Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet

PART 1

Introduction to the IPA

1. What is the International Phonetic Alphabet?

The aim of the International Phonetic Association is to promote the scientific study of phonetics and the various practical applications of that science. For both these it is necessary to have a consistent way of representing the sounds of language in written form. From its foundation in 1886 the Association has been concerned to develop a system of notation which would be convenient to use, but comprehensive enough to cope with the wide variety of sounds found in the languages of the world; and to encourage the use of this notation as widely as possible among those concerned with language. The system is generally known as the International Phonetic Alphabet. Both the Association and its Alphabet are widely referred to by the abbreviation IPA, but here 'IPA' will be used only for the Alphabet. The IPA is based on the Roman alphabet, which has the advantage of being widely familiar, but also includes letters and additional symbols from a variety of other sources. These additions are necessary because the variety of sounds in languages is much greater than the number of letters in the Roman alphabet. The use of sequences of phonetic symbols to represent speech is known as transcription.

The IPA can be used for many different purposes. For instance, it can be used as a way to show pronunciation in a dictionary, to record a language in linguistic fieldwork, to form the basis of a writing system for a language, or to annotate acoustic and other displays in the analysis of speech. For all these tasks it is necessary to have a way of designating sounds unambiguously, and which is generally understood. The purpose of this Handbook is to provide a practical guide to the IPA and to the conventions associated with it.

Phonetics, like any science, develops over time. New facts emerge, new theories are created, and new solutions to old problems are invented. The notational system of any science reflects facts and theories, and so it is natural that from time to time the Alphabet should be modified to accommodate innovations. The Alphabet presented in this Handbook is the version revised by a Convention of the International Phonetic Association held in Kiel in 1989, subject to a subsequent set of minor modifications approved by the Council of the Association in 1993. Despite these and earlier changes, the Alphabet today shows striking continuity with the Association's Alphabet as it was at the end of the 19th century. The development of the IPA has, throughout the history of the Association, been guided by a set of 'Principles', and these are listed in Appendix 1.

2. Phonetic description and the IPA Chart

Behind the system of notation known as the IPA lie a number of theoretical assumptions about speech and how it can best be analyzed. These include the following:
— Some aspects of speech are linguistically relevant, whilst others (such as personal voice quality) are not.
— Speech can be represented partly as a sequence of discrete sounds or 'segments'.
— Segments can usefully be divided into two major categories, consonants and vowels.
— The phonetic description of consonants and vowels can be made with reference to how they are produced and to their auditory characteristics.
— In addition to the segments, a number of 'suprasegmental' aspects of speech, such as stress and tone, need to be represented independently of the segments.

The IPA is summarized in the 'IPA Chart', reproduced as the unnumbered center pages of this issue, and the structure of this chart reflects the above assumptions. The following subsections provide a very brief introduction to phonetic description in the context of these underlying assumptions, while referring to the relevant parts of the Chart. This introduction can only deal with a few important points, and readers who need a more thorough treatment of phonetic description should consult phonetics textbooks.

2.1 Linguistically relevant information in speech

Although phonetics as a science is interested in all aspects of speech, the focus of phonetic notation is on the linguistically relevant aspects. For instance, the IPA provides symbols to transcribe the distinct phonetic events corresponding to the English spelling *refuse* (['rɛfjus] meaning 'rubbish' and [rɪ'fjuz] meaning 'to decline'), but the IPA does not provide symbols to indicate information such as 'spoken rapidly by a deep, hoarse, male voice'. Whilst in practice the distinction between what is linguistically relevant and what is not may not always be clear-cut, the principle of representing only what is linguistically relevant has guided the provision of symbols in the IPA. The need to go further, however, is now recognized by the 'Extensions to the IPA' presented in Appendix 2.

2.2 Segments

Observation of the movements of the speech organs reveals that they are in almost continuous motion. Similarly the acoustic speech signal does not switch between successive steady states, but at many points changes gradually and at others consists of transient events. Neither the movements of the speech organs nor the acoustic signal offers a clear division of speech into successive phonetic units. This may be surprising to those whose view of speech is influenced mainly by alphabetic writing, but it emerges clearly from (for instance) x-ray films and acoustic displays.

