudiments) of a number of languages, with their phonological systems presented glottographically, not glottosophically.

Every phonic realization, besides carrying a conceptual, strictly linguistic MESSAGE, also contains a collection of additional ‹extra-linguistic› INFORMATION. This information includes the age, gender, personality, state of mind, social group, and

—sometimes— even the health of the speaker. Individuals also have their own physical characteristics, which make it possible to recognize people by listening to their voices while not seeing them, and even when not hearing clearly what they are saying.

Thanks to the radio, television, movies, telephones, CD's, and cassette recorders, not to mention ease in (and often, the necessity of) moving from one country to another, it has been recognized that the phonic aspect is much more important in the study of languages than the written aspect (which is quite secondary and dependent upon the spoken aspect). But for a long time (and many continue in this manner even now), study was focused upon morphosyntactic structure, <idiomatic expressions>, and a language's literature, using only books and without much (or any) concern for the phonic structure – the PRONUNCIATION of that language. But if one wishes to UNDERSTAND a speaker of a given language, if one wishes to BE UNDERSTOOD, for reasons of work or merely for tourism, it is indispensable to learn pronunciation. The distinctive elements, or PHONEMES, should be used appropriately, thus resulting in a sufficient level of phonological competence.

3.1.5. It is still commonly believed that the only way to acquire a good pronunciation of a foreign language is to live for a long time in the country where that language is spoken. Certainly, this method can lead to a certain fluency in the language. However, practice has shown that (except for unusually talented people, who are rather rare) the method is of almost no use for acquiring a good pronunciation. Everyone knows some people who have lived for decades in a country they were not born in, and who nevertheless continue to speak with a broad <foreign accent>. The accent of most English-speaking people who use a foreign language is notoriously bad, even though they can frequently understand others and make themselves understood in that language, at least to some extent.

Languages do not differ only in their grammar, their vocabulary, and their idioms. They also differ a great deal in the SOUNDS of which they are composed, in the way these sounds are combined to form words and sentences, in the quality of stresses and their distribution, as well as in the way voice pitch rises and falls (ie intonation). Inevitably, when people study a foreign language (unless they have a special talent or special training), they transfer into the new language the phonic habits of their mother tongue. Contrary to what some continue to think, the desired <new habits> do not arise spontaneously, but require a bit of work, as with any other activity.

Language teaching has improved noticeably ever since it was realized that pronunciation, like any other aspect of linguistic study, can be studied scientifically. It was realized, in other words, that the true nature of any sound used in any language can be described precisely. PHONETICS works effectively side by side with the imitative abilities of the learner, guiding and indispensably complementing the entire endeavor. Phonetics makes available simple and complete scientific descriptions, thereby making it possible to abolish definitively reprehensible statements like <this sound cannot be described>, or <this sound must be learned directly by imitating your teacher>.

The phonetic method

3.2.1. The true difficulties in studying a spoken language are rarely well-understood. However, when they are clearly explained in detail, it becomes much easier to overcome them. Indeed, depending upon the particular phonological system of one's own language, *phonic interference* can be of four types: (1) certain phonemes are not distinguished (‹hypodifferentiation›), (2) more phonemes are distinguished than actually exist (‹hyperdifferentiation›), (3) phonemes are distinguished in a different way than in the language in question (‹reinterpretation›), (4) certain phones are confused with others (‹substitution›).

The PHONETIC METHOD consists in understanding completely the articulatory possibilities of the PHONO-ARTICULATORY APPARATUS and in working to recognize and produce freely a large number of sounds which have been systematically analyzed. The teacher explains how certain sounds are produced, and then pronounces them repeatedly for the students, who try to imitate them. Specially prepared diagrams are used in the explanations: OROGRAMS (from Latin *ōs*, *ōris* ‹mouth›, not from Greek *ὄρος* –*orós*– ‹mountain›), which are articulatory sections, usually sagittal (but also occasionally frontal, therefore, a sort of LINGUOGRAM), and PALATOGRAMS. These diagrams show the position of the various organs involved in the articulation of different consonant sounds. For vowel sounds, the VOCOGRAM (or vowel quadrilateral) is used. (The vocogram employed here is one with modernized shape and internal subdivisions; the result is a decisive improvement on the ‹official› shape. This last diagram has the defect of sacrificing precision on the altar of tradition; in fact, it stems from ideas which were partially erroneous from the outset – all of this will be seen later, while treating the voicoids, in § 7.)

If the teacher is reasonably competent (both in perceiving and [re]producing the various sounds), it will be possible to help students when their efforts fail to reach desired levels of performance. Useful techniques include repeating the correct sound multiple times while alternating it with the incorrect one, giving further advice, suggesting additional exercises, and obviously referring back to the articulatory diagrams (and showing accurate symbols).

An early and detailed treatment of intonation is more than appropriate, even though it is usually relegated to the later sections of textbooks, and often treated hurriedly. In fact, every tonogram should be analyzed in its smallest details (just like any other articulatory figure), in order to ‹discover› all of its characteristics and thereby acquire full familiarity with the concepts treated.

3.2.2. The *first difficulty* that must be overcome concerns the ability to PERCEIVE THE SOUNDS of the language sufficiently well. Individuals differ quite a bit in this regard, but it is not rare to find people who can easily distinguish a large number of sounds while noting even small nuances, who can reproduce at will sounds heard on other occasions, and who can compare mentally new sounds with familiar ones.

When beginning the study of a foreign language, there is a fatal and inevitable

tendency to <hear> the sounds of the foreign language in terms of those of one's mother tongue. However, with a bit of work, it is possible to begin to compare the sounds of the language studied with other familiar sounds, and to discover in this manner similarities and differences.

To succeed in pronouncing a foreign language well, the very first objective is to train the ear to recognize new sounds. Practice will definitely improve any ear, whether good or bad. The important thing is to notice the differences, and the fact that differences exist. However, there is no way of avoiding a certain amount of work in the form of organized exercises. In the study of languages, working to recognize many sounds and many nuances of these sounds is fundamental. It is clear that there is no way to imitate and use the sounds of a foreign language in an appropriate and timely fashion, without first managing to tell these sounds apart.

One of the most effective ways to work at recognizing new sounds involves listening to the dictation of <nonsense> words. These are invented words or syllables, which can contain both familiar sounds and unfamiliar ones. An alternative is to use words in a language unknown to the learner, who is expected to transcribe these words unambiguously, using the symbols of the phonetic alphabet. For use of this method, it is obviously necessary to have previously explained phonetic notation to the student through comparisons and articulatory figures. The tables should be used as well, since they are indispensable for classifying, comparing, and grouping the different linguistic sounds. Following the approach we have described here, it will become possible to associate every symbol to the corresponding sound (or rather, to the corresponding category of sounds), and vice versa.

3.2.3. From the transcriptions of the student, the teacher will be able to see whether the student has heard the dictation or audio recording correctly or not. The use of recordings has the advantage of permitting an unlimited number of identical repetitions, and it can include sounds which the teacher might not be capable of pronouncing adequately. Since the words used are not understood (due to being nonsensical or from an unfamiliar language), the student is forced to concentrate exclusively on the sounds, without relying on phonic redundancy and the knowledge of word syllable structure – information which would have aided the student in recognizing the words had they been in a familiar language.

If the student does not <hear> correctly, and as a result transcribes sounds which are different from the ones which were produced, the teacher will pronounce the word again, alternating between use of the erroneous sounds and of the correct ones. In this way, the student's ability to perceive the sounds will improve. It is clear that in this case, a teacher capable of reproducing the sounds actually transcribed by the student is superior to a prepared audio recording, obviously containing only the sounds considered in advance.

The exercises for the student should involve steadily greater details, according to the principle of gradually progression from tasks involving simpler problems. Exercises will be particularly useful if they have been prepared while considering the special needs of the language studied. However, the exercises are also essential for practice in general phonetics. Another possibility with great potential would

be the employment of audio recordings which anticipate and sound out probable mistakes by the student. Recordings of this kind could be prepared for particular languages and for learners speaking a specific mother tongue, whose problems of linguistic interference will be fairly predictable.

Of course, the best way to perform *meaningless dictations of nonsense words* is to record well (and with high quality) some not too long strings, starting from monosyllables, up to bi- and tri-syllables. In class (but especially during an examination, for good measure) all necessary repetitions are played aloud at random, but taking good note of their order (after prearranging the recordings, which have been previously chosen from several samples, already classified according to levels of difficulty). A good store of such strings (possibly as *mp3* files) will guarantee both that they are performed exactly alike every time they are <repeated> (without mishap), and that students will not easily recognize something on which they may have already been working.

Even teachers/examiners will benefit from this, as they will not risk to put up a poor show or to be contested because of some accidental modifications – and they themselves have the possibility of knowing exactly what they are using. But, most of all, it is easier to collect a vast choice of previously controlled materials, in order to deal with all possible aspects, with no likelihood of any English speakers dictating for instance something like [μ], or even [ʁ], instead of (a supposed <cardinal>) [u], or vice versa, &c.

3.2.4. The *second difficulty* to overcome in the course of acquiring a good pronunciation in another language is to PRODUCE THE SOUNDS of the new language. As is easy to verify, every language has a certain number of characteristic sounds, and usually some of these are not present in the student's mother tongue. Therefore, the student must learn to produce the new sounds, and this necessity implies new (and often unfamiliar) articulatory habits, which can only be acquired with particular exercises. The process is sometimes long and complicated, but will infallibly result in a rich and well-deserved reward, in terms of ease in understanding and in making oneself understood (not to mention appreciated!). It is important to keep in mind that there are many cases of sounds which are similar, but which nonetheless are not exactly the same, or are distributed differently; this is also a problem which should be approached systematically.

It is important for those who study another phonic system to understand their own perceptive and articulatory capacities. This understanding can be reached by working at recognizing the sounds and phonemes of one's own language. The realizations of these phonemes will vary quite a bit even among people speaking the <same language>, resulting in sounds which are, to a greater or lesser extent, different from each other, and more or less <correct>. It is very useful to be constantly aware of one's own phonic emissions, as well as those of other people. An example would be becoming conscious of whether we say normally (for now, it will be sufficient to understand that the different *symbols* show that different *sounds* are indicated): *yes* [ˈjɛs] or [ˈjɛs] (or perhaps: [ˈjes]); or also *will* [ˈwɪlː], or [ˈwɪlː] (or perhaps: [ˈwɪlː]), or: [ˈwɪɸː], or [ˈwɪɸ], or [ˈwɪʁ], &c).

3.2.5. It is also very important to notice what pronunciations are used by people we hear, as we learn to recognize analytically the traits that tell us, for example, that a certain person is –say– from London, New York City, Sydney, or Glasgow. This means not being satisfied with recognizing the accent from the phonic emissions taken as a whole, but instead identifying the particular vowel and consonant sounds used, given that these sounds either differ from –or coincide with– our own. The purpose here is not to see who has the better (or the worse!) pronunciation, but rather to discover the differences between the pronunciations, so that we can exercise our linguistic capacities: first to *perceive*, then to *reproduce* (*imitate*), and finally, to *produce*.

After a bit of phonetic exercise, it becomes possible to acquire good phonetic control. This entails full awareness of the sounds we emit and the ability to choose which sounds and articulations to use while speaking. We have, therefore, the ability to choose between <correct> and <incorrect> sounds in the phonic inventory of a foreign language, while eliminating and replacing the sounds of our mother tongue, which we would tend to use instead (because of the strong phonic interference of one system with the other, on both phonological and phonetic levels).

Of course, it is also possible to choose to use more <correct> sounds within one's own language. It is possible, for example, to eliminate those sounds, or groups of sounds, which have a (more or less) strong regional connotation. This allows one to work towards fulfilling the (fully legitimate) desire of acquiring a pronunciation more like the neutral pronunciation – spoken especially by professionals on the better radio and television channels. It is curious that these professionals are sometimes criticized harshly by people who do not know themselves what neutral pronunciation is, and who speak with extremely broad regional accents.

3.2.6. There is a *third difficulty* along the way for those who study the pronunciation of a language (whether it is a foreign language, or one's own): it consists in knowing where to use the sounds acquired. This is the knowledge of the correct DISTRIBUTION OF SOUNDS of a given language in the speech chain – ie in phrases, not just in isolated words. Memory is very important for this task, but the phonetic method once again comes to the aid of the scholar (and the conscientious learner – in practice, there is little difference between the two). The phonetic method makes it possible to SEE THE SOUNDS which need to be pronounced, and which are heard on recordings, already in their correct order in syllables, words, and sentences characteristic of the spoken language (in any variety, whether NEUTRAL – a term preferable to <standard>, cf § 1.11– or REGIONAL). The method succeeds in this goal by representing sounds and combinations of sounds with symbols of the phonetic alphabet.

Moreover, the use of phonetic transcription (and phonemic transcription, whether given separately, or better, together) has, as is well-known, the advantage of showing which sounds need to be pronounced, together with their precise qualities, and of showing assimilations and elisions where these occur, while marking clearly the position of stress in words and sentences. At the same time, the learner is not confused or distracted by traditional orthographies, which often do not

reflect at all (or any longer) the phonic structure of the language.

Transcriptions which also show the melodic curves of sentences are the most useful, since they are the most complete. This is INTONATION, which can be the only distinguishing element, in cases like *today* /tə'deɪ./, *today?* /ɛtə'deɪ?/, &c.

3.2.7. The *fourth difficulty* to be overcome involves acquiring the ability to pronounce *complete sentences* in the language studied with sufficient ease and rapidity. After having completed the preceding phases, it should be possible to recognize and to reproduce, without too much difficulty, the isolated sounds of the new language (or of the mother tongue). Until every new sound can be produced with sufficient ease and fluency (both when pronounced in isolation and when present in arbitrary combinations), the learner will not reach a (sufficiently) CORRECT AND NATURAL LEVEL OF PERFORMANCE in the speech chain. To arrive at this point, it is necessary to work at pronouncing all the sounds (especially the more difficult ones) in syllables and groups (especially the more complicated ones), starting with the simplest problem points and progressing towards the more complex ones.

These exercises are essential, since by working hard at pronouncing sounds and combinations of sounds, not only does it become possible to pronounce them correctly, but one also acquires practice in coordinating all of the tongue, lip, &c, movements which are necessary in order to pass smoothly from one sound to another without hesitations or mistakes.

We should not forget that children perform numerous auditory and articulatory exercises in the course of their early years, and this occurs before they have learned to recognize, select, and put into practice the characteristic sounds of their language.

Once one is capable of articulating smoothly all of the sounds of the new language, in all possible combinations, it becomes possible to acquire a fully natural delivery, pausing only where native speakers would pause. As a matter of fact, a series of words should be pronounced joined together, because the words belong to the same RHYTHM GROUP and are indissolubly connected on the semantic, pragmatic, and morphosyntactic levels. It is important not to stop in the middle of such series, because this creates undue divisions. In order to bring the whole program to a worthy conclusion, rhythm and intonation appropriate to the meaning of utterances should be used, whether in short or long sentences, up to daily (unplanned) conversation.

3.2.8. Finally, there is a *fifth difficulty*, which involves the acquisition of the *meaningful aspect* of language, and is therefore a difficulty of a different nature than the others we have considered. It involves WRITING and the traditional conventions governing spelling, which often have little sense today, or at least, do not any more. The problems with spelling have to do with the fact that writing tends to reproduce itself by inertia, without adapting itself to the natural and inevitable evolution of the spoken language (which is the truly fundamental aspect). Writing is (as we have pointed out previously) a decidedly secondary phenomenon, with respect to speech. However, an exorbitantly excessive value is currently attributed

to writing, owing to the practical utility of being able to fix on paper, or other (more) <modern> instruments, what would otherwise inevitably be lost. Thus, it is possible, by <snail mail>, by fax, or by E-mail, to send documents all around the world, and with great rapidity. Because of writing, and (traditional) printing, we can read ancient documents and make thousands of copies of them. We can extend the diffusion of learning and preserve works of genius. At the same time, we can also take notes and file them in order to help us take care of our daily tasks.

Of course, no one would think of denying the importance of writing. However, it needs to be emphasized that it is only a derivative of true language – ie spoken language. This point remains valid even though the great works of humanity, which have been preserved through writing, are of superior quality, in part due to being more formally and stylistically structured (including the possibility of still further structurings in later phases), despite having been codified in a system quite inferior in quality to speech.

As a matter of fact, written codification can only render an incredibly limited part of what we manage to communicate, unconsciously and involuntarily as well, through spoken language. In writing, the richness and genuineness of spoken language is inevitably lost; it can be partially recovered by steadily more precise phonetic transcriptions (or, in E-mail, with <emoticons>, or <smileys>, such as for example :-) ;-) :) :o) :-(: :o(<:-(>:-()). In these cases, we come back to writing, but using symbols which are much more accurate than those in standard orthography. We have, therefore, something which is TRANSCRIBED, with multiple possibilities of use, whether GLOTTOGRAPHICAL or GLOTTODIDACTIC.