For example, the movements and the acoustic signal corresponding to the English word *worry* will show continuous change. Figure 1 presents a spectrogram of this word. Spectrograms are a way of making visible the patterns of energy in the acoustic signal. Time runs from left to right, and the dark bands reflect the changing resonances of the vocal tract as the word is pronounced. In the case of the word *worry*, it is clear that the pattern ebbs and flows constantly, and there are no boundaries between successive sounds. Nonetheless the word can be segmented as [wɛɹi] – that is, as [w] + [ɛ] + [ɹ] + [i].

This segmentation is undoubtedly influenced by knowledge of where linguistically significant changes in sound can be made. A speaker could progress through the word making changes: in a British pronunciation, for instance, [wɛɹi] *worry,* [hɛɹi] *hurry,* [hæɹi] *Harry,* [hæti] *Hatty,* [hætə] *hatter.* There are thus four points at which the phonetic event can be changed significantly, and this is reflected in the analysis into four segments. Languages may vary in the points at which they allow changes to be made, and so segmentation may have to be tentative in a first transcription of an unknown language (see section 9). Nonetheless there is a great deal in common between languages in the way they organize sound, and so many initial guesses about the segmentation of an unfamiliar language are likely to be right.

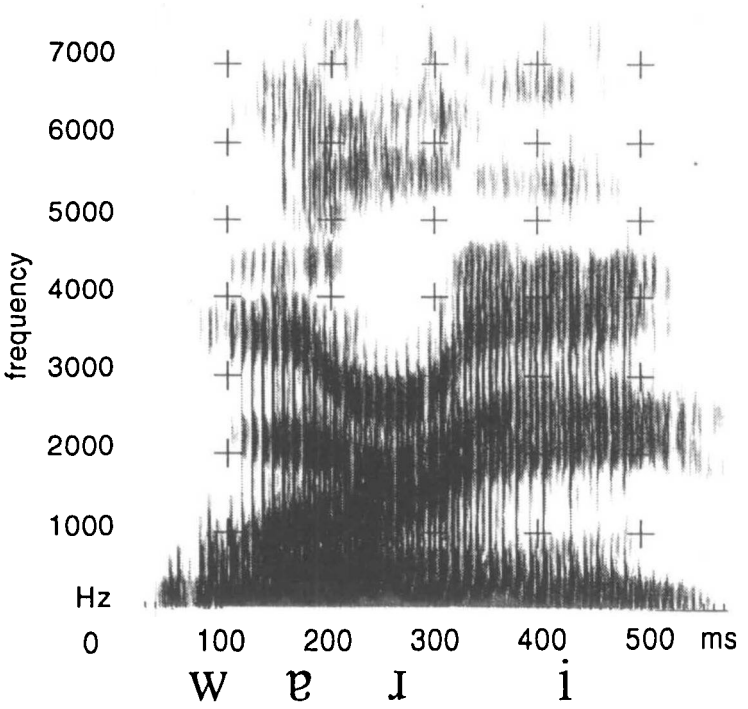


Figure 1. Spectrogram of the word *worry*, spoken in a Southern British accent.

Phonetic analysis is based on the crucial premise that it is possible to describe speech in terms of a sequence of segments, and on the further crucial assumption that each segment can be characterized by an articulatory target. 'Articulation' is the technical term for the activity of the vocal organs in making a speech sound. The description of the target is static, but this does not imply that the articulation itself is necessarily held static. So, for example, [ɹ] (as in the word *worry* above) is described as having a narrowing made by the tongue-tip near the back of the alveolar ridge (the flattish area behind the upper teeth). The tongue-tip actually makes a continuous movement to and from that target, as reflected in the dipping pattern of higher resonances on the spectrogram in Figure 1 at around time 200-300 ms. In other sounds, a target will be held for a fixed amount of time. The important point is that the use of segments and associated 'target' descriptions allows for a very economical analysis of the complex and continuously varying events of speech.

2.3 The consonant-vowel distinction

Broadly, speech involves successive narrowing and opening of the vocal tract, the passage through which the air flows during speech. This can be seen clearly in an example such as *banana* ([bəˈnænə] or [bəˈnɑnə]), in which the vocal tract is closed three times (first by the lips and then twice by the tongue), each closure being followed by an opening of the vocal tract. The successive openings are the basis of syllables, and the word *banana* consists therefore of three syllables. The open part of the cycle is regarded as the center, or nucleus, of the syllable.