While studying a language thoroughly, orthography is also important – but it should not be the only concern of teachers and learners...

3.2.9. In conclusion, we cannot say enough on the importance of avoiding (saying, and particularly) writing mistaken things about phonetics. Those who prepare linguistics and grammar texts should become well-informed, instead of spreading erroneous ideas and confusing beginners. The <correct> alternative would be to refrain from writing about topics with which one is insufficiently familiar. Specialization should have some purpose, and the serious specialists in the field would be more than happy to provide help – for the good of everyone, and especially of knowledge itself!

Moreover, it should be emphasized that the *true phonetic method* arrives at writing only after all of the phonic structures (vowels, consonants, and intonation –including tonemes, if these are present– of a given language) have been adequately learned. Consequently, according to the true spirit of the method, success in pronouncing the language fluidly (and therefore, in already knowing how to speak it for purposes of communication) should be reached before moving on to the <orthography> and all of its (sadly well-known) problems.

Instead, in <traditional> teaching (ie through the written language, and at times without ever hearing any native voices!), as well as in more <modern> approaches (communicative, with dialogues, but again, far too often without listening to <original> sound materials), the impact of writing is so preponderant that it is un-

avoidably deleterious. Of course, interference of writing with phonic production is normally inevitable, and it unfortunately limits even one's perceptual abilities. This occurs because of confusion due to false analogies, both because of differences in orthographic conventions with respect to those of one's mother tongue, and because of excessive irregularities in the <orthographic> systems themselves.

To make the learning of a foreign language's pronunciation simpler and more <natural>, the five difficulties (mentioned here, in § 3.2.2-8) should be systematically overcome, one by one. These tasks are already complicated on their own, and distractions owing to orthography should be avoided – the orthography should only be employed when learning to write words and sentences which can already be produced <naturally>. The problems with orthography should be effectively restricted to those which native speakers encounter, namely, finding out (while being ready for many surprises) how words and sentences one can hear and say are traditionally written. Instead, unfortunately, writing constitutes the most non-sensical obstacle systematically encountered in the study of languages!

3.2.10. The worth of the phonetic method is obvious. However, today's society is unfortunately one which gives first priority to writing, while ignoring pronunciation completely. Therefore, the least harmful compromise is to give the spelling together with the transcriptions. This approach will only work if a sufficient *trans-reading* occurs first, ie reading of the phonetic and phonemic transcriptions (in order to study carefully the relationship between taxophones and phonemes). The spelling can then be considered, but only to see how the word is officially written.

The fifth difficulty, that of writing, should be overcome with special materials designed for the task. These materials should use the phonemes as the point of departure, when explaining how to write things which the learner already knows how to transcribe. Unfortunately, <tradition> does exactly the opposite: it starts with graphemes, and tries to deduce from them which phonemes are used in practice in the words of a particular language. Obviously, in English, the endeavor is very arduous; but it is not without problems in French, as well. Other languages are not completely <phonetic> either; even with Spanish there are uncertainties and traps; &c.

Considering these problems, in order to fully apply the <natural phonetic method> (thus postponing writing till the end), in *HP_r*, at least, we have used the order: (1) [fσ·ɔnz̥, fɜ·ɔnz̥] (American and British pronunciations), (2) /fɔʊnɪmz/, (3) *graphemes*; however, we have not explained systematically the passage: [σɔ, ɜɔ] → /ɔʊ/ → ⟨o⟩, which will be more appropriately treated in books on the particular languages.

Instead, the intention of *NPT/HP_h* is not to fully explain the pronunciation of any particular language. It aims, rather, to make known the potential of phonetics and tonetics through examples which are mostly English and Italian or Spanish, but also from a certain number of other languages. Therefore, the order used is: (1) *graphemes*, (2) /fɔʊnɪmz/, (3) [fσ·ɔnz̥, fɜ·ɔnz̥] (although at times, element 2 or 3 will be missing, if less pertinent to the exposition). The spelling introduces

the example, while the transcriptions show it for what it really is; for this reason, they should be analyzed and read with great care.

As we have already pointed out more or less directly, the real ‹enemies of phonetics› are, ironically, those who ought to know it well as an integral and fundamental part of their work: teachers of languages, authors of grammars (or *grammatographers*), experts in teaching languages (or *glottodidacts*), glottologists, glottochroniclers (or *historians of language*), linguists, dialectologists, speculative phonologists, and acoustic phoneticians. The greatest embarrassment is that even those who call themselves phoneticians (or ‹phonetists›, and worse if ‹phoneticists› – even ‹phonetes› would be a little better, given its pedantic sound) far too often do not mind their ‹fines and phones›, nor their ‹tines and tones› (leaving aside the traditional *p*'s and *q*'s)!

The matter is scandalous, but it should not be a surprise. Indeed, things are exactly the same in every other field: what is it that ruins religion, the judicial system, the health system, politics, bureaucracy, skilled labor, and sports? Precisely those very people who are responsible for administering them! Besides 12% who try to do things seriously, all of the rest of these ‹operators› range from being inept to being harmful.

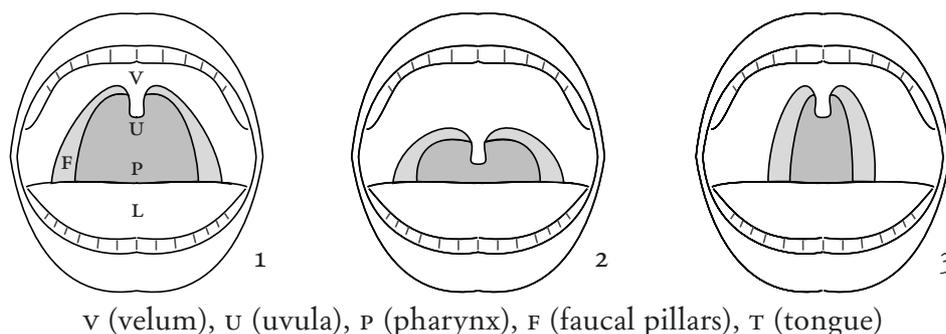
3.2.11. All things considered, to fully master another language, according to the principles of Natural Phonetics, is a very pervasive operation. In fact, besides entering the phonic system confidently, it becomes inevitable for the learner to acquire the typical facial expressions of the new language as well (and even gestures become more... consonant).

4. The phono-articulatory apparatus

4.1.1. To study the pronunciation of a given foreign language, or also the neutral pronunciation of one's mother tongue, it is necessary to have a fairly clear idea of the structure and functions of all of the PHONO-ARTICULATORY ORGANS. For this purpose, it will be indispensable, not merely useful, to provide diagrams which show clearly and objectively everything that can be useful for a simple and clear, yet precise study. Overly detailed descriptions of various anatomical parts of the phono-articulatory organs are superfluous and distracting, so we will only present what is essential.

To begin, it will be helpful to examine, with some care, what can be seen of the reader's own articulatory apparatus. The best way to do this is to open one's mouth as wide as possible while shining a penlight or a small flashlight inside, and to inspect the oral cavity, through the use of a handheld mirror placed in front. Keeping the tongue consistently lowered, it will be possible to see what has been simplified in *fig 4.1*. Here frontal views of the open mouth are given, first with the velum raised (1), then with the velum lowered (2, as occurs while breathing from the nose, even when the mouth is open); finally, the contraction of the faucal pillars (3, which produces the faucalization found, for example, in Sicilian pronunciation, especially around Catania).

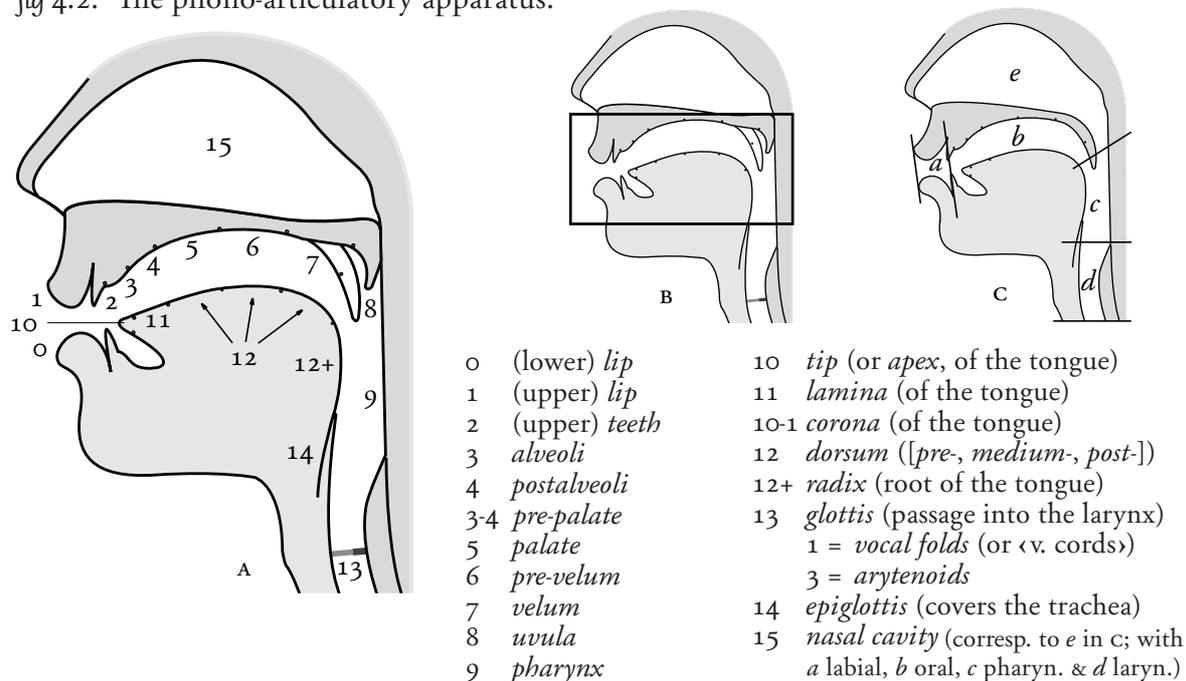
fig 4.1. Front view of the open mouth.



4.1.2. *fig 4.2* is an OROGRAM of the organs involved in the phono-articulatory process (down to the vocal folds, A.13; while *fig 4.3* shows the phono-respiratory mechanism, below the folds). It is helpful to compare *fig 4.1* and *fig 4.2* in order to see where the useful parts are found in both cases (we have numbered these parts only in *fig 4.2*, while in *fig 4.1*, we have labeled the less familiar parts with their initial letter). Thus, we have: the LIPS (0 & 1), the FRONT TEETH (the upper ones, which are more important phonetically, 2), the TONGUE (10-12, T), the PALATAL VAULT (3-7, V), the UVULA (8, U), and the PHARYNX (9, P) – although it should be

kept in mind that the proportions and the degree of jaw opening are different in the two figures.

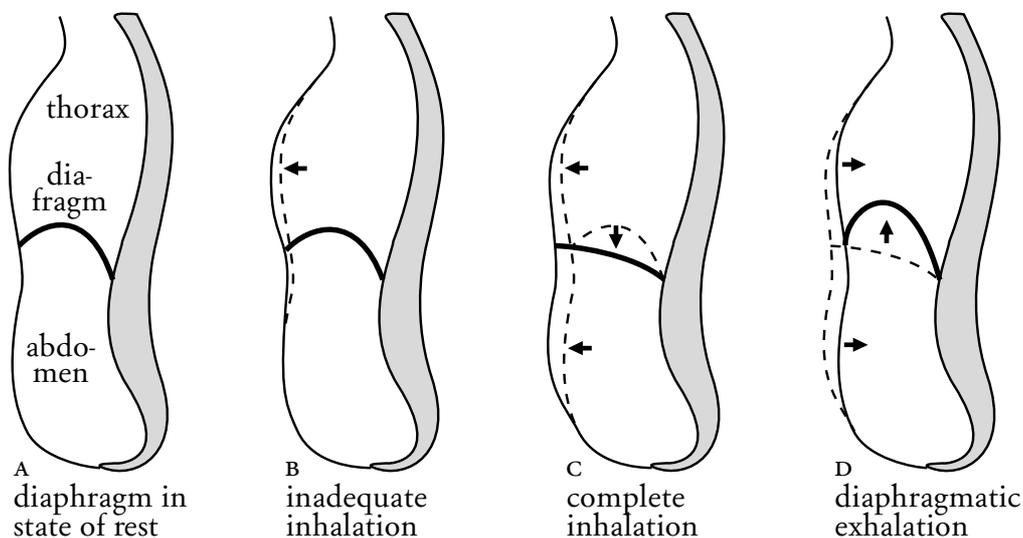
fig 4.2. The phono-articulatory apparatus.



In fig 4.1, it is also possible to see the faucal pillars (F, and their contraction, in the rightmost diagram). If they had been included in fig 4.2, that diagram would have become overly complicated, given that there the ALVEOLAR RIDGE (or *alveoli*, 3), the PALATE (or *hard palate*, 5, with bone structure), and the VELUM (or *soft palate*, 7, without bone structure) are all clearly visible, not to mention the nasal cavity (15), the vocal folds (together with the arytenoid cartilages, 13), the epiglottis (14), and other parts and cavities which will be seen below.

In fig 4.2, the distinct parts of the phonoarticulatory apparatus are visible: in (A) the general *zones* of articulation are given (with *articulators* and *resonators*); in (B),

fig 4.3. Respiratory mechanism.



the *fundamental area* for articulating vowel and consonant sounds is delimited. In (c) the five *resonators* are shown: the labial (*a*), oral (*b*), pharyngeal (*c*), laryngeal (*d*), and nasal (*e*) cavities.

It will be advantageous to learn the terms used in describing articulations and phonation (ie, the flow of expiratory air [essential for speech] in the passage through the glottis, and its various positions, cf fig 4.4 & fig 11.5). Moreover, the reader should examine as carefully as possible the phonatory organs and their movements while producing sounds of the mother tongue. The purpose is to attempt (with the help of the diagrams which will follow) to <feel> not just the positions, but also the *movements* taking place while producing sounds, since it is difficult or impossible to look at these organs while speaking normally.

In this manner, it is possible to acquire sufficient KINESTHESIA (ie awareness of organ movements and positions in the course of articulating sounds) to become capable of controlling, to the extent necessary, one's own phono-articulatory apparatus. The objective is to become able to produce, with a little practice, any sound, no matter how <strange>, of any language or dialect.

4.1.3. Most linguistic sounds are produced using expiratory air, which comes from the lungs and then encounters various obstacles formed by the articulatory organs situated in the various places of articulation. The principal FIXED ORGANS are: the TEETH, the ALVEOLAR RIDGE, and the PALATE.

The TEETH, particularly the UPPER TEETH, play a role in the articulation of *t* and *d* in most languages of the world (including Spanish, French, Italian), through contact with the tongue tip, as in Italian *dato*, as well as in the English *f* and *v* in *five*, here through contact with the lower lip.

The upper teeth are embedded in the ALVEOLI (ie *alveolar ridge*), which thereby constitutes the zone of the palate situated immediately behind the teeth. It can be easily touched by the tip of the tongue, or by a finger, since it is in a fairly prominent position. The English *t* and *d* in *today* and the *n*'s in *none* (and the Spanish *n*'s in *nones* as well) are articulated with the tip of the tongue touching the alveolar ridge.

The PALATE (or hard palate), is the part of the palatal vault behind the alveolar ridge, and it has a bone structure (hence the name: *hard* palate). The Spanish *ñ* in *baño*, is articulated with the back of the tongue against the palate.

4.1.4. The principal MOVABLE ORGANS are: the *lips*, the *tongue*, and the *velum*. The LIPS coming together produce the *p* in *pipe*. The TONGUE is the most important organ, given that it plays a role in most linguistic sounds. In fact, it is well-known that even extremely small alterations in the position of the tongue can change the resulting sound to the point that it becomes unrecognizable, or changes into a different sound. The principal parts of the tongue are: the TIP, the BACK, and the ROOT. However, in the descriptions of articulatory phonetics which follow, we will use more specific terminology; in the meantime, the reader is advised to take another look at fig 4.2.

The VELUM, or soft palate (since it has no bone structure), is normally lowered

while breathing through the nose (in white, in fig 4.2, as can also be seen with the nasal articulations, § 6.3.1 & § 9.8-9 & § 10.2 & fig 10.1). Instead, while speaking, it is usually raised so as to touch the upper part of the pharyngeal wall, thereby preventing air from escaping through the nasal cavity and the nostrils. The velum is naturally lowered for nasal sounds for the precise purpose of causing air to escape through the nose. The closing and opening of the velum can be easily seen by opening the mouth and breathing, first through the nose (the white part), and then through the mouth (the grey part). In this way, the velum can be seen to retract, closing the passage leading to the nasal cavity and thus assuming its most frequent position in speech. These positions are also shown in fig 4.1. As we have already mentioned, NASAL SOUNDS, for example *man*, *song*, *onion*, involve lowering the velum, just as in breathing through the nose.