Sounds like [b] and [n] which involve a closed, or nearly closed, vocal tract, are 'consonants'. Sounds like [ə] and [ɑ] which involve an open vocal tract are vowels. More precisely, any sounds in which the flow of air out of the mouth is impeded at least enough to cause a disturbance of the airflow are consonants. So a sound such as [s], in which the 'hissing' that can be heard results from the airflow being made turbulent, is as much a consonant as [b]. Conversely any sounds in which the air flows out of the mouth unimpeded are vowels. The distinction between consonant and vowel is fundamental to the way segments are described in the framework underpinning the IPA.

It follows from the definitions of 'consonant' and 'vowel', and from the origin of the syllable in the repeated opening and narrowing of the vocal tract, that vowels are well suited to playing the role of syllable nuclei, and consonants are well suited to defining the margins of syllables. The relationship between syllables and type of sound is not, however, totally straightforward. For one thing, a sound which is a consonant may nonetheless act as a syllable center. So in a common pronunciation of the English word *button* as [bʌtn̩] there are two syllables, but the nucleus of the second is a consonant. Conversely in the word [jɛt] *yet*, the first sound, if prolonged, is very similar to the vowel of [hid] *heed,* and does not involve a narrowing severe enough to produce friction. Because [j] plays the same role in the syllable as sounds which are by definition consonants (e.g. [b] in [bɛt] *bet*), it is useful to include it in the class of consonants and to describe it accordingly.

On the IPA Chart, there are separate sections for vowels and for consonants, reflecting different techniques for describing them. The different techniques arise from the more closed articulation of consonants and the more open articulation of vowels.

2.4 Consonants

Because consonants involve a narrowing or 'stricture' at an identifiable place in the vocal tract, phoneticians have traditionally classified a consonant in terms of its 'place of articulation'. The [t] of *ten*, for instance, requires an airtight seal between the upper rim of the tongue and the upper gum or teeth. Phonetic description of place of articulation, however, concentrates on a section or 'slice' through the mid-line of the mouth, the so-called mid-sagittal plane, and in this plane the seal is made between the tip or blade of the tongue and the bony ridge behind the upper front teeth, the alveolar ridge. The sound is therefore described as alveolar. Figure 2 shows a mid-sagittal section of the vocal tract, with the different places of articulation labeled. As further examples, the [p] of *pen* is bilabial (the closure is made by the upper and lower lips), and the [k] of *Ken* velar (made by the back of the tongue against the soft palate). Other places of articulation are exemplified in section 3.

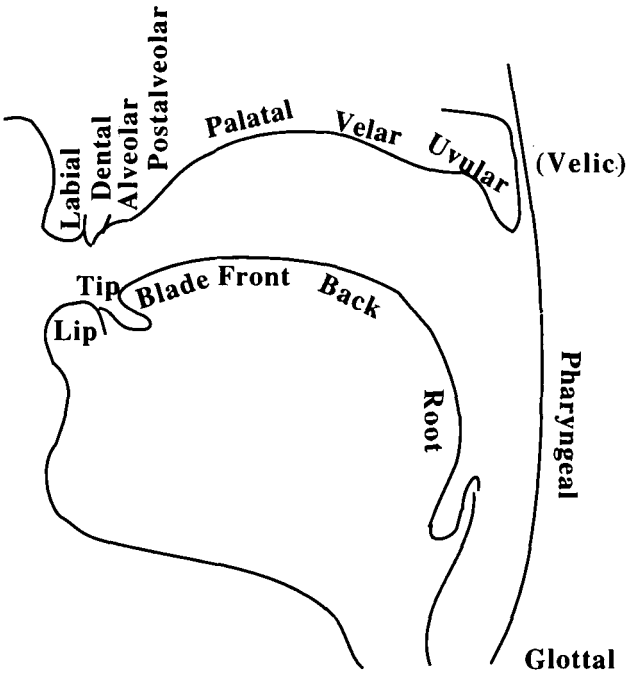


Figure 2. Mid-sagittal section of the vocal tract showing labels for place of articulation