4.1.5. It is helpful to subdivide the tongue further, as is indicated in fig 4.2: the TIP (OR *apex*), the BLADE (OR *lamina*, the part which ends up against the alveolar ridge when the mouth is closed and the tongue tip is touching the lower front teeth), and the BACK, which can usefully be subdivided into three parts: the FRONT (OR *predorsum*), the MIDDLE (OR [*medium*]*dorsum*), the BACK (OR *postdorsum*); and finally the ROOT (OR *radix*).

The TIP is relevant, for example, in the articulation of Spanish *t* (where it is against the front teeth), or in the English *t* (here against the alveolar ridge). The BLADE is used in Italian *s*, the PREDORSUM (together with the blade) in *ch* in *church*, the (MEDIUM)DORSUM in Spanish *ñ* in *sueño*, or in Italian *gn* in *sogno*, the POSTDORSUM in *k* in *book*. The ROOT is used in PHARYNGEAL sounds, such as Arabic *ħ* in *ħaal*. The predorsum, by itself, without involvement of the blade, produces PREPALATAL sounds, such as the first *n* in *onion*, or that in Russian *net* (*нет*, in Cyrillic).

It is good to subdivide the palatal vault as well into smaller and more precise parts. Thus, starting with the right extremity of the velum, we distinguish the UVULA, which is used in articulating the various *r*'s found in French *recevoir*, or German *rauchen*, or also a good number of the types of <French *r*'s> found in many defective Italian pronunciations. (Almost all of) the rest of the soft palate is termed the VELUM, whereas the part where it begins to harden (because of the bone structure) is called the PREVELUM. Moving forward, there is the PALATE, in the strict sense of the term, and afterwards, the PREPALATE (all visible in fig 4.2). In all of these zones, contact with the tongue occurs between its back and the subdivisions of it just now detailed.

As can be seen in fig 4.2, the PREPALATAL ZONE is subdivided further into the ALVEOLAR and POSTALVEOLAR ZONES, which can play a role in articulations involving the tongue tip or blade.

4.1.6. We have already seen that in Spanish and Italian, the *t* is articulated with the tongue tip against the front TEETH – in English, instead, it is articulated with the tip against the ALVEOLI, or alveolar ridge. In the Indo-European languages of (northern) India, there is also a POSTALVEOLAR articulation, while, in many of the

non-Indo-European languages of (southern) India, the articulation is even farther back, ie APICO-PALATAL. It is helpful to be able to indicate all of these articulations clearly: respectively, [t, t̪, t̺, t̻].

The *ch* in *check* is defined as POSTALVEO-PALATAL PROTRUDED: the principal contact with the blade occurs in the POSTALVEOLAR ZONE, in accompaniment with a (secondary) PALATAL COARTICULATION, with the center of the tongue moved towards the palate, and with contact along the sides up to the center of the palate. At the same time, the LIPS are simultaneously thrust forwards. All of this will be much more clear after examining the articulatory figures, especially those in 8 & 10-11. In Spanish, and also in the broader regional accents of northern Italy, the lips are not protruded. In this case, the sound is only POSTALVEO-PALATAL, and the difference can be usefully indicated by employing two distinct symbols: [tʃ, tʃ̟] respectively.

The ⟨palatalized⟩ (or ⟨soft⟩) *t* in Russian is formed by contact between the pre-dorsum and the PREPALATE ([tʃ]); while in other languages, the contact is between the (medium)dorsum and the PALATE – for example, [gʲ] in Spanish *ya*, and [c] in Greek *kēkīdi* (κηκίδι). The English *k* and *ck* in *key* and *lucky* is articulated with the PREVELUM; while the *c* of *court* involves contact between the postdorsum and the VELUM. In Arabic, there exists a sound which is still farther back, [q], produced by contact between the postdorsum and the UVULA, as in *suuq*.

The vocal folds

4.1.7.1. The articulations of *t* and *d* in *two* and *do*, of *ch* and *j* in *chain* and *Jane*, and of *f* and *v*, in *fan* and *van*, are the same within each pair. In *t* and *d*, the tongue tip blocks the expiratory air, forming an obstacle by being placed against the alveolar ridge. In *ch* and *j*, air is blocked between the blade and the postalveolar area. Finally, in *f* and *v*, the contact is between the lower lip and the upper teeth; this contact does not, however, completely block the air, which manages to pass through the interstices between the teeth. Therefore, within each pair, the articulations are identical. Yet the sounds in each pair are nonetheless different to such an extent that they are the sole means of differentiating between the words in each of the pairs given above (these sounds constitute therefore different *phonemes*).

The difference lies in the involvement, or lack of involvement, of vocal fold vibration (as we have already mentioned, the vocal folds are also called, less exactly, ⟨vocal cords⟩; they are indicated with the number 13 in fig 4.2) caused by the passage of expiratory air, which adds *voicing* to a given articulation. In fig 4.2.A, it can be seen (13) that the folds, corresponding to 1-, and the arytenoids, corresponding to -3, are situated along an almost horizontal line in the front part of the larynx. The vocal folds are membranes similar to two thin rubber bands, and are found in the larynx at the height of the ⟨Adam's apple⟩ (note that despite the picturesque name, women have one, too). They are joined together in the front, while each one is firmly connected to one of the arytenoid cartilages in the back. The arytenoid cartilages can be widened or narrowed. Therefore, the vocal folds can be

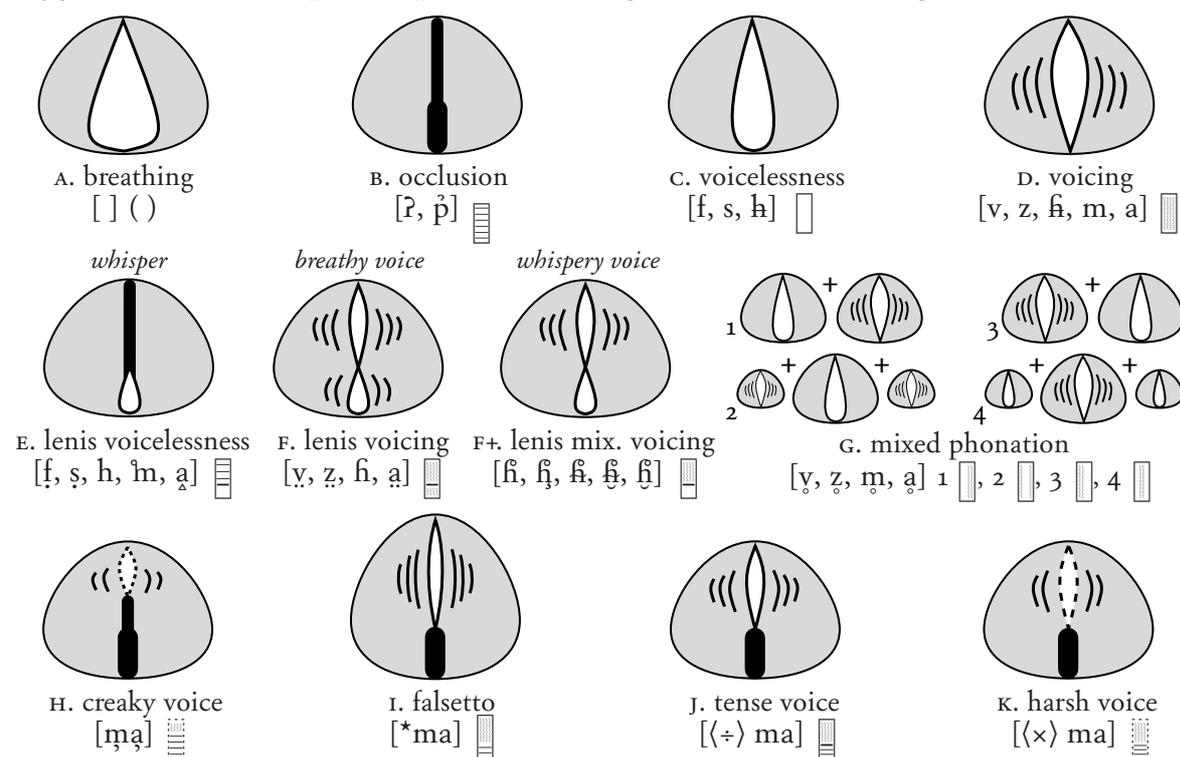
open, assuming a position similar to an upside-down V (or rather, a *lambda*: Λ); or closed along their whole length, resulting in a shape like that of an I.

The space between the open folds is called the GLOTTIS; for this reason, *glottal* and *laryngeal* are generally synonyms, even though LARYNGEAL is usually the preferable term. In fig 4.4, the folds constitute the upper $\frac{2}{3}$, and naturally the arytenoids are the remaining lower $\frac{1}{3}$. When the two of them are joined together, the folds are indicated by a thinner and longer line than that representing the arytenoids.

On average, male vocal folds are 2 cm long (and vibrate more than 100 times a second – 100 % [\times cycles per second]); female vocal folds are 1.5 cm long (and vibrate more than 200 times a second – 200 %); those of children are 1 cm long (and vibrate almost 300 times a second – 300 %); while those of newborns measure less than 0.5 cm in length (and the cry of a baby vibrates more than 400 times a second – 400 %).

4.1.7.2. fig 4.4 shows various positions of the glottis: *breathing* (maximal opening, A), full *occlusion* (maximal closure, B, [ʔ]), *voicelessness* (the maximal opening which occurs in speech, C, [f, s, ʰ]), and *voicing* (maximal vibration, D, [v, z, ɦ, m, a]). There are also the two lenis types, for which the arytenoids are separated: *lenis voicelessness* (with the folds in contact with each other, but without vibration, E, [f, ʃ, h, m, a]), and *lenis voicing* (with the folds vibrating, F, [v, z, ɦ, a]). The auditory difference between [h, ɦ] (*ie* approximants) and [ħ, ɦ̥] (*ie* constrictives) lies in the fact that more air escapes in the case of the constrictives. Moreover, [ɦ̥] is *intermediate* between [h] and [ɦ], although fig 4.4.G shows non-lenited phones (cf fig 11.5, as well, and all of the examples provided).

fig 4.4. Phonation types and positions of the glottis (and their laryngoids).



There is also *mixed phonation*, which consists of the combination, in successive phases, of the two fundamental types of voicing and voicelessness. This phonation type can be seen in the combinations of smaller laryngograms given in G.1-4, [ʏ, z̥, m̥, ḁ] and [f̥]. Therefore, the *mixed* (or ‹intermediate›) type of phonation is not a true phonation type, as are the voiced and voiceless ones (either normal or lenis), but is rather a change in the course of the production of a phone. Within a single segment, the speaker moves from one type of phonation to the other, in various ways according to the context.

In English, the voiced consonants (in diphonic pairs [consisting of a voiced consonant and the corresponding voiceless one]), when followed by a pause or by a voiceless consonant, have the first half voiced and the second half voiceless: *five* /'faɛv/ [ʼfaɛv̥], *five times* /'faɛv 'taɛmz/ [ʼfaɛv̥ 'tʰaɛmz̥] (with laryngoids such as [ʁ]). In German, the diphonic voiced stops, when after a pause or a voiceless consonant, have their first half voiceless and their second half voiced (while before pause or other consonants, including voiced ones, they are neutralized as voiceless): *Bad* /'bat/ [ʼbat̥], *das Bad* /das'bat/ [das'bat̥] (with [ʁ]). In *HPr*, there is a similar figure (fig 1.1), which is a bit simplified, but nevertheless gives further details specifically on mixed phonation.

In the languages which have mixed consonants between voiced phones, the central half (ie the two quarters in the middle) is voiced, while the two quarters on both ends are voiceless. We will give examples from broad and less broad regional Italian accents, while recognizing that gradations towards the lenis type (voiced or voiceless) are also possible. Thus, in the Italian of Naples, we find: *fate* /'fate/ [ʼfaɾɾɛ] (with [ʁ]), but [-tɛ] in less broad accents; while we have: *fante* /'fante/ [ʼfanɾɛ] in the less broad accents, but [-ɛ] in the broader ones.

The strength of stress, in its true sense of syllable prominence, can also cause oscillations. This occurs, for example, in (Mandarin) Chinese, where we have /p/ [p, ɸ, ɸ]: *bàba* /'papa/ [ʼpɸ.bɸ], *bàgōng* /'pa'kuŋ/ [ʼɸɸ'kʊŋ] (with [ʁ]).

4.1.8. Additionally, we have creaky voice and falsetto, which we explain here for reasons of completeness, even though they are more technical peculiarities. In a full study, these phonation types can be left to a successive phase of exhaustive review, with paraphonics integrated as well (cf § 14.1).

In *creaky voice* (or laryngealization, fig 4.4.H, [m̥, ḁ]), the arytenoids are firmly shut, as are the parts close to the folds. In consequence, the vibration is less strong and less normal, producing a rapid series of glottal beats, as well as the impression of lower pitch and considerably higher tension with respect to normal.

In Mandarin Chinese, creaky voice is associated with the lower parts of the tones, thus with 3 ([˩]) and the ‹half 3› ([˨˨]), as well as the full 4 ([˩˩]), but not the ‹half 4›, [˨]; cf *HPr* § 11.3.3.1-5 & fig 11.10). In Danish, creaky voice is used distinctively: *mor* ‹mother› /'mʊʊʌ/ [ʼmʊʊʌ], *mord* ‹murderer› /'mʊʊʀʌ/ [ʼmʊʊʀʌ] (cf § 17.36).

In *falsetto* (fig 4.4.I, [*ma]), the folds lengthen and are strongly tensed, with little output of air, producing the impression of much higher pitch and tension than normal. Falsetto typically extends over an entire syllable and is marked with an in-

dependent diacritic placed before the syllable, not combined with the particular symbols as occurs with the other phonation types we have seen.

When falsetto is linked to the intonation, it is written after the relevant intonemes. An example occurs in Hindi, where it occurs with the interrogative (/ʔ/ [·'··*]) and with the suspensive (/;/ [· _ ·· *]) – in Hindi, these intonemes have their posttonic syllables in the upper parts of the middle and high bands (cf *HP* fig 10.5).

For *paraphonic* purposes, *creaky voice* is marked with ⟨ʔ⟩, and *falsetto* (ie ⟨false voice⟩) with ⟨*⟩. Finally, we will give the positions producing tense voice and harsh voice. In *tense voice*, fig 4.4.J, ⟨÷⟩, the arytenoids are firmly shut and do not vibrate, with an increase in the general level of tension, while the folds vibrate regularly. The result is an impression of force and pitch which are to some extent greater than normal, and a rather ⟨metallic⟩ timbre.

In *harsh voice*, fig 4.4.K, ⟨×⟩, the arytenoids are also firmly closed and do not vibrate. At the same time, the false vocal cords are lowered, thus coming into fairly tight contact with the folds (or ⟨true vocal cords⟩), eliminating –or compressing– Morgagni's ventricle (fig 4.5.B). As a result, the folds vibrate less freely and there is generally greater tension, giving the impression of quite a bit of effort. In *trembling voice*, ⟨ʔ⟩, the pitch oscillates very quickly within single syllables.

In *whisper*, fig 4.4.E, ⟨_Δ⟩, the folds are in contact and the arytenoids are separated; both of them do not vibrate. In *breathy voice*, or lenis voicing, fig 4.4.E, ⟨·⟩, the arytenoids are open and the folds touch; both vibrate, giving an impression of lower pitch and less tension than normal. In *whispery voice*, fig 4.4.F+, ⟨[∞]⟩, the folds touch and vibrate, while the arytenoids are separated and do not vibrate. Here, the impression is of slightly higher pitch and greater tension than normal.

4.1.9. The vocal folds can be completely closed, so that expiratory air is blocked and no vibration can occur (fig 4.4.B). When people make a physical effort of short duration, such as lifting a heavy weight, they tend to hold their breath, closing their vocal folds tightly. Afterwards, when the folds are opened, the air which had been compressed escapes abruptly, producing a LARYNGEAL STOP (or ⟨glottal stop⟩).

This sound, represented by the symbol [ʔ], is also heard in coughing. Given the tight closure of the folds, it is neither voiced, nor voiceless, while all other sounds are voiced or voiceless, or at most intermediate between them (as we will see). Therefore, even though it has no ⟨voicing⟩ (produced by vocal fold vibration), it would not be entirely appropriate to speak of voicelessness, which is characteristic of the position with the folds opened (fig 4.4.C). For this reason, it makes sense to place the laryngeal stop in a position –in the tables– halfway between voicing and voicelessness, even though it could be considered more like a voiceless sound, especially from an auditory point of view.

As we have just mentioned, the vocal folds can also be kept separated (fig 4.4.C), with the result that air passes freely through the glottis without making the folds vibrate. The resulting sounds are VOICELESS, such as [p, t, k, tʃ, f, θ, s, ʃ] in *pen, ten, Ken, chain, fane, thane, same, shame*.

It is also possible for the folds to be held gently close to one another (fig 4.4.D) so that when air passes, they vibrate. Sounds produced in this manner are VOICED

– examples include the English vowels (and the vowels in most languages), and the consonants [b, d, g, dʒ, v, ð, z, ʒ] in *big, dig, get, jet, vat, that, bays, beige*. Moreover, the following are usually voiced: [ɹ/ɹ̥, m, n, ŋ, l, ʎ, j, w], as in *ram, ran, rang, Billy, Bill, yet, wet*.

4.1.10. It is very important to learn to distinguish clearly voiced sounds from voiceless ones. For this purpose, it will be helpful to practice controlling the <voicing> used by repeating out loud, and without interruptions, prolonged sequences of continuous consonants of the same articulatory position, while adding and removing vibration of the vocal folds. For example: [fff vvv fff vvv fff, sss zzz sss zzz sss] (also representable as [f:: v:: f:: v:: f::, s:: z:: s:: z:: s::]), &c, and vice versa. While doing this, one can and should check the presence and absence of voicing by covering the ears with the hands – at this point, the voiced parts will produce a strong *buzz*.

If one places the palm of one's hand on the throat, the voiced parts produce *vibration* there which is equally perceptible (by touch). It is reasonable, therefore, to put one hand on the ear and the other on the throat, so that both the buzz and the vibration can be felt during the voiced period; during the voiceless period, instead, neither will be present. Only voiced sounds can be sung, and for this reason, if a musical note can be produced while pronouncing a sound, the sound is undoubtedly voiced.

Besides the presence or absence of <voicing>, voiceless and voiced sounds, such as *p* and *b* [p, b], differ in another way which helps to distinguish them. Voiceless sounds are emitted with greater energy, and air escapes with greater force, since the glottis is open. Instead, during voiced sounds, the air makes the folds (which are closed) vibrate, and thereby the air loses a part of its force.

4.1.11. When breathing normally, the glottis reaches a position (A) which is open still more than the position for voiceless sounds, as shown in fig 4.4. The position (E) is what occurs while <whispering> (or rather, and better, in VOICELESS LENIS phonation) in the place of the voiced sounds (D), while the voiceless sounds remain unchanged (C). As a matter of fact, while whispering, there remains a fairly perceptible difference between voiceless and voiced phonemes for this very reason. Thus, in English, *simple, riches, refuse* remain different from *symbol, ridges, reviews*, &c.

The position (F) is that of <breathy voice> (or rather, of VOICED LENIS phonation), in which the folds are brought close to one another, as with voiced sounds, but the arytenoids are kept apart, as while whispering, but both of them vibrate. Therefore, this position is *intermediate* between voiced and voiceless, since the air causes the folds to vibrate only partially. As a result, the vibrations are less energetic than in voiced sounds. In addition, there can also be vowels with mixed phonation, especially in unstressed syllables which contain the sound (often referred to as <indistinct>) of the central vocoid [ə]: we thus have [ə̤], partially devoiced.

4.1.12. To change one's voice PITCH, the frequency of vibration needs to be changed. This involves tightening the vocal folds, to reach a HIGHER pitch, or loos-

ening them, for a LOWER pitch. The mechanism is the same as that of a guitar string, for which the vibrating portion is lengthened or shortened by pressing the strings down on the handle in points which are different distances away from the headstock.

While whispering, it is possible to maintain some distinctions among different tones and also different intonation groups – if instead these characteristics should come to be lost, communication would be seriously compromised. An interplay of tension of the articulations and the larynx helps to maintain auditorily the differences in pitch, by providing a partial compensation for the lesser distinctions in <tonality> proper.

Resonators (5 phono-articulatory cavities)

4.2.1. Considering in greater depth the PHONO-ARTICULATORY APPARATUS, we can say that it determines the sounds and their specific timbres. Expiratory air is essential for nearly all the sounds used in the languages of the world. In any case, it is the only way to produce full words and sentences. It is produced in the pulmonary cavity, and becomes variously characterized by passing through the FIVE CAVITIES, which constitute the phono-articulatory apparatus. These are, in the order in which they are encountered by the expiratory air: the LARYNGEAL, PHARYNGEAL, ORAL, LABIAL, and NASAL cavities (fig 4.2.C, respectively, *d, c, b, a, e*; while normally, the orograms are meant to be examined moving from what is most external [ie labial, oral, and nasal], to what is most internal [ie pharyngeal and, visible only in (more) extended orograms, laryngeal], down to the pulmonary and abdominal cavities, visible in fig 4.3).

The LARYNGEAL CAVITY (fig 4.2.C.d) determines both articulatory and phonatory characteristics. In fact, the phones [h, ħ, ʔ] can be considered both as MANNERS OF ARTICULATION (namely, the approximant –[h, ħ]– and stop –[ʔ]– manners), in the laryngeal place of articulation, and as particular TYPES OF PHONATION, or states of the glottis. However, considering that many languages use these phones as segments, which can be distinctive, it is better to include them with the consonant elements, and to put them in the consonant table.

Moreover, the laryngeal cavity also produces the various types of phonation (cf § 4.1.7-8). The other phones, which have their own articulation determined by the other four non-laryngeal cavities, also have phonation as a constituent part.

fig 4.5. Vocal folds (or <vocal cords>, 3) and <>false vocal cords> (1). The space in between is Morgagni's ventricle (2).



The activity of the laryngeal cavity is also fundamental for the (prosodic) production of pitch and its variations, which are determined by differences in tension,

lengthening, and vibration.

In addition, the shape and dimensions of Morgagni's ventricle (ie the space between the true and false vocal folds, cf fig 4.5) determines a good part of the individual characteristics of the voice; these characteristics are then accentuated *or* modified by the supra-laryngeal cavities.

4.2.2. The PHARYNGEAL, ORAL, and LABIAL cavities (fig 4.2.c.c-b-a) constitute together a resonator of extremely variable shape. In fact, these three cavities can operate together, to determine and accentuate a particular effect, or they can operate contrary to one another, to some extent neutralizing some of their individual effects.

The resonance produced is different depending on whether a particular cavity is wide or narrow. For example, the back of the tongue can be raised towards the uvula, or the root can be pushed backwards towards the pharynx. The result is a narrowed area in the pharyngeal cavity, but a widened one in the front part of the oral cavity. Or instead, the tongue can be moved towards the prepalate, thus widening the pharyngeal cavity and the back part of the oral cavity. When the lower and upper articulatory organs are in full contact, all of the area to the front of the articulation is excluded from participating in the formation of the sound.

The LABIAL CAVITY helps to determine the timbre of sounds, according to the shape of the lips. Instead of being in their <normal> position, they can also be ROUNDED (or *round*), or PROTRUDED forward; in these cases, the articulatory channel constituted by the pharyngeal and the oral cavities is lengthened. On the other hand, the lips can be SPREAD, thereby reducing the length of the channel and producing contrary effects.

4.2.3. The articulatory channel is constituted by the laryngeal, pharyngeal, oral, and labial cavities. The resonance induced by the various shapes it assumes can be further modified by lowering the velum, thereby adding the activity of the NASAL CAVITY (fig 4.2.c.e). The air, in this case, escapes from both the mouth and the nose, or possibly only from the nose, if the oral or labial cavity is closed.

The different articulations are produced by touching, or by moving close together, two (or more) points along the two parts composing the articulatory apparatus. One of these is the UPPER PART, which goes from the upper lip, along the palatal vault, to the pharyngeal wall, and even to the arytenoid cartilages in the back part of the larynx. The other is the LOWER PART, which goes from the lower lip to the (tip, back, and root of the) tongue, and when relevant, all the way to the vocal folds, in the front part of the larynx.

For the various possibilities, see the chapters and figures which treat the different segmental articulations.

The lips

4.2.4. As we have already mentioned briefly, the lips are very important in the articulation of many vowel and consonant phones (vocoids and contoids, cf § 6.1.1-2, § 6.2.1-6.3.7 & ¶ 8-10). They can assume several fundamental positions, which can be combined with different levels of jaw opening as well (cf fig 8.1-2, fig 8.6-7, fig 8.8 & fig 8.9).

There is the position with NEUTRAL lips, unmarked (or *normal*, fig 8.9), but equally functional, by virtue of its opposition to the other marked positions. Moreover, the lips can be SPREAD (also in fig 8.9) by pulling back the corners of the mouth, as in the more front vowels, such as the [i] in *bee*. Apart from rounded (cf § 4.2.5), they can also be PROTRUDED, by moving them outwards and away from the teeth; this occurs with certain contoids, such as [ʃ] in *she* (cf fig 9.1). In this case, we do not have rounding, but rather protrusion, and as a result this is the PROTRUDED position (cf § 11.3 and fig 9.1), characterized by the lips being noticeably extended outwards.

As will be seen further ahead, in the phonetic reality of the languages of the world, and in the *canIPA* alphabet, [ʃ] is the unmarked articulation. This remains true even though the lips are protruded, because this particular articulation is most commonly found in languages with protrusion present. It is related to two other articulations with differing lip positions: [ʃ̥], without protrusion, and [ʃ̠], with over-rounding (by far the least frequent). This series, [ʃ̥, ʃ, ʃ̠], is analogous to [ɸ, ɸ̥, ɸ̠] (including the corresponding voiced consonants and stopstrictives), but not entirely to [ɸ, ɸ̥, ɸ̠], in which the central element of the series is bilabialized and the third one is rounded.

As for the lip positions *in* the orograms, it is important to take care not to confuse the various types. The reader should also consider fig 8.3, as well as fig 9.1, where an external prospective (a profile view) is given, without the schematic simplifications found in other diagrams.

In the orograms, what is mostly seen is the central part of a sagittal cross section (from the center of the lips, back to the uvula and thence to the larynx). Since these diagrams already contain an extensive quantity of information, it was necessary to exclude the nuances involving the corners of the mouth to avoid compromising their usefulness by encumbering them excessively.

Returning to the particular case of [ʃ̥, ʃ, ʃ̠], we wish to emphasize that [ʃ̠] is also characterized by lip protrusion, which is very evident (see the lower part of fig 9.1). The non-protruded [ʃ̥], on the other hand, has no protrusion or rounding (but nonetheless a different position from [j], where the lips are spread and the corners of the mouth are pulled back). Here the fact that the teeth are very close together also causes the lips to be quite close to one another – but this should not be taken to imply that an active use of the lips is involved.

Therefore, when we specify that a certain *C* is *unrounded*, we simply wish to make clear that the lip position is normal (ie indifferent). This does not necessarily imply that the position is the perfect opposite of the rounded or protruded position (those positions instead need to be indicated exactly). If ⟨unrounded⟩ *C* are

followed by rounded phones, then they will also have a labial coarticulation which cannot be ignored. However, this coarticulation also occurs with rounded or protruded C, where it is added to the labialization already present. As a result, the proportion distinguishing the groups remains the same.

Instead, in the overrounded [ʃ], the rounding is accentuated, but not the protrusion, so that the lips are pushed outwards less. Besides, they are only a bit closer together than in the neutral position, of [ʃ̄], where the jaws are so tight that the teeth almost touch. Comparing the labiograms given on the right-hand side of fig 9.1, we see that the lip position of [ʃ̄] is quite close to that of [s]. The only difference is that, in this last case, it is possible to glimpse the tongue through the teeth, since the tongue is farther forward than in the lamino-postalveolar position.

4.2.5. The **ROUNDED** (or *round*, fig 8.9) position normally has a bit of simultaneous protrusion as well. This protrusion can be seen in the /uu/ in *blue*, and in a few contoids, such as [w] in *web*. The lip position here is virtually the same as that for [ʃ]; this last contoid is considered <overrounded> because, contrary to the case of vocoids, contoids usually have little rounding.

More rarely, there are rounded phones which are not protruded – ie, the lips are not extended forwards, but remain close to the teeth (ie **VERTICAL ROUNDING**, fig 8.9). An example of this phenomenon is seen in the vocoid [ɣ] in Norwegian *ny* [nɣ]. Another instance is partial rounding, by which we mean less rounding than would normally be expected for a given level of jaw opening. This occurs in certain languages and dialects, such as the English [ɔ] in *good* [ˈɡʊɔ̄d]. In cases where it is useful to indicate this fact, the term **HALF-ROUNDED** (or *half-round*, fig 8.9) can be used, and the symbol can be written with a diacritic mark, as in [ˈɡʊɔ̄d̄, ˈɡʊɔ̄d̄], *good*.

The **LIP POSITIONS** –ie **NEUTRAL**, **SPREAD**, and **ROUNDED**– can be combined with different levels of **JAW OPENING**, ranging between **CLOSE** and **OPEN** positions, corresponding to the six levels of opening of the vocoids (again, cf fig 8.1-3, fig 8.6-7, fig 8.8 & fig 8.9).

It is important to keep **ROUNDED** articulations (where the rounding is secondary, even though it is distinctive in many languages – examples include [ɥ; w; ŋ; t̪, d̪] and [y, u; ø, o; œ, ɔ]), or **BILABIALIZED** (as [ɸ, ɹ; dɸ, tɸ], with *vertical* rounding), and **PROTRUDED** ones (with *secondary* protrusion, such as [ʃ, ʒ; tʃ, dʒ]) carefully distinguished from **LABIAL** articulations. For these, the lip position may be the *primary* articulation, such as in [m; p, b; ɸ, β], or a *simultaneous* and *equal* articulation, as in [ɱ; kp, ɸ], where two different articulations –ie [ɱ; k, g] & [m; p, b]– are joined together in a single sound.

5. The classification of sounds

5.1. It is not possible to learn a sport, an art, or a profession, without knowing the various skills involved, as well as the technical terms which refer to them. In the same manner, it is not possible to become familiar with the great quantity and diversity of sounds used in different languages, without knowing how these sounds are produced, how they are classified, and the technical terms that competent linguists and phoneticians use to refer to them.

Every linguistic sound is a member of one of two principle classes, traditionally called VOWELS and CONSONANTS. These terms will be used here only in the traditional sense, or in the technical, more scientific one, which is proper to FUNCTIONAL PHONETICS (ie phonology, phonemics, or phonematics), with reference to the value that a given sound has with respect to others nearby.

Therefore, phonetically, we will use two different terms, so that no doubts or ambiguities can arise between the various levels of analysis of a particular language: VOCOIDS and CONTOIDS.

5.2. We can define VOCOIDS, in articulatory terms, as those sounds, generally voiced, which are produced by air which does not encounter any obstructions by the articulatory organs, even partial ones (besides the vocal folds), nor constrictions great enough to produce air friction. At the same time, the articulatory position is maintained for sufficient time so as to render the sound stable and clearly perceptible (and capable of lengthening, if necessary). The characteristic sound of a vocoid mainly depends, in practice, on how the tongue is raised or lowered and moved forwards or backwards (and also on whether the lips are rounded). For the mechanism of raising and lowering the back of the tongue, the movement of the lower jaw is clearly fundamental (given that the base of the tongue is solidly fixed to the lower jaw). For this reason, while raising or lowering the back of the tongue (shown in the vocograms), we have two complementary movements, which together make the articulations and transitions between different sounds easier. In reality, it would be possible to communicate fairly well even if we spoke with our teeth tightly clamped – it is certainly not very comfortable, but it is easy to give it a try with a given word or sentence...

All the <vowels> in the following examples are VOCOIDS: *city*, *let*, *all*, *son*. There are two VOCOIDS in the <diphthongs> occurring in words spelled either *no*, *be*, *my*, or *now*, *boy*, *day*. Certain grammatical traditions call the diphthong (present in words such as Spanish or Italian *pausa* /au/) a <falling diphthong>, because there is greater intensity on its first element. The same (antiquated) grammatical traditions also call a <rising diphthong> the sequence in Spanish or Italian *piano* /ja/, because the vocalic element has greater intensity.

5.3. From a phonic point of view (both phonetic and phonological), however, [ja, wa] cannot be considered true <diphthongs>, given that they consist of a consonant element, [j, w], followed by a vowel, [a] (or any other vowel), just like sequences such as [ka, sa, la, ma].