On the IPA Chart, symbols for the majority of consonants are to be found in the large table at the top. Place of articulation is reflected in the organization of this consonant table. Each column represents a place of articulation, reflected in the labels across the top of the table from bilabial at the left to glottal (consonants made by the vocal cords) at the right. The terms 'bilabial' and 'labiodental' indicate that the consonant is made by the lower lip against the upper lip or against the upper teeth; otherwise it is normally assumed that the sound at a named place of articulation is made by the articulator lying opposite the place of articulation (so alveolars are made with the tip of the tongue or the blade (which lies just behind the tip)). The exception to this is the term 'retroflex'. In retroflex sounds, the tip of the tongue is curled back from its normal position to a point behind the alveolar ridge. Note that except in the case of fricatives only one symbol is provided for dental / alveolar / postalveolar; if necessary, these three places can be distinguished by the use of extra marks or 'diacritics' to form composite symbols, as discussed in section 2.8. For example, the dental / alveolar / postalveolar nasals can be represented as [n̪ n n̠] respectively.

The rows of the consonant table, labeled at the left side by terms such as plosive, nasal, trill, and so on, reflect another major descriptive dimension for consonants, namely 'manner of articulation'. Manner of articulation covers a number of distinct factors to do with the articulation of a sound. One is the degree of stricture (narrowing) of the vocal tract involved. If the articulation of the plosive [t] is modified so that the tongue tip or blade forms a narrow groove running from front to back along the alveolar ridge, instead

of an airtight closure, air can escape. The airflow is turbulent, and this creates sound of a hissing kind known in phonetics as frication. Such a sound is called a fricative. In this case the resultant sound would be [s] as in *sin*. Other fricatives include [f] (as in *fin*) and [ʃ] (as in *shin*). If even less narrowing is made in the vocal tract, an approximant will result, in which the airflow is not turbulent and no frication is audible. Approximants are exemplified by the sound [j] at the start of *yet*, and the first sound in *red* in most varieties of English ([ɹ], [ɻ], or [ʋ] according to the variety).

'Manner of articulation' also includes important factors such as whether the velum (the soft part of the palate at the back of the mouth) is raised or lowered. If it is lowered, as for the sounds [m] and [n] in *man*, the resonances of the nasal cavity will contribute to the sounds. Consonants where this happens are called nasals. Laterals (lateral approximants such as English [l] in *let* and lateral fricatives such as Welsh [ɬ] in *llan* 'church') are sounds where air escapes not in the mid-line of the vocal tract but at the side. Trills are sounds like [r] in Spanish *perro* 'dog' in which the air is repeatedly interrupted by an articulator (in this case the tongue tip) vibrating in an airstream. A very short contact, similar in duration to one cycle of the vibration of a trill, is called a tap such as the [ɾ] in Spanish *pero* 'but'.

A further important factor in the description of consonants is not shown in the column or row labels. This is whether the consonant is voiced or voiceless. In voiced consonants the vocal cords are producing acoustic energy by vibrating as air passes between them, and in voiceless ones they are not. Where a cell in the table contains two symbols, the one on the left is for a voiceless consonant, e.g. [p], and the one on the right is for the voiced counterpart, [b]. Voicing distinctions are actually more fine-grained than implied by this two-way distinction, so it may be necessary to add to the notation allowed by the two basic symbols. For instance, the symbolization [ba pa p^ha] implies consonants, respectively, in which the vocal cords are vibrating during the plosive closure, vibrating only from the release of the closure, and vibrating only from a time well after the release (giving what is often known as an 'aspirated' plosive). Where a cell contains only one symbol, it indicates (with one exception) a voiced consonant and is placed on the right. The exception is the glottal plosive [ʔ] (as the vocal cords are closed, they are unable simultaneously to vibrate).

It should be clear that the consonant table is more than a list of symbols; it embodies a classificatory system for consonants. It allows the user to ask a question such as 'how should I symbolize a voiced sound involving complete closure at the uvula?' (The answer is [ɢ].) Or conversely, what sort of a sound is [ʝ]? (The answer is one which is voiced, and in which frication can be heard resulting from a narrowing between the tongue body and the hard palate.)