Generally speaking, vocoids constitute the nucleus, or center, of the syllables containing them. This routinely occurs when a vocoid lies between two non-vocoids (ie consonants). Instead, when two (or more than two) vocoids come together in a single phono-syllable, the first vocoid inevitably constitutes the syllable nucleus: *my* ['maɹ] (English), *tower* ['tʰaʊə] (British English), *mai* ['ma:i] (Italian). However, in *quit* ['kwɪt] (English), or *può* ['pwɔ] (Italian), there is a single syllable with a single vocoid, respectively [ɪ, ɔ]. In fact, [w] is a contoid, as will be seen below (§ 5.5-6 & fig 5.1). On the other side, we find that Spanish *pie* /'pje/ ['pjɛ] sometimes is pronounced as ['piɛ], or that fairly systematically, in the broad regional accent of Naples, the Italian word *può* /'pwɔ/ ['pwɔ] is pronounced as [p'u'o], which has two vocoids and two phono-syllables, in opposition to having only one grapho-syllable in the official interpretation. From a phonic point of view, also Spanish *día*, *buey* /'dia, 'bwei/ are monosyllabic.

In the same way, there is only one phono-syllable in Italian *miei* ['mjɛ:i], *mia* ['mi:a]. But grammatical tradition, which constantly seeks to force things excessively into fixed schemes, considers *miei* ['mjɛ:i] monosyllabic and *mia* ['mi:a] polysyllabic! In fact, uncritical consideration of the grammatical structure –while following a method lacking scientific objectivity and rationality– leads to the grave error of identifying grammemes (ie grammatical morphemes) with syllables. The words in question are divided grammatically as *mi-a*, *mi-o*, *mi-e*. For this reason, grammatical tradition, equipped only with a few confused concepts (absurdly based upon writing conventions!), is obliged to <invent> a bisyllabic nature for these words, distinguishing them in this manner from *miei* (since it does not have alternations with similar forms), which is monosyllabic indeed. All of this occurs despite the <bulkier> appearance of *miei*.

Of course, while in *bit* ['bɪt], *bet* ['bɛt], *bat* ['bæt], &c, we have true vowel elements, and true diphthongs in *beat* ['bi:t], *bait* ['bɛɪt], *bite* ['baɪt], we can certainly not say that *beaut* ['bjʊt, 'bjʊt] contains a diphthong, as far as [jV] is concerned, just as in *yet* ['jɛt], or [wV] in *wet* ['wɛt]. On the other hand, *beaut* and *you* (with [jʊu, 'jʊu]) do have a real diphthong as far as their total vowel elements are concerned: /ʊu/ [ʊu/ɥu].

5.4. All sounds lacking the characteristics of vocoids, instead, are called CONTOIDS. This definition is not tautological at all – rather, it is fully scientific, since *x* and *non-x* are very important categories.

Consequently, the <consonants> in *take* and *Dutch* are contoids, given that they completely block the passage of air through the articulatory apparatus. The English *l*'s in *lill* ['lɪl] are also contoids, because they block the air partially. Nasals, such as *m*, *n*, *ng* in *man*, *song*, are contoids, because the air does not escape out of the mouth, but rather through the nose, due to the lowering of the velum.

We also have contoids in the case of sounds which, while not blocking the pas-

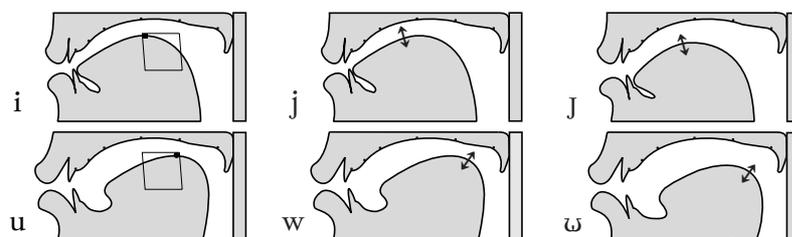
sage of air out of the mouth, are produced by moving articulatory organs close together in such a way that a constriction is created which in turn generates friction. This friction can be heard clearly both in voiceless and in voiced sounds, thus for example in the [f, v, θ, s, z, ʃ] in *five, these, seize, shoes*.

Sounds produced by moving the articulatory organs together close enough that (a lesser degree of) friction can be heard in voiceless sounds, are also contoids (as, for example, [h, ɦ] in *hat*). In the corresponding voiced sounds, this light friction is covered up by the voicing (which is produced by vibration of the vocal folds), as in a frequent variant pronunciation of *h* [ɦ] (occurring between voiced sounds) as in *behind*, or as [β, δ] in Spanish *abonado* [aβo'naðo].

5.5. This last group (ie the approximants) also includes sounds such as [j, w] in *yes, pure, win, quite*, which are clearly consonantal, but have an articulation corresponding –or, in the case of English, fairly similar– to particular high vocoids, such as (at least one part of) the <vowels> in *me, do* ['mɪi, 'dʊu/'dʒu]. These sounds are shorter in length than [i, u], and, even more importantly, the position is less static (more mobile) during their articulation – although, as we have seen, in *me* and *do*, we actually have two real diphthongs: [iɪ, uɪ/ʊɪ]. In the case of Spanish *quién* or Italian *ieri, uomo*, we have ['kjen; 'jɛ:ri, 'wɔ:mo], respectively.

The orograms of [j, w] could seem <static>, just like those of [i, u] – this is due to technical constraints on iconic representation. In fact, they are nearly identical, except that the tongue back is slightly higher and more peripheral (more towards the front for [j] and more towards the back for [w]). The difference in question is due to the fact that the jaw is in movement, and will shortly become more open in transition towards the next phone (and, naturally, the lip rounding of [w] also lessens as the jaw opens). As a matter of fact, [j, w] are fundamentally movements (corresponding to [i, u]), which pass through the characteristic positions of [i, u], initially surpassing them slightly (as indicated above), and which then move back to positions a bit below them. Because of these considerations, it is fully reasonable to consider them to be the <same> articulation (as they actually are), in terms of the position (ie height and frontness) of the tongue. The essential difference is that [i, u] are naturally static, while [j, w] are dynamic (this is shown by the double arrows, in fig 5.1).

fig 5.1. Articulatory comparisons between the vocoids [i, u] and the contoids [j, w] (& [ɹ, ʊ]).



5.6. Therefore, fig 5.1 provides a comparison between the articulations of [i, u] and of [j, w]. The differences indicated above are (more) evident if one juxtaposes the relative figures (perhaps after enlarging them and making slides out of the printed figures). In any case, it is a good rule to show orograms for vocoids togeth-

er with their vocograms (the slightly tilted quadrilateral at the center of the figure). Normally, the vocoids of a particular language are most effectively treated by using only vocograms, suitably enlarged (as will be seen later). It is better to reserve the use of orograms, with vocogram inserted in the middle, for the general presentation of all possible vocoids (as in \mathcal{G} 8).

The position shown in *fig 5.1* refers to the basic articulation, which occurs in particularly careful and precise pronunciation. Examples include *yeast* [ˈjɪɪst], *womb* [ˈwʊm, ˈwʊm] (the first form is British, the second American, cf *HPr* \mathcal{G} 2 of *HPr*), or Italian *baia*, *kiwi* [ˈba:ja, ˈki:wi]. In normal speech, especially when the nearby vowels are not high, the articulation is characterized by less evident tongue elevation. However, there is no need to employ different symbols, since the results are still sufficiently different from the two <attenuated> articulations which may correspond to /j, w/ – we represent these with [j] (semipalatal) and [w] (semivelar rounded), and have given them on the right-hand side of *fig 5.1*.

5.7. We will add a pair of diagrams, which will be useful for becoming more familiar with the articulatory apparatus and improving one's perception of sounds. *fig 5.2* shows the position of *rest*, when a person does not speak, and breathes through the nose.

fig 5.2. Breathing with the mouth closed.
Position of the articulators at rest.

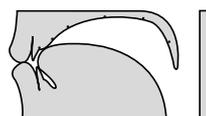


fig 5.3 exemplifies the use of *handphones*, or <hand-headphones>. We refer to an instrument which is very useful for hearing better the sounds that we ourselves produce, whether we are concentrating on vowels or consonants, or on pitch and intonation. It suffices to place one's hands so that they produce a sort of channel, with one hand in front of the mouth (while leaving some room in front of the lips), and passing in front of the cheek to the point where that hand connects to the other hand, which covers the ear (leaving some room here, too, so that the sound wave will not be smothered). The purpose is to funnel what we say into this <channel> so that the sounds produced can be heard more clearly. It is definitely more difficult to describe this useful <instrument> than to put it into action – it is usually enough to give it a couple of tries, guided by *fig 5.3* as well.

A similar though somewhat less intense result can be obtained by putting an obstacle such as a board or a fairly large book about four inches in front of the mouth. In this way, the sound wave will rebound towards the speaker (instead of being propagated forward and lost), and will therefore arrive better at the ears.

fig 5.3. The use of *handphones*
for hearing the reader's
own sounds better.



6. A gradual approach

6.0. In this introductory chapter, we will present the fundamental categories, with a simplified treatment limited to the most basic elements. These categories constitute the minimum necessary to proceed scientifically with phonetics (even though the information given in the preceding chapters is nevertheless rigorous).

In what follows, every part will be gone into in greater depth and with added detail, helping the reader to arrive at a more complete knowledge of the subject.

Vowels

6.1.1. The back of the tongue is the fundamental element in vowel production. It moves in two different directions: HIGH–LOW and FORWARD–BACK. Consequently, the combination of these two elements produces a *quadrilateral*, which gives us the fundamental *VOCOGRAM*, used for showing –inside it– the positions of vowels of a given language. On the left side of fig 6.1.1, there are three orograms indicating the zone of vocoid articulations; these orograms are steadily more schematic, moving downwards. The first, on top, is the most realistic, while the third, at the bottom, is a quadrilateral.

On the right-hand side of fig 6.1.1, the upper diagram is an orogram which shows the tongue LOW and CENTRAL, as in the pronunciation of *a* [a] in most languages. The upper outlines of the positions of *i* [i], HIGH and FRONT, and *u* [u], HIGH and BACK, are also given – as they occur in most languages. The points are connected and contained in the white (or transparent) quadrilateral, which is given enlarged in the figure below (the *vocogram*, on the lower part of the right-hand side).

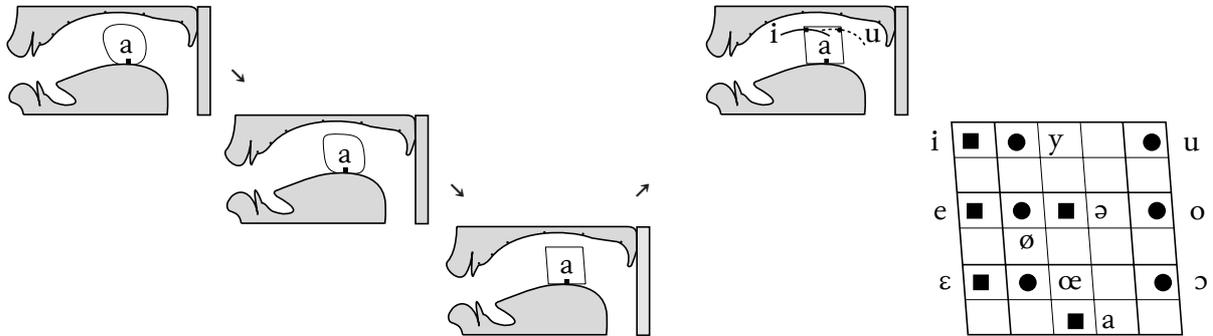
In the large quadrilateral, 11 vowels have been placed, shown by (square and round) MARKERS. The ROUND ones refer to vowels articulated with ROUNDED lips, while the SQUARE ones naturally represent vowels with unrounded –either SPREAD or neutral– lip position.

The symbols [i, a, u] correspond to Spanish *i, a, u*, as in *utilidad* /utili'dad/ (or Italian *utilità* /utili'ta/), while [e, o] are the <closed> vowels of Portuguese, as in *vê, povo* /'ve, 'povu/ (or Italian *tre, sono* /'tre, 'sono/); [ɛ, ɔ] are the (stressed) <open> vowels of Portuguese, as in *pé, pó* /'pɛ, 'pɔ/ (or Italian *sette, otto* /'sette, 'otto/). Note also German *Kamm, Tag* /'kam, 'tak/, *viel, Kuh* /'fi:l, 'ku:/, and –but closer– *Weg, Boot, weg, Loch* /'vek, 'bo:t, 'vek, 'lɔx/. The Italian words written *corressi* and *volto* have two different meanings corresponding to two different pronunciations: (*se*) *corressi* <(if) I ran> /kor'ressi/, and (*io*) *corressi* <(I) corrected> /kor'rɛssi/; (*il*) *volto* <(the) face> /'volto/, and (*io*) *volto* <(I) turn around> /'vɔlto/. Consequently, the two GRAPHEMES <*e, o*> can each represent two different phonemes: /e, ɛ/ and /o, ɔ/.

The vowels of English, Italian, French, German, Spanish, Portuguese, Russian,

Arabic, Hindi, (Mandarin) Chinese, Japanese, and Esperanto, can be seen in \mathcal{C} 2-13 of *HPr*. The Italian vowels can also be seen in *M^aPI* and *DⁱPI*. The vowels of more than 300 languages and dialects can be found in \mathcal{C} 16-23 (here).

fig 6.1.1. The articulatory extent of vowel sounds.



6.1.2. fig 6.1.1 (the vocogram part) contains three more vowels /y, ø, œ/, which are rounded, and for this reason have circular markers. These vowels are *almost* like /i, e, ε/ with lip rounding added. However, the tongue is a bit farther back than it is in /i, e, ε/, and in fact, these rounded vowels are a little centralized in the vocograms. /y, ø, œ/ occur in many languages, such as French: *lune*, *deux*, *seul* /'lyn, 'dø, 'sœl/, or German: *Füße*, *Öl*, *zwölf* /'fy:zə, 'ø:l, 'tsvœlf/ (as well as in several Italian dialects, particularly Lombardian, Piedmontese, and Ligurian). The first German example also has an instance of [ə], which is generically placed in the center, at the height of [e, ø, o] (cf fig 6.1.1). However, ⟨[ə]⟩ has many different realizations in the different languages, which we will render with more appropriate symbols in the relevant phonosyntheses (\mathcal{C} 16-23, and of course in *HPr*). These can include taxophonic variants in languages such as English and German (\mathcal{C} 2 & \mathcal{C} 5 of *HPr*).

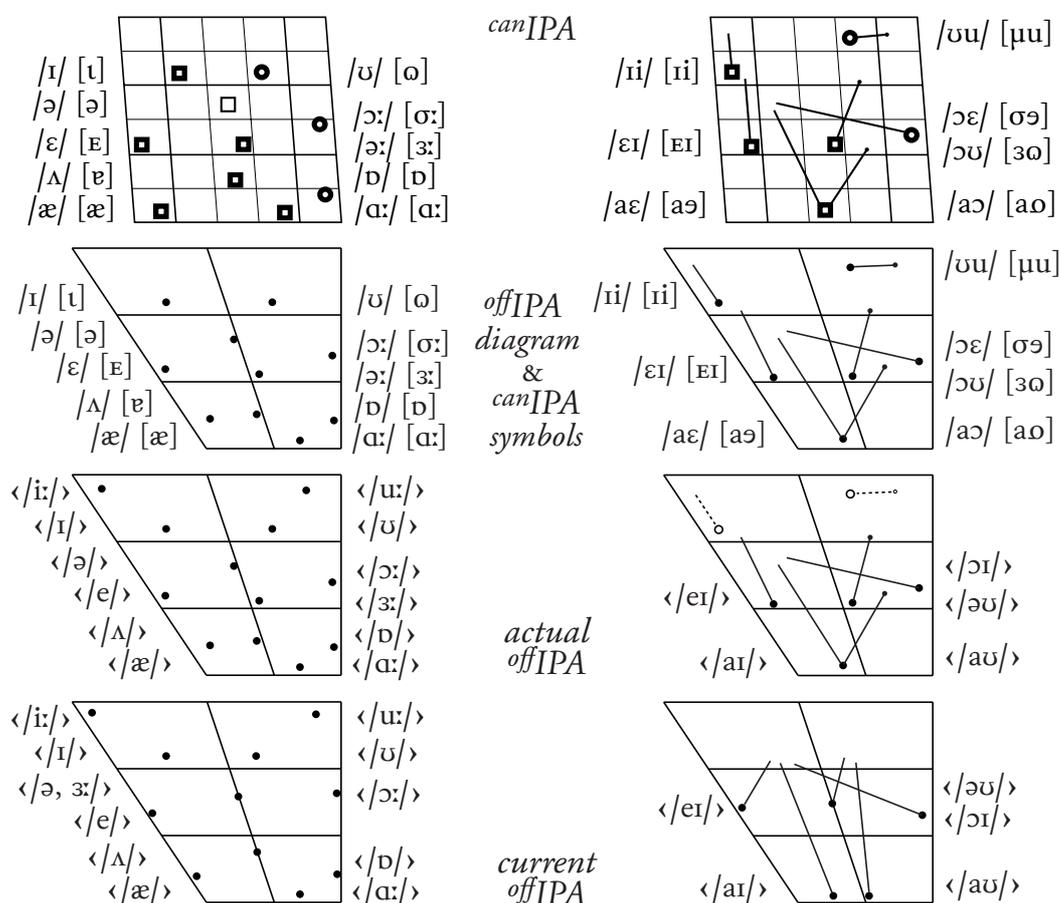
The symbol /'/, (an uncurved apostrophe) placed immediately before a syllable, indicates STRESS. The CHRONEME, /:/, indicates distinctive lengthening of the preceding vowel – for example, in German there is a contrast between *Stadt* /'ʃtat/ <city> and *Staat* /'ʃtat:/ <State>. When the same symbol occurs in PHONETIC TRANSCRIPTIONS (in brackets, [], instead of in PHONEMIC TRANSCRIPTIONS, which are written between slashes, / /), it is called a CHRONE, and indicates length which is not distinctive. An example of non-distinctive lengthening is that occurring in Italian word-internal stressed unchecked syllables: *seme*, *solo* ['se:me, 'so:lo].

In conclusion, vowels consist of three fundamental elements: RAISING (of the tongue and jaw), ADVANCING (of the back of the tongue), and lip ROUNDING (or its absence).

As a first approach to the vowel phonemes of English, which are many more than in Spanish (5) or in Italian (7), we reproduce a simplified version of the vocograms of neutral British English (from \mathcal{C} 2 of *HPr*), showing only its monophthongs (9 + *schwa* /ə/) and diphthongs (7), with no combinatory variant, and excluding centering diphthongs as well (which, of course, are not left out in *HPr*). This is done to enable the comparison with other similar figures currently found in phonetics or linguistics textbooks. We then present them both in our own

vocograms and in the official quadrilaterals (but keeping our symbols) for a quicker comparison (followed by both an actual and current application of *offIPA* criteria and symbols, too).

fig 6.1.2. Four versions of simplified monophthongs and diphthongs of neutral British English.



Voicing

6.1.3. Voicing is the <voice> given to vowels and certain consonants by the vibration of the vocal folds (which are located in the larynx, cf § 4.1.7-12 & fig 4.4).

Voicing can therefore be present or absent, giving rise to two principal TYPES OF PHONATION: VOICED and VOICELESS consonants.

To give a few examples, the consonants present in *man*, *ring*, *dig*, *jazz*, *these*, *leisure* are voiced: /'mæn, 'ɪŋ, 'dɪg, 'dʒæz, 'ðɪz/ and /'leɪzə, 'li:zə, 'leɪzə/. The Spanish or Italian /ɲ, ʎ/ are also voiced, and in neutral Italian pronunciation, they are always geminated between vowels, just like the consonants written doubled in the official orthography: *sogno*, *foglio*, *mamma*, *babbo*, *oggi* /'soŋno, 'foʎlo, 'mamma, 'babbo, 'ɔdʒdʒi/. However, in other languages, /ɲ, ʎ/ are generally found without gemination, as in Spanish: *mañana*, *calle* /ma'ɲana, 'kaʎe/, or Portuguese *ninho*, *filho* /'niɲu, 'fiʎu/.

6.1.4. The other fundamental group of consonants is that of VOICELESS consonants, as seen in *pack*, *teach*, *south*, *fish* /'pæk, 'ti:tʃ, 'səʊθ, 'fɪʃ/. Of course, we have *fishy* /'fɪʃi/, while in neutral Italian, /ʃ/ is geminated between vowels: *pesce* /'peʃʃe/. Gemination occurs even in foreign words adapted into Italian, such as the word *cachet* /kaʃʃe/, which in French is /kaʃe/. It is interesting to note that Italians also pronounce the orthographic geminates of foreign languages as true phonic geminates, as in the English name *Billy* /'billi/, instead of /'bili/.

Consonant gemination is distinctive in Italian, as the following examples demonstrate: *cade*, *cadde* /'kade, kadde/, *tufò*, *tuffò* /'tufo, 'tuffo/, *nono*, *nonno* /'nɔno, 'nɔnno/, *caro*, *carro* /'karo, 'karro/. In neutral Italian, there is also gemination in cases such as *è vero* /ɛv'vɛro/, *ho sonno* /ɔs'sonno/, *a casa* /ak'kaza/, *blu mare* /blum'mare/, *così forte* /koziff'forte/, *tornerà domani* /torner'ɔd do'mani/, *città balneare* /tʃit'tab balne'are/. This kind of gemination is better defined as *co-gemination*, cf § 12.14 (& ¶ 3 of *HPr*, & ¶ 5 of *M^aPI*, & § 1.6 of *DⁱPI*).

Consonants

6.2.1. We will now see how the consonants are produced. As we have seen, the articulation of vowels is determined by the back of the tongue, with its up/down movements (complemented by closing and opening of the jaw), as well as its front/back movements, and also by the possibility of lip rounding. With consonants, instead, the space available is greater. In fact, it extends from the lips all the way to the larynx (cf fig 6.2).

In the table of fig 6.2, the names across the top are the principal PLACES OF ARTICULATION, ranging from the lips to the larynx. The names to the left of the rows, instead, indicate the principal MANNERS OF ARTICULATION. Intersections between the rows and columns can then produce various consonant sounds, and the number is often doubled due to the possibility of adding voicing (ie the voiced PHONATION TYPE – cf § 4.1.7-12).

fig 6.2. Simplified table of consonant sounds.

	bilabial	labiodental	dental	alveolar	postalveolar	postalveo-palatal protruded	palatal	velar	velar rounded	uvular	laryngeal
nasal	m			n			ɲ	ŋ			
stop	p (b)		t (d)	ʈ (ɖ)				k (g)			
stop-strictive						tʃ (dʒ)					
constrictive		f (v)	θ (ð)								
grooved con.			s (z)			ʃ (ʒ)					
approximant					ɹ		j	w			h
trill										ʀ	
lateral				l			ʎ				

(symbols in brackets –or standing alone [except *h*]– are voiced)

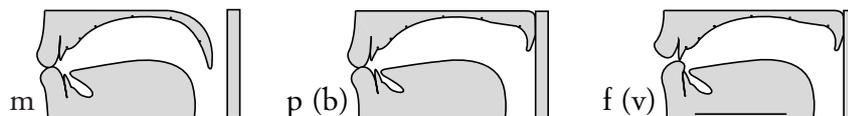
All the British English consonant phonemes are given in the table, including the voiced elements forming diphonic pairs (given in parentheses). The consonants [ɲ; t̪; (d); r, ʀ; λ] also appear; these are not found in English (and are therefore given in italics), but are very important in certain other languages. All of these articulations are given in fig 6.3-9 (and again, from another perspective, in fig 6.10-16).

Of course, in \mathcal{C} 9-10 all possible contoids are given, while all the English phonemes and phones are dealt with thoroughly in \mathcal{C} 2 of *HPr* (including American English and other accents such as –both American and British– mediatic and an ‹international› one, besides Canadian, Australian, New Zealand, Cockney, and traditional and affected British English).

Places of articulation

6.2.2. Here we consider the most important PLACES (or *points*) of articulation according to a structural and typological point of view (further on, we will see many more). The most external ones are BILABIAL (/m; p, b/), as in *my pub* /mæˈpʌb/, and LABIODENTAL (/f, v/), as in *five* /ˈfaɪv/. These articulations are particularly easy to see.

fig 6.3. Bilabial and labiodental articulations.



Immediately afterwards, we encounter the places: DENTAL ([t, d; θ, ð; s, z], fig 6.4), as in *the thing*, *seize* /ðəˈθɪŋ, ˈsiːz/, and Spanish *data*, *zona* /ˈdata, ˈθona/ (in American Spanish we have /ˈsona/); ALVEOLAR ([n; ɲ; ɟ; r; l], fig 6.5), as in *today* /təˈdeɪ/, and Spanish or Italian *rana*, *luna*, Sp. /ˈrana, ˈluna/, It. /ˈrana, ˈluna/. In English, /t, d/ are alveolar (as we have already seen), as is the Castilian Spanish /s/. In phonemic (or phonological) transcription, the same symbols are used as above: *today* /təˈdeɪ/, *casas* /ˈkʌsɪs/. However, in truly useful phonetic transcriptions, more precise symbols are to be used, [t̪, ɟ̪; ʃ̪], as will be seen later (\mathcal{C} 8-10 & \mathcal{C} 16-23, and in *HPr*), although they are not official.

fig 6.4. Dental articulations.

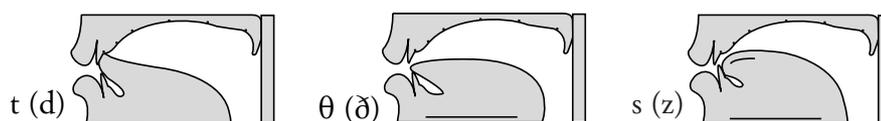
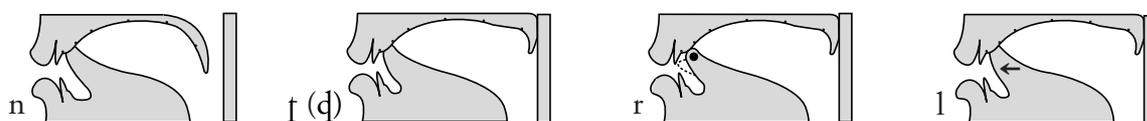


fig 6.5. Alveolar articulations.



We make a quick note that the phonemic transcription for English –as used here– is intentionally a bit different from what is normally found in English textbooks. Our notation is, in fact, less abstract than ⟨/tə'deɪ/⟩ (or ⟨/tə'dei/⟩ which was used, still less concretely, not too long ago). The English pronunciation, for example of *my*, is not [ˈmaɪ], just as in Italian *mai*; instead, the second element is possibly more like the vowel of *yes* /jɛs/. It is true that the symbol [ɪ], which we have not yet introduced, represents a more open sound than [i]. But it is not sufficiently open to indicate adequately the true pronunciation of the English diphthong /æ/ (here given in phonemic transcription, not in a more concrete phonetic one). In the chapter on English of *HPr* (ie Φ 2), it will be possible to find all of the details and nuances necessary for describing (and hence, for learning and teaching) the genuine pronunciation of English.

6.2.3. Returning to the places of articulation, we now have the POSTALVEOLAR location (fig 6.6), which is still farther back than the alveolar one. It occurs in British English *rain* /ˈɹeɪn/ [ˈɹeɪn], although we diaphonemically use /ˈɹeɪn/ (American English [ˈɹeɪn]). It is quite clear that the British articulation is postalveolar (in spite of the misleading official term ⟨retroflex⟩, which means the same thing, although saying it in a more complicated way). However, in part because of a less clear official terminology, even British and American phoneticians often exchange the symbols, using [ɹ] for the neutral American *r*, which is not postalveolar, but a slightly postalveolarized prevelar approximant, that we indicate exactly with the symbol [ɹ̠].

The following place of articulation, which officially (but very dangerously) is called ⟨postalveolar⟩, naturally risks being confused with the preceding articulation (which is legitimately POSTALVEOLAR) – a common fate with those who entrust their fate to overly simplistic definitions. In reality, we have here a compound articulation. It is not merely POSTALVEOLAR, but also has two simultaneous articulatory components (ie *coarticulations*): one which is PALATAL and another which is LABIAL. fig 6.6 (on the right) shows the articulation of the (respectively, voiced and voiceless) consonants *church*, *judge* /ˈtʃɜːtʃ/, ˈdʒʌdʒ/. As can be seen, there is a point of contact (in the postalveolar zone, indicated in black [for reasons we will soon see when we move on to manners of articulation]), and a point of proximity of the articulatory organs (at the palate), as well as (fairly visible) protrusion of the lips.

The descriptions of this articulation are usually among the worst (and this goes for the MANNER as well). In fact, perhaps thinking to make things easier by (excessive) simplification, the articulation is often described as ⟨palatal⟩ (as an alternative to ⟨postalveolar⟩, already seen). In reality, its proper definition is POSTALVEO-PALATAL PROTRUDED, precisely because each of its three components is fundamental. For example, in Spanish, we encounter an articulation without lip protrusion, which is therefore simply POSTALVEO-PALATAL. It is useful to indicate this slightly different articulation with a symbol of its own (as we have already mentioned, and will again). The symbol used is a suitably modified version of the one used for the articulation with lip protrusion, so that the relationship between the articulations

is preserved in the symbols, without, however, confusing them together. In phonemic transcriptions, the more general symbols are employed in all cases, thus, we have Spanish *chachachá* /tʃatʃa'tʃa/ [tʃatʃa'tʃa].

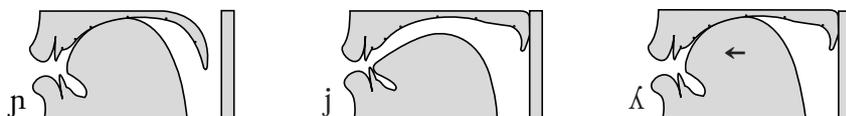
Although it is more complex, this clearer definition surely helps the reader to fully understand the mechanism of its articulation; and the consequential knowledge and phonetic richness leads to much more satisfying practical results. In fact, phonetics should not be carried out unwillingly, proceeding only by memorization. Phonetics is an artistic science, and as such, should be «savored» and «lived» in the best and most creative way (as we have already pointed out in § 0.4).

fig 6.6. Postalveolar and postalveopalatal protruded articulations.



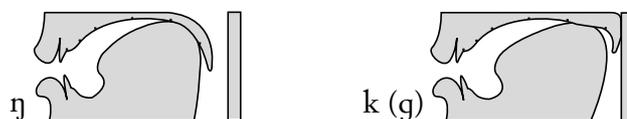
6.2.4. We next come to the true PALATAL place of articulation (fig 6.7), as with Italian /ɲ, j, ʎ/ in Castilian Spanish *sueño, ciencia, calle* /'sweno, 'θjenθja, 'kaʎe/, or in Italian *gnocco, paio, foglia* /'ɲokko, 'paio, 'foʎʎa/. English has /j/ in *yes, unit* /'jes, 'juɪnɪt/.

fig 6.7. Palatal articulations.



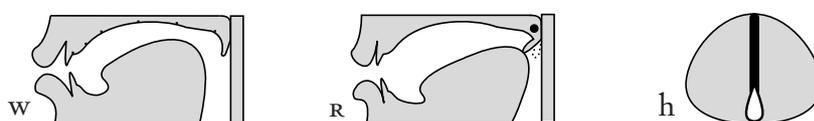
We then have the VELAR place (fig 6.8). The velar nasal, /ŋ/, is a phoneme in English (occurring between vowels as well): *sing, singing* /'sɪŋ, 'sɪŋɪŋ/. Moreover, there are the velar stops, /k, g/, as in *cat, get* /'kæt, 'gɛt/. In Spanish and Italian, [ŋ] only occurs as a contextual variant of the phoneme /n/, as in Sp. *congreso* /kon'greso/ [kon'gɾɛ'so] or It. *congresso* /kon'gresso/ [kon'gɾɛ'sso].

fig 6.8. Velar articulations.



Adding lip rounding (as in /u/), we obtain the VELAR ROUNDED place of articulation (fig 6.9, on the left), as in /w/ in *wit, one* /'wɪt, 'wʌn/, or in Spanish *cuatro* /'kwatro/, or Italian *uomo* /'wɔmo/.

fig 6.9. Velar rounded, uvular, and laryngeal articulations.



Farther back, we find the UVULAR place (fig 6.9, in the middle), which we will exemplify with the voiced trill, /ʀ/. It is advisable to use this symbol in phonemic transcriptions of French and German, even though the most frequent actual realization in these languages is not a trill (as will be seen later, or particularly in the chapters of *HPr* dedicated to these two languages, ie \mathfrak{G} 4-5). The purpose of this choice of phonemic symbol is to make it particularly evident that the articulation is uvular (and not alveolar, /r/, or postalveolar, /ʀ/): French *rare* /'ʀa:ʀ/, and German *rein* /'ʀaen/.

The last place of articulation (in this simplified table) is the LARYNGEAL place, most commonly represented by /h/ (fig 6.9, on the right), as in English *hat* /'hæt/, and German *Hans* /'hans/.

Manners of articulation

6.3.0. Now, in order to fully master the table of fig 6.2 (which can be pictured mentally as well, since it is fairly simple – though new to those who have never done phonetics), we will move on to the seven fundamental MANNERS OF ARTICULATION, using the same consonants, but from this opposing perspective.