Not all cells of the consonant table contain symbols. The gaps are of three kinds. Where the intersection of a manner and a place of articulation define a sound which is thought to be impossible to produce, such as a velar trill, the cell is shaded. Unless phoneticians have made a mistake in judging such a sound to be impossible, no symbols will be needed for any of the shaded cells. An unshaded gap, such as the velar lateral fricative, may indicate that the sound in question can be produced, but has not been found in languages. It is always possible that a language will be discovered which requires the gap to be filled in. A case of this kind is the velar lateral approximant [ʟ], which only became generally known among phoneticians in the 1970s when it was reported in

Kanite, a language of Papua New Guinea. An unshaded gap may also occur where a sound can be represented by using an existing symbol but giving it a slightly different value, with or without an added mark separate from the letter. A symbol such as [β], shown on the chart in the position for a voiced bilabial fricative, can also be used to represent a voiced bilabial approximant if needed. In a similar way, no symbols are provided for voiceless nasals. A voiceless alveolar nasal can be written by adding the voiceless mark [̥] below the letter [n] to form an appropriate composite symbol [n̥]. Many of the gaps on the chart could be filled in this way by the use of diacritics (sections 2.8 and 3). The formation of this kind of composite symbol is discussed further in the section on diacritics below.

2.5 Non-pulmonic consonants

All the symbols in the main consonant table imply consonants produced using air from the lungs ('pulmonic' consonants). Whilst some languages rely exclusively on air from the lungs for sound production, many languages additionally use one or both of two other 'airstream mechanisms' to produce some of their consonants. Symbols for these sounds are given in a separate box below and to the left of the main consonant table. These sounds are exemplified in Section 3.

The more common of the two non-pulmonic airstream mechanisms used in languages, the 'glottalic', involves closing the glottis, and squeezing or expanding the air trapped between the glottis and a consonant stricture further forward in the vocal tract. If the air is squeezed, and therefore flows outwards — abruptly when a closure further forward is released, or briefly but continuously through a fricative stricture — the sound is known as an 'ejective'. Ejectives are symbolized by the appropriate voiceless consonant symbol with the addition of an apostrophe, e.g. [p'], [s']. If instead the air between the glottis and a closure further forward is expanded, reducing its pressure, air will flow in to the mouth abruptly at the release of the forward closure. Usually the closure phase of such sounds is accompanied by vocal cord vibration, giving '(voiced) implosives' such as [ɓ]. If it is necessary to symbolize a voiceless version of such a sound, this can be done by adding a diacritic: [ɓ̥].

'Velaric' airstream sounds, usually known as 'clicks', again involve creating an enclosed cavity in which the pressure of the air can be changed, but this time the back closure is made not with the glottis but with the back of the tongue against the soft palate, such that air is sucked into the mouth when the closure further forward is released. The 'tut-tut' or 'tsk-tsk' sound, used by many English speakers as an indication of disapproval, is produced in this way, but only in isolation and not as part of ordinary words. Some other languages use clicks as consonants. A separate set of symbols such as [ǂ] is provided for clicks. Since any click involves a velar or uvular closure, it is possible to symbolize factors such as voicelessness, voicing, or nasality of the click by combining the click symbol with the appropriate velar or uvular symbol: [k͡ǂ g͡ǂ ŋ͡ǂ], [q͡!].

2.6 Vowels

Vowels are sounds which occur at the center of syllables, and which, because they involve a less severe narrowing of the vocal tract than consonants, cannot easily be described in terms of a 'place of articulation' as consonants can. Instead, they are classified in terms of an abstract 'vowel space', which is represented by the four-sided figure known as the 'Vowel Quadrilateral' (see the Chart, next to bottom left). This space bears a relation, though not an exact one, to the position of the tongue in vowel production, as explained below.

Figure 3 shows a mid-sagittal section of the vocal tract with four superimposed outlines of the tongue's shape. For the vowel labeled [i], which is rather like the vowel of *heed* or French *si* 'if', the body of the tongue is displaced forwards and upwards in the mouth, towards the hard palate. The diagram shows a more extreme version of this vowel than normally found in English at least, made so that any further narrowing in the palatal region would cause the airflow to become turbulent, resulting in a fricative. This extreme vowel is taken as a fixed reference point for vowel description. Since the tongue is near the roof of the mouth this vowel is described as 'close', and since the highest point of the tongue is at the front of the area where vowel articulations are possible, it is described as 'front'.