The PLACE and the MANNER of articulation are two of the *three* components constituting the consonants – the third, already examined (§ 4.1.7-12 & 6.1.3), is the TYPE OF PHONATION, particularly the distinction VOICED vs VOICELESS.

We will now move through the table, from the top downwards, so that we can see these MANNERS of articulation. The presentation will follow a quite precise physiological and articulatory logic, as we shall see.

6.3.1. *Nasal* (1). Lowering the velum, we open the passage to the nasal cavity, thus allowing expiratory air to escape from the nose. The result is the NASAL manner of articulation, which is combined with a closure produced somewhere in the mouth (in this table, in the bilabial, alveolar, palatal, or velar places). However, these articulations should certainly not be called <stops> (the next manner we will consider), since nasal sounds are continuous, not momentary. Notwithstanding the closure in the oral channel, air can continuously escape through the nose, and the sound can be prolonged as long as expiratory air remains available.

The four nasal consonants we have considered are [m, n, ŋ] in English *man*, *singing* (or [m, n, ɲ, ŋ], as in Spanish *mar*, *no*, *caña*, *tengo*, or in Italian *mai*, *no*, *ragno*, *lungo*), and they are voiced. We group them together in fig 6.10 so that it can be easily seen that the velum is lowered in all four of them.

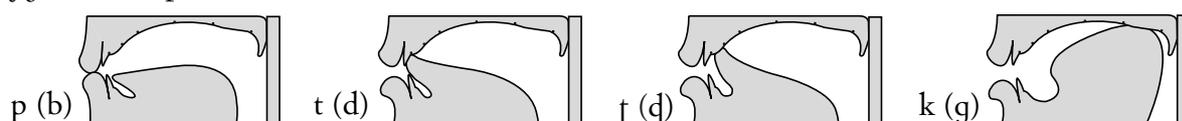
fig 6.10. Nasal articulations.



6.3.2. *Stop* (2). If instead the velum is raised (as in all of the manners which follow), and a closure occurs, we have the STOP manner of articulation (fig 6.11). Here we have voiced and voiceless consonants, as in [p, b; t, d; k, g]: *pen, Ben; two, do; cot, got*; and [t, d] *diente* (Sp.); *dente* (It.).

In all the figures given to illustrate the manners of articulation, the reader should pay particular attention to what they have in common (even between different places of articulation) – these common features are precisely the characteristics of the manner in question.

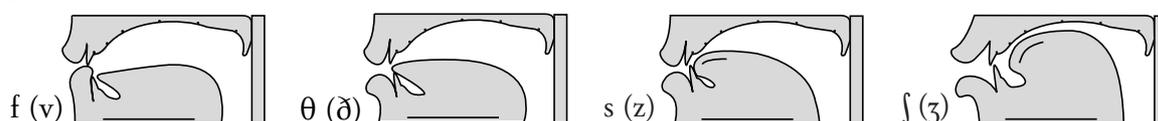
fig 6.11. Stop articulations.



6.3.3. *Constrictive* (3). For now, it will be convenient to skip the manner which is <halfway> between the preceding manner and this one (and indicated in the table as 2+3, since it results from a combination of those two manners in a single sound – the reason will be seen shortly). We therefore come to the CONSTRUCTIVE manner of articulation, characterized by the speaker bringing the articulatory organs sufficiently close together that there is an audible noise of air friction. The constrictive manner is characterized by this friction, which however differs quite a bit in sound, depending upon the place of articulation. In the table of fig 6.2, we have four DIPHONIC PAIRS of constrictives (which appear in fig 6.12), ie [f, v; s, z; θ, ð; ʃ, ʒ], as in *five, seize, this thing, ash, rouge* /'faev, 'sɪz, ðɪsθɪŋ, 'æʃ, 'ruʒ/. As we have indicated, a diphonic pair consists of voiceless and voiced elements, sharing the same place and manner of articulation.

The term CONSTRUCTIVE is clearer and more appropriate, since it is articulatory in nature, and therefore easier to put into concrete relationship with the production of the sounds in question. However, due to a sort of pernicious inertia, the term <fricative> is still more common (the term is auditory and semantically much less transparent).

fig 6.12. Constrictive articulations.



6.3.4. *Stopstricative* (2+3). The combination of manners 2 and 3 produces the STOPSTRICATIVE manner, which naturally derives from *stop + constrictive*. The more common term <affricate> is not articulatory, but rather auditory, and therefore less evident and less easily concretized. Instead, the new term *stopstricative* immediately communicates the exact nature of the sound by virtue of its compound structure: the sound is composed of a first part which is incomplete, firmly joined to a second part, which characterizes it.

In the table, we have one diphonic pair of stopstrictives, [tʃ, dʒ], as in *match*, *age* /'mætʃ, 'eɪdʒ/. The mechanism is a combination of the stop manner (2) and the constrictive manner (3), with a total *length* corresponding to that of a *single* segment, *not* to the sum of two segments. A duration equivalent to that of two segments is found instead in SEQUENCES /ts, dz; tʃ, dʒ/, such as, for example, *cats*, *heads* /'kæts, 'hedz/, or French *patchouli*, *adjectif* /patʃu'li, adʒɛktif/.

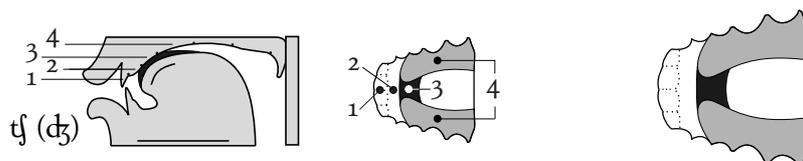
It is important to pay careful attention to the distinction between the stopstrictive symbols, /tʃ dʒ/, which are *monograms*, and the symbols for *sequences*, /tʃ, dʒ/, which are similar, but clearly not identical – cf § 7.2 as well. For instance, in English, we have *patchouli*, /'pætʃəli, pətʃu'li/ and *adjective*, *agent* /'ædʒɪktɪv, 'eɪdʒənt/. The two successive phases of the articulation are, in fact, HOMORGANIC (ie produced in the same place of articulation). What occurs here is the combination of two different manners: the first half is a stop, corresponding in place of articulation to the constriction of the second half.

The best symbols for indicating stopstrictives are MONOGRAMS, as [tʃ, dʒ], which make three fundamental points quite clear: that the sound is a SINGLE sound, and not two sounds in sequence (even though it is composed of two distinct phases), with the NORMAL *duration* of one segment (in fact, for instance, in Italian it is possible to have phonemic oppositions such as the one between *mogio* <downcast> and *moggio* <bushel>: /'mɔdʒo, 'mɔdʒɔ/), and HOMORGANIC, as was mentioned above – it is therefore not a simple combination of [t, d] with [ʃ, ʒ], as can unfortunately be read in certain linguistics texts (and even phonetics texts!).

In fig 6.13, the first phase is marked in black, while the second one is in grey (as with all the other articulations). The first phase is the *stop* phase, and the second is the *constrictive* one, with the articulatory organs close together, but without occlusion of the passage of air. The two diagrams on the right-hand side of fig 6.13 show the mechanism from another point of view: that of PALATOGRAMS.

Comparing the orogram of [tʃ, dʒ] with that of [ʃ, ʒ] (fig 6.12), it is possible to see the difference between the constrictives and the stopstrictives, at least for the case of the postalveopalatal (protruded) place of articulation. Both of these, in our figures, contain a horizontal line at the bottom, which by convention represents the noise common to the two manners. Instead, a curved line, at the height of the blade, represents (also by convention) a longitudinal groove. This groove is formed between the blade of the tongue and the part of the palatal vault it approaches and partially touches. It is through the groove that air escapes, causing the hissing noises which characterize these GROOVED SOUNDS.

fig 6.13. Stopstrictive articulations.



6.3.5. *Approximant* (4). The next manner, following the table of fig 6.2, is the APPROXIMANT manner. It is distinguished from the CONSTRICTIVE manner (3) be-

cause the articulatory organs are less close together, and as a result, they produce a less apparent noise. In fact, this noise is mostly heard only in the voiceless sounds, while in the voiced ones it is usually <covered over> by the voicing produced by vocal-fold vibration.

fig 6.14 gives the orograms of [ɹ, j, w], in which the amount of space between the back of the tongue and the palatal vault is clearly visible. In the orthographic systems of different languages, [j, w] are found written both with <vowel> graphemes and <consonant> graphemes: *use, yes, quite, wet* /'ju:s, 'jɛs, 'kwaɛt, 'wɛt/ in Italian, *ieri, uomo* /'jɛri, 'wɔmo/. Both are voiced.

In the table of fig 6.2 (and fig 6.14, on the right), we have [h], as well. Although it is mostly foreign to the Romance languages, it is nevertheless very important in many other languages: English *hut* /'hʌt/, German *Hut* /'hu:t/. It is voiceless, and produced in the glottis by opening the arytenoids. It therefore has no oral articulation of its own (but cf § 11.3).

fig 6.14. Approximant articulations.



6.3.6. *Trill* (5). The second to last manner in the table is the TRILL manner. It regards sounds which produce a pair of rapid tapping contacts of the tongue tip against the alveolar ridge, in the case of [r] in Italian *rana* /'rana/, or of the uvula against the postdorsum, as in the [ʀ] theoretically possible for French *rue* /'ʀy/ or German *Rast* /'rast/. In Spanish, the alveolar trill is typically longer: *rana* /'r:ana/ (sometimes we find </'rrana/>). Both are voiced, and both are shown in fig 6.15, where the tapping contacts are indicated schematically by the dark balls, and more concretely by the dashed outlines (more easily visible in the magnified versions on the sides).

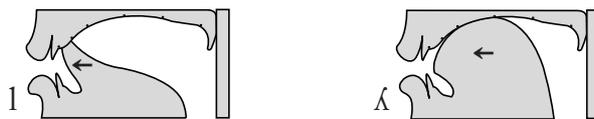
Later on, we will also encounter <trills> with only one tapping contact (these are called TAPS, § 9.22). It will be seen, in any case, that the grapheme *r* does not represent a strong or weak trill at all, in many languages, but rather a constrictive or an approximant.

fig 6.15. Trill articulations.



6.3.7. *Lateral* (6). The last manner is the LATERAL one, in which the tongue, while touching a point on the palatal vault, contracts laterally, thereby permitting air to pass out by the sides of the tongue. fig 6.16 shows the laterals [l, λ], as in *lily* /'lɪli/, or in Castilian Spanish *calle* /'kaʎe/, or Italian *luglio* /'luʎʎo/. English and many other languages do not have any [λ] sound.

fig 6.16. Lateral articulations.



Prosodic elements

6.4.1. While speaking of the vowels (§ 6.1.2), we have already mentioned the distinct role that segment DURATION (also called LENGTH or QUANTITY) can have in certain languages. Normally, the CHRONEME, /:/, is placed after a vowel when it is necessary to indicate length (as we have seen in § 6.1.2, in the case of German *Stadt* /'ʃtat/ <city> and *Staat* /'ʃta:t/ <State>). At times, differences in duration are combined with differences in timbre, as we find, again in German, with *offen* /'ɔ:fən/, *Ofen* /'o:fən/. Duration can also be associated with diphthongization, as in English *bee*, *two* /'bi:, 'tu:/. Too often, these last example are still transcribed </bi:, tu:/>, as if they were actually long monophthongs (and, unfortunately, they are also often transcribed without a stress mark, as if monosyllables could not be either stressed or unstressed).

PHONEMIC LENGTH of consonants is better indicated by doubling, or more technically GEMINATING the symbol. This is especially true of languages such as Italian, where –phonetically as well– the consonants in question are truly GEMINATE, extending over two different syllables ([CC], and not merely <lengthened> consonants, [C:]): *vanno*, *detto*, *faccio*, *passo*, *carro*, *gallo* /'vanno, 'detto, 'fatʃtʃo, 'passo, 'karro, 'gallo/. It is thus important to avoid transcriptions such as </van:o, 'det:o, 'fatʃ:o, 'pas:o, 'kar:o, 'gal:o/> (or, even worse, </fatʃ:o/>). Let us also note English: *penknife*, *bookcase*, *this seat* /'pɛnnaɪf, 'bʊkkɛɪs, ðɪs'si:t/.

PHONETIC LENGTH (which is not distinctive) of single elements, whether vowels or consonants, is marked with the CHRONE, [ː], or with the SEMI-CHRONE, [ˑ] (when less duration is present): English *car*, *card*, *cart*, *cardigan* /'kha:(ɪ), 'kha:(ɪ)ɔ̃, 'kha:(ɪ)t̪, 'kha:(ɪ)ɔ̃dɪgən/ /'ka:ɪ, 'ka:ɪd, 'ka:ɪt, 'ka:ɪdɪgən/ (the diaphoneme /ɪ/ corresponds to [ɪ] in American English, but to <zero> in British English), *sea*, *seed*, *seat*, *seeding* /'sri, 'sriɔ̃, 'si:t̪, 'si:t̪ɪŋ/ /'sri, 'si:d, 'si:t, 'si:dɪŋ/.

Stress

6.4.2. Word STRESS (as well as that of RHYTHM GROUPS, or *stress groups* – the first term is preferable) is marked by /'/ in front of the syllable in question: *finally* /'fæ:nəli/ (and certainly not in front of the stressed vowel, </fæ:nəli/>, nor above the vowel, </fá:nəli/>). Secondary stress, which is weaker (and generally, phonetic and not phonemic, ie without distinctive value), is denoted by /,/: *dynamite* /'daɛnə-maɛt/ (not </d'aɛnəməaɛt/>, nor with the graphemic grave accent placed above the vowel, </dá:nəməàɛt/>).

Especially in Romance studies, terminological inertia has dragged obviously un-

scientific names through time from the Roman era to the present, and so we must insist, once again, that <tonic> is completely inappropriate in the sense of STRESSED. The word *tonic* clearly refers to the *tone* (pitch) of a syllable, not to its *stress*. The Romans took their terminology for syllable PROMINENCE from Greek, where prominence was *tonal* (determined by *pitch*), even though in Latin prominence was *intensive, stress-based*. All terms of this sort without scientific foundation should be rigorously avoided, since they cannot fail to produce dangerous conceptual misunderstandings.

In the case of stress position, it is also good to use scientific and objective terminology. We will therefore speak of FINAL-STRESSED words (stressed on the last syllable, rather than <oxytone>), ie with stress on the last syllable: *ago, again, replace, kangaroo* /ə'gɔʊ, ə'gɛ(ɪ)n, ɪ'plɛɪs, [ɪ]kæŋgə'ɹuʊ/, Spanish *terminó, convoy, tendría, tomar* / [ɪ]termi'no, kom'boi, ten'dria, to'maɾ/, Italian: *partirà, partirai, ferrovia, Manin* / [ɪ]parti'ra, [ɪ]parti'rai, [ɪ]ferro'via, ma'nin/.

Next we have PENULTIMATE-STRESSED words (stressed on the last but one syllable, <paroxytone>): *apparent, deductive, evolution* /ə'pæ.rɪənt, dɪ'dʌktɪv, [ɪ]ɛvə'lʊʃən/ or / [ɪ]iivə-/ , Spanish: *termino, mañana, hermoso* /ter'mino, ma'ɲana, er'moso/, Italian: *ritorno, domani, principi* <principles> (also written *princípi*) /ri'torno, do'mani, prin'tʃipi/; PREPENULTIMATE-STRESSED ones (stressed on the last but two syllable, <proparoxytone>): *dedicate, cumbersome, curiosity* /'dɛdɪ[ɪ]kɛɪt, 'kʌmbə.ɪsəm, [ɪ]kjʊə.ɪ'ɒsɪti/, Spanish: *término, régimen, regímenes* /'termino, 'rɛximen, rɛ'xime'nes/, Italian: *ritornano, domenica, termino, principi* <princes> (also written *príncipi*), *fabbrica* /ri'tornano, do'menika, 'termino, 'printʃipi, 'fabbrika/.

Much less frequently, we encounter words STRESSED ON THE FOURTH TO LAST SYLLABLE: *prosecutor, definitely* /'pɹɒsɪ[ɪ]kjʊtəɹ, 'defɪnɪtli/, Italian: *terminano, fabbricalo* /'termina[ɪ]no, 'fabbrika[ɪ]lo/; on the FIFTH TO LAST: *cumulatively, positivism* /'kjʊm.jələtɪvli, -[ɪ]ləɪtɪvli, 'pɒzətɪ[ɪ]vɪzɪm/, Italian: *fabbricamelo* /'fabbrikame[ɪ]lo/; and on the SIXTH TO LAST as in the very rare Italian form *fabbricamicelo* <build it for me there, or by means of that, or out of that> /'fabbrika[ɪ]mitʃe[ɪ]lo/ (actually, a form made up purposely as an example, just to set a linguistic record).

Sentence stress

6.4.3. It is advisable to consider as SENTENCE STRESS, or *ictus*, every case of word stress which remains stressed in sentence context, and does not become reduced. When stress reduction actually occurs, it is a phonetic (rather than a phonemic) phenomenon, as in Italian *tre gatti* <three cats> /treg'gatti/ [treg'gattɪ], where the isolated /'tre/ loses its stress when placed in a rhythm group. In English such a reduction does not occur; as a matter of fact, we can easily have examples such as: *Then three nice black cats ran out* /ðɛn 'θɹi 'naɛs 'blæk 'kæts 'ɹæn 'aʊt./.

It is preferable to avoid using the term <sentence stress> to refer to the sentence FOCUS; this last notion refers to the word, or words (and therefore concepts), which in a given utterance are communicatively more PROMINENT; in fact, they are highlighted by virtue of being new to the conversation (as opposed to being

already given, or known). Sentence stress and focus are in fact two distinct attributes, although they are not necessarily incompatible. In fact, they can both be present in the last stress group, even though this possibility is statistically the least frequent: *I never said that was true* /æ'nɛvəɪ 'sɛd 'ðæt wəz'tɹuː/ (or in Italian *Non ho mai detto che questo fosse vero* <I never said that was true> /nonɔmmai'detto kekwestofosse'vero./).

In practice, it is much more probable that the sentences above would be said as /æ'nɛvəɪ 'sɛd, 'ðæt wəz'tɹuː./, or better /æ'nɛvəɪ 'sɛd, 'ðæt wəz'tɹuː./, or /æ'nɛvəɪsɛd. ɹ'ðæt wəz'tɹuː./ (and /nonɔmmai'detto, kekwestofosse'vero./, or /nonɔmmai'detto, kekwestofosse'vero./, or also /nonɔm'maidetto. ɹkekwestofosse'vero./).

Therefore, a concrete utterance (which is sufficiently long) will have multiple *ictuses*, *ie* *protonic* syllables and one or more *tonic* syllables (in the rigorous sense of *stressed syllables* in the *intoneme*). At the same time, the utterance can also have one or more points which are communicatively *highlighted* (*ie* the *sentence foci*), and these are generally expressed by different proportions of stress and pitch.

The sentence *These are the new co-workers of my neighbor Roberta* /ðɹiːzəɪðə'njʊu 'kɔʊwəɹ:kəɹz əvmaɛ'nɛɪbəɹ ɹə'bɛɹɹtə/ can be variously realized, with single or multiple highlights. We can therefore encounter /ðɹiːzəɪðə'njʊu 'kɔʊwəɹ:kəɹz, əvmaɛ'nɛɪbəɹ ɹə'bɛɹɹtə/, or also /ðɹiːzəɪðə'njʊu, 'kɔʊwəɹ:kəɹz, əvmaɛ'nɛɪbəɹ ɹə'bɛɹɹtə/, or possibly /ðɹiːzəɪðə'njʊu 'kɔʊwəɹ:kəɹz, əvmaɛ'nɛɪbəɹ, ɹə'bɛɹɹtə/, or else also /ðɹiːzəɪðə'njʊu, 'kɔʊwəɹ:kəɹz, əvmaɛ'nɛɪbəɹ, ɹə'bɛɹɹtə/. Of course, similar subdivisions are possible for the corresponding Italian sentence: *Questi sono i nuovi colleghi della mia vicina Roberta* <These are the new co-workers of my neighbor Roberta>, /kwestisonoi'nwɔvi koll'ɛgi dellamiavi'tʃina ro'berta./, or /kwestisonoi'nwɔvi koll'ɛgi, dellamiavi'tʃina ro'berta./, or also /kwestisonoi'nwɔvi, koll'ɛgi, dellamiavi'tʃina ro'berta./, or possibly /kwestisonoi'nwɔvi koll'ɛgi, dellamiavi'tʃina, ro'berta./, or else also /kwestisonoi'nwɔvi, koll'ɛgi, dellamiavi'tʃina, ro'berta./.

In any case, the elements highlighted can also be grammemes, in cases such as particular contrasts. With the examples above, we can have /'ðɹiːz, əɹðə'njʊu 'kɔʊwəɹ:kəɹz əvmaɛ'nɛɪbəɹ ɹə'bɛɹɹtə/, or /ðɹiːz'ɑɹɹ, ðə'njʊu 'kɔʊwəɹ:kəɹz əvmaɛ'nɛɪbəɹ ɹə'bɛɹɹtə/ (with *are* highlighted), or even /ðɹiːzəɪðə'njʊu'kɔʊwəɹ:kəɹz əv'maɛ, 'nɛɪbəɹ ɹə'bɛɹɹtə/ (with *new* destressed, but with *my* highlighted, for some particular reason). In Italian: /'kwesti, sonoi'nwɔvi koll'ɛgi dellamiavi'tʃina ro'berta./, or /kwesti'sono, i'nwɔvi koll'ɛgi dellamiavi'tʃina ro'berta./ (with *sono* <are> highlighted), or /kwestisonoi'nwɔvi koll'ɛgi della'mia, vi'tʃina ro'berta./ (with *mia* <my> highlighted).

Some kind of attenuation can occur in parts of the sentence rendered <parenthetical>, as in /ɹəvmaɛ'nɛɪbəɹ ɹə'bɛɹɹtə.ɹ/, where *of my neighbor Roberta* is spoken as a sort of afterthought. Again, similar possibilities occur in the Italian example given: /ɹdellamiavi'tʃina ro'berta.ɹ/ *della mia vicina Roberta* <of my neighbor Roberta>.

Tones

6.4.4. Certain languages have distinctive TONES; these are called, logically enough, TONEMES. Distinctive tones imply that when the pitch of a syllable chang-

es, its meaning can change as well. Let us look at, for example, the three basic ton(em)es of the African language Yoruba (cf fig 6.17): *ró*, *ro*, *rò* /ro, 'ro, _ro/ <to drape, to till, to think>.

In fig 6.18, the four ton(em)es of Mandarin Chinese are shown: *mā*, *má*, *mǎ*, *mà* /ma, 'ma, ,ma, `ma/ <mother, hemp, horse, to curse>. Examining these fairly simple examples, it becomes clear that the graphic signs used are capable of referring to (quite) different tonetic realities in different languages. In *HPr*, in § 11 on Chinese, it will be seen that the presentation here is actually rather simplified (cf fig 11.10-20 of *HPr*).

Although our notation system seeks the greatest level of concreteness possible, our tonemic symbols are nonetheless rather <theoretical>. They are, however, much less so than those of other systems, including the system used in the recent reform of the International Phonetic Alphabet, which is expounded in the volume mentioned below. Our tonetic symbols, like our phonetic ones (as will be seen later, in the more detailed sections), are less mysterious, given that they have a real connection to the tonal values they represent. The symbols used before the reform were better –tonemically– even though they were still insufficient from a tonetic point of view.

fig 6.17. The three Yoruba tonemes.

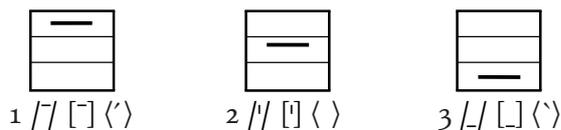
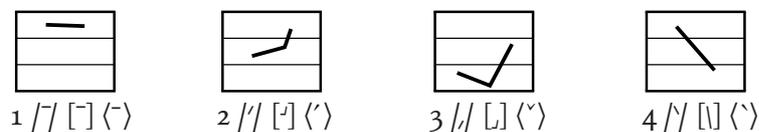


fig 6.18.
The four (Mandarin)
Chinese tonemes.



Intonation

6.4.5.1. We will now concisely introduce the bare essentials of INTONATION. In fact, all languages have their own intonation system, and phonetics should therefore not be treated without examining intonation as well. Unfortunately, it is often left out entirely, even in descriptions of particular languages or in transcriptions of sentences or passages! A notably bad example of this omission is given by the recent <official manual> of the International Phonetic Association: *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet* (found in the bibliography).

In every language the THREE MARKED INTONEMES (/, ? ;/) and the UNMARKED PREINTONEME (the normal //, without a special symbol) should be clearly indicated with appropriate symbols (both on a phonetic, or rather, TONETIC level, and on a phonemic, or TONEMIC one). The *intoneme* involves the final stressed syllable of an utterance and the syllables around it (cf fig 6.19), while the *preintoneme* is what is found before the intoneme in the same intonation group (cf fig 6.20). In the example *his cousin's name is Bartholomew* /hɪz'kʌzɪnz 'neɪm ɪz bɑ:ɾθɒləmjʊ/,

the intoneme is constituted by the full name of *Bartholomew*, while the preintoneme is everything prior to it: *his cousin's name is...*

The example of *Bartholomew* is particularly interesting because it allows us to consider the four ideal components of an intoneme: the *pretonic* syllable (*Bar-*), the *tonic* syllable (*-thol-*), and the two *posttonic* ones (*-omew*). The pronunciation of this example normally provides a reasonably adequate realization of the schematic tonal movements shown in fig 6.19-20 (which besides the unmarked preintoneme and the three marked intonemes, give the important interrogative preintoneme, / ξ /, which is marked, and the continuative intoneme, / \downarrow – which is unmarked). If the example were *his cousin's name is Dick* /hɪz'kʌzɪnz 'neɪm ɪz'dɪk/, the intoneme would be *is Dick*. The tonic and posttonic syllables would consist of only one syllable (*Dick*). In consequence, the ideal movement shown in the diagrams (for the case with four syllables) would be compressed, not just horizontally, but inevitably in terms of the vertical range as well. When only one syllable is present (as in the answer to a question like *what is his cousin's name?* – *Dick*), the result is a fusion of the expected pitch patterns which maintains the characteristic movements, but in an attenuated form.

The intonation schemes of the British school are among the few to have some practical use; but precisely for the reasons considered here (and in general), they are sometimes decidedly excessive. In fact, for [\cdot ' \cdot] or [\cdot ' \cdot '] (cf fig 6.19), they give diagrams like $\bar{\downarrow}$ or $\bar{\downarrow}$ when there is only one short voiced element: for example for [ɪ] in *Dick* – if the result were truly as extended as their diagrams show, it would rather sound like a police siren!

fig 6.19. The four intonemes of neutral British English.

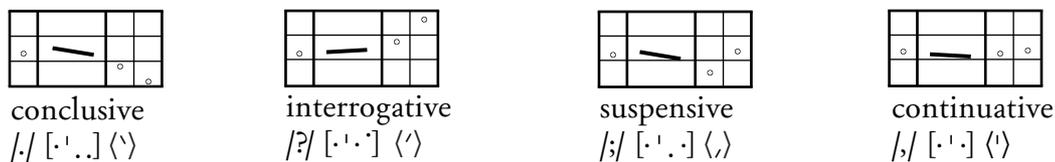
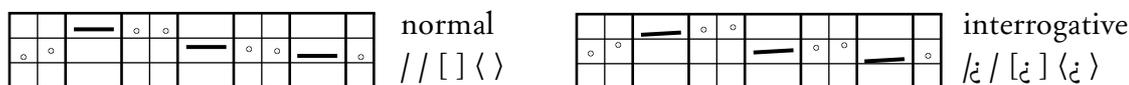


fig 6.20. Two preintonemes of neutral British English.



6.4.5.2. The preintoneme and the intoneme taken together are usefully called an INTONATION GROUP. We use examples such as *My favorite dictionary*, or *That patient thinks he's Giuseppe Verdi*, to show that the parts of an intonation group do not necessarily respect word boundaries. In fact, the intonemes in these utterances are, respectively: /ɹət'dɪkʃən.ɪ./ and /i'veɹə.ɪdi./ (*-rite dictionary* and *-pe Verdi*). The preintonemes, on the other hand, are /ðæts mæ'feɪv/ and /ðæt'peɪʃənt 'θɪŋks (h)ɪz'dʒʊ'seɪp/ (*My favo-* and *That patient thinks he's Giusep-*). The full examples are: /ðæts mæ'feɪv.ɹət 'dɪkʃən.ɪ./ and /ðæt'peɪʃənt 'θɪŋks (h)ɪz'dʒʊ'seɪpi 'veɹə.ɪdi./

It will be seen that our transcriptions are not subdivided pedantically along word boundaries. That practice is still quite common (in the best case, motivated

by hopes of helping the reader). It is much more useful to subdivide transcriptions into rhythm groups, as we have done, instead of giving things like </'ðæt ɪz 'maɪ 'feɪnrət 'dɪksənri/>, or </'ðæt 'peɪsnt 'ðɪŋks hɪz dʒu:'sepi 'veədi/>, where the stresses and some un-reduced forms (for current <weak forms>) are also unnatural (ie in the cases of </ɪz 'maɪ/ in the first example and </hɪz/>, at least, in the second, which are weakened in normal speech, both articulatorily and prosodically).

Another (not unimportant!) counsel regards the fact that <sounds have no capitals>; note that, for other reasons, the traditional orthographies of languages such as Arabic and Hindi, and Chinese and Japanese as well, have no capital letters. Children can easily tell that there is no phonic difference between *smith* and *Smith*, or between Italian *franco* and *Franco* – both of the English examples are pronounced exclusively /'smɪθ/, and the Italian examples are both pronounced /'franko/. And yet, even in textbooks, all too often we find (printed, as well) atrocities such as </Dʒu:'sepi 'Veə(r)di/ and also </'Maɪ/ absurdly derived from writing conventions! (Here *My*'s transcription is given with a capital letter, because it is the first word in the sentence. Moreover, the transcription of *Giuseppe* uses a capital letter because the word is a proper name, and the result is an inappropriate and ambiguous digram, *Dʒ*, instead of a slightly less forced *Dʒ*, which would at least represent the unity of the sound [dʒ] better.)

6.4.5.3. fig 6.21 will be a useful explanatory tool in order to understand more explicitly the use of tonograms (given that we are not all musicians or singers, for whom the analogy with a musical score is obvious). Let us observe, then, the graphemic text, to which we have given the form of the intonation curve. Normally this curve is shown with the lines and dots of tonograms, but here we have used a more <intuitive> approach. We show just four examples, based on the segment *see you on Saturday* (in neutral British pronunciation expressly to compare them with $\bar{\square}$ and $\bar{\square}$, seen above). These examples contrast pairwise: a *conclusive* utterance is contrasted with an *interrogative* one (of a total question), and a *suspensive* utterance with a *continuative* one.

fig 6.21. An iconic way to introduce people to 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>

In the case of the last two sentences, the semantic importance of what follows (given in parentheses) is fundamental, whether it is expressed out loud, or instead

remains implicit. In any case, the suspensive intoneme is characterized by decidedly greater and more immediate anticipation, while this is lacking with the continuative. This difference, and certainly not their syntax, explains the difference in intonation between the third and fourth examples.

Applying the movements of the three intonemes to a slightly different example, we see that in neutral (better than ‹standard›) British English, the *conclusive intoneme* is falling (./ [·' .]), of the type shown in fig 6.19 (and also in three examples in fig 6.21): *Christian* /'kɹɪstʃən./ ['kɹɪstʃən.]. The *interrogative intoneme* is rising (/ʔ/ [·' ·']), as in the question *Christian?* /'kɹɪstʃənʔ/ ['kɹɪstʃən·']. The third intoneme, the *suspensive*, is used to create a sort of anticipation, or ‹suspense›. In neutral British pronunciation, it is falling-rising, /;/ [·' ·.]: *Although his name's Christian*, -/'kɹɪstʃən;/ ['kɹɪstʃən:]—*he's no good Christian at all*.

6.4.5.4. In fig 6.20 (as well as in the second example of fig 6.21), we have the *interrogative preintoneme*, /ǰ /, as well. This preintoneme is a modification of the normal preintoneme, and it anticipates on the rhythmic-group syllables of the preintoneme the characteristic movement of the interrogative intoneme (although in an attenuated form). Obviously, in the part specifically dedicated to the topic, we will be more explicit and more exhaustive (cf § 13.8-34). Here, we remark only that the interrogative preintoneme is the same in all types of questions, whether these are TOTAL questions, like *Is his cousin's name Christian?*, or PARTIAL ones (containing a question word, such as *why, when, who, how...*), such as *Why is his cousin's name Christian?*

We must warn the reader that, contrary to what grammar books and writing-based teaching imply, not all questions have an interrogative intoneme, nor should they. In fact, partial questions, in order to sound truly natural and authentic, should be pronounced with a conclusive intoneme (or at most, with the unmarked *continuative* intoneme, with pitch in the mid band, which will be seen in greater detail later on): *Why is his name Christian?* /ǰ'waɪz hɪz'neɪm 'kɹɪstʃən./ [ǰ'waɪz hɪz'neɪm 'kɹɪstʃən.] (or /'kɹɪstʃən./ ['kɹɪstʃən·], with a continuative intoneme).