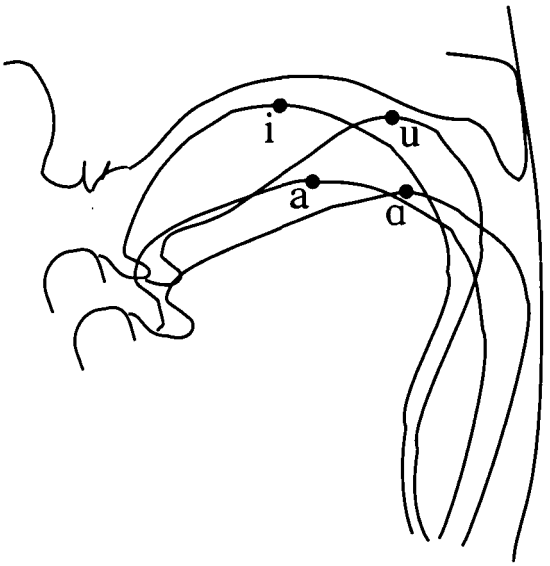


Figure 3. Mid-sagittal section of the vocal tract with the outline of the tongue shape for each of four extreme vowels superimposed.

Conversely, for the vowel labeled [ɑ], which is rather like the vowel of Standard Southern British or General American English *palm,* the tongue body is displaced downwards and backwards, narrowing the pharynx. The most extreme version of this vowel, made so that any further narrowing in the pharynx would result in a fricative, is taken as a second fixed reference point. The space between the tongue and the roof of the mouth is as large as possible, so this vowel is described as 'open', and the tongue is near the back of the mouth, so it is described as 'back'.

If the tongue body is raised as close as possible at the back of the mouth, just short of producing a velar consonant, and (as is common in languages) the lips are simultaneously rounded and protruded, the close back vowel [u] results (see Figure 3), which is similar to the vowel of French *vous* 'you' or German *du* 'you'. And if a vowel is produced in which the highest point of the tongue is at the front of the mouth and the mouth is as open as possible, the result is [a]. This is rather like the quality of the vowel in *cat* in contemporary Standard Southern British English (other dialects may have less open or less front qualities). These two extreme vowels may also be regarded as fixed references.

The first part of Figure 4 shows that joining the dots representing the highest point of the tongue in these four extreme vowels gives the boundary of the space within which vowels can be produced. For the purposes of vowel description this space can be stylized as the quadrilateral shown in the second part of Figure 4. Further reference vowels can now be defined as shown in the third part of Figure 4. Specifically, two fully front vowels [e] and [ɛ] are defined between [i] and [a] so that the difference between each pair of adjacent vowels sounds the same; and similarly, two fully back vowels [ɔ] and [o] are defined to give equidistant steps between [ɑ] and [u]. The use of auditory spacing in the definition of these vowels means vowel description is not based purely on articulation, and is one reason why the vowel quadrilateral must be regarded as an abstraction and not a direct mapping of tongue position.

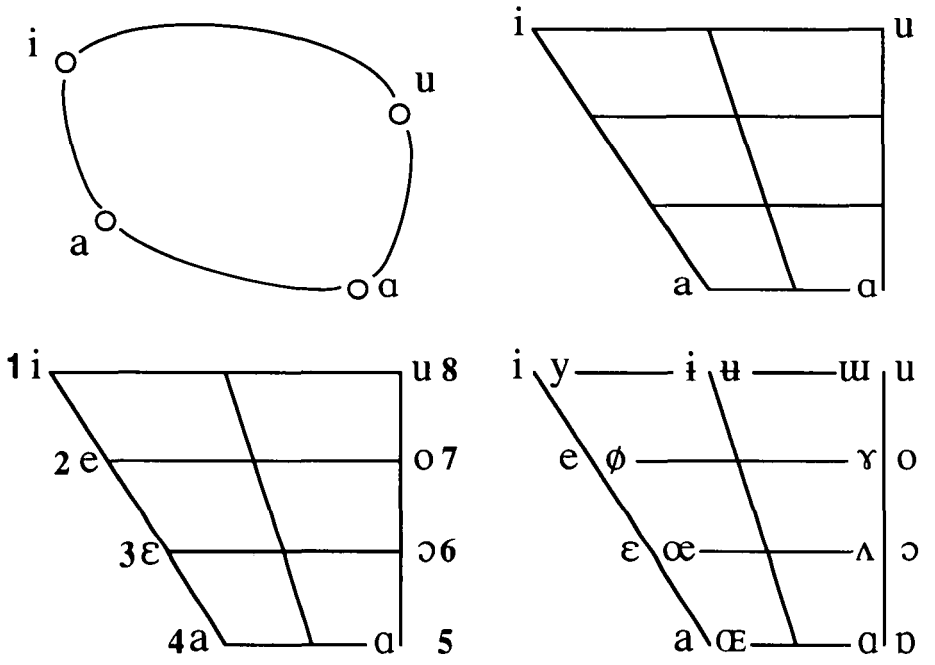


Figure 4. The vowel quadrilateral and cardinal vowels. Above, the relation between the vowel quadrilateral and the vowels shown in Figure 3; below, the primary cardinal vowels and all cardinal vowels.

There are now four defined vowel heights: [i] and [u] are close vowels, [e] and [o] are close-mid vowels, [ɛ] and [ɔ] are open-mid vowels, and [a] and [ɑ] are open vowels (note

that in this last pair the difference in letter shape is important, signifying a front vowel and a back vowel respectively). The vowel space can be seen to be taking on the form of a grid. The eight reference vowels are often numbered as shown, and are known as the 'primary cardinal vowels'. 'Cardinal' in this sense refers to points on which the system of description hinges. These and other vowels are exemplified in Section 3, though it must be remembered that only occasionally will a vowel in a language coincide exactly with one of these extreme peripheral reference points.

So far, lip activity has been largely ignored. In the back series of cardinal vowels ([ɑ ɔ o u]) lip-rounding progressively increases, from none on [ɑ] to close rounding on [u]. By convention unrounded vowels are placed to the left of the front or back line of the quadrilateral, and rounded vowels to the right. Conversely in the front series [a ɛ e i] the lips are neutral for [a], and become progressively more spread through the series to [i]. Lip activity is, however, independent of tongue position, and many languages exploit this in their vowel systems. To complete the descriptive system a set of 'secondary cardinal vowels' are therefore defined which differ only in lip position from their primary counterparts. These are shown in addition to the primary set in the fourth part of Figure 4. So, for example, of the close vowels [i y ɯ u], [i ɯ] have spread lips and [y u] have closely rounded lips; and of the open-mid vowels [ɛ œ ʌ ɔ], [ɛ ʌ] have slightly spread lips and [œ ɔ] have open rounding. A further two secondary cardinal vowels are defined; these are the close central vowels [ɨ] and [ʉ].

The complete set of IPA vowel symbols is shown in the quadrilateral on the Chart. In addition to the cardinal vowels already discussed, which lie on the outside edge of the quadrilateral, there are symbols for central vowels, and for vowels at a number of intermediate locations. Since the vowel space is continuous it is a matter of chance whether a vowel in a language exactly coincides with one of the reference points symbolized on the quadrilateral. If detailed phonetic description is required, most vowels have to be placed in relation to a reference vowel, for instance 'a vowel centralized and lowered from cardinal [e]'. This description can be symbolized by adding diacritics (see section 2.8) to the cardinal vowel symbol: [ë̞]. However, in most circumstances the simple transcription [e] will be fully sufficient.

2.7 Suprasegmentals

A number of properties of speech tend to form patterns which extend over more than one segment, and/or to vary independently of the segmental targets. This is particularly true of pitch, loudness, and duration. These properties are often referred to as 'suprasegmentals', and part of the process of phonetic analysis is the separation of these properties from the rest of the speech event. The IPA provides a separate set of symbols for suprasegmentals, to be found on the Chart at the bottom right corner.

Pitch variation, for instance, can operate over complete utterances to convey meaning additional to that of the words in what is generally termed 'intonation'. This is true in all languages, though the complexity of the intonational system varies across languages. The symbol [‖] can be used to mark the end of the domain of an intonation pattern, and [|] to demarcate a smaller unit. The symbols [↗ ↘] for 'global rise' and 'global fall' respectively may also be useful for intonation, although a complete intonational transcription will require symbols not provided on the IPA Chart.

In the majority of languages, usually called tone languages, one domain of pitch variation is the word or syllable, and this pitch variation serves to distinguish words in much the same way as their segmental make-up does. In Thai, for example, [kha: ˥˩] with a falling pitch (indicated by the diagonal part of the symbol following the segments) means 'servant' and [kha: ˩˦] with rising pitch means 'leg'.

The IPA has two different sets of symbols for indicating tones. In languages in which lexical contrasts are predominantly dependent on the pitch movement on each syllable, such as Thai and the various forms of Chinese, so-called tone letters are often used. These letters, as in the Thai examples, indicate the tone of the preceding syllable by a vertical stroke with a line preceding it. The vertical stroke is assumed to represent five possible pitch heights within the speaker's range, and the position of the line shows the height and movement (if any) of the pitch on the preceding syllable. The tone letters are often used to indicate general tone movements. For example, if there is only one falling tone in a language, and no strong reason to draw attention to the particular level of its endpoints, it can be noted as going from the highest to the lowest level. Thus a transcription of the Standard Chinese word for 'scold' is [ma˥˩], although most Chinese speakers will not produce this syllable with a fall extending through their whole pitch range. It is also possible to use the tone letters to show more detailed transcriptions for certain purposes. Thus, the Thai high tone can be transcribed with the symbol [˥]; but measurements of the fundamental frequency in high tone syllables show that there is actually a rise and a fall in syllables of this sort, so the tone could be represented as [˦˥˦].

The other IPA system for transcribing tone has often been used for languages in which tonal contrasts depend predominantly on the pitch height in each syllable. There are three diacritics, corresponding to high [́], mid [̄] and low [̀] tones, which can be placed above the segment bearing the tone. Thus the three tones in the West African language Yoruba can be transcribed as exemplified in the phrases [ōbá] ('he/she met'), [ōbā] ('he/she hid'), [ōbà] ('it perched'). Notice that these tone symbols must not be interpreted as iconic; that is, although the 'acute accent' [́] looks rising, it in fact means 'high'. To represent a rising tone it is necessary to combine a 'low' and a 'high', and similarly for other contour tones. So a syllable such as [e] occurring on a rising tone is [ě], and on a falling tone is [ê]. On the other hand the 'tone letters' such as [e˥] (meaning 'high') and [e˩˥] (meaning 'rising') are directly iconic.

The chart shows the tone letter [˥] as if it were equivalent to [̋], the extra-high tone symbol in the other set of symbols, and so on down the scale. But this is done only to simplify the layout of the chart. The two sets of symbols are not comparable in this way. The four tones of Standard Chinese are often symbolized as [ma˥] (mother), [ma˥˩] (scold), [ma˧˥] (hemp), [ma˨˩˦] (horse). If they were transcribed in the other system they would be [má mâ mǎ ma᷈].

The symbols [ꜛ] for upstep and [ꜜ] for downstep are used to indicate modifications (raising or lowering) of the pitch that have been indicated by ordinary tone symbols. Upstep occurs, for example, in Hausa in that the last of a series of high toned syllables before a low tone is pronounced with a higher pitch than the others. Thus the Hausa word [túrántʃí] (English) has three high tones with the same pitch when said by itself. In the phrase [túránꜛtʃí nè] (it is English), the raising of the high tone can be indicated by the modifier [ꜛ] as shown. Downstep occurs in the Akan word [ɔ̀kɔ́ꜜtɔ́] (crab), which has a downstepped high tone on the last syllable. This tone is demonstrably a high in that it

has the same pitch level as an initial high tone in a following word.

Symbols are also provided for indicating the relative prominence or stress of syllables, differing segmental length, and syllable divisions. The exact nature of syllable prominence or stress varies from language to language, but the IPA provides for up to three degrees of prominence to be indicated; in [ˌpærəsaɪˈkɒlədʒi] *parapsychology* the highest level occurs on the fourth syllable, and the second highest on the first syllable, while the remaining unmarked syllables are less prominent (a further division among these may be inferred from vowel quality, those syllables with [ə] being least prominent in English). Segmental length can be marked on a continuum from short to long as [ĕ e eˑ eː], though the possibility exists to show even greater length as [eːː]. Syllable divisions, which it may be useful to indicate for phonological reasons or where the syllable division determines phonetic difference as in [naɪ.t̪eɪt] *nitrate* versus [naɪt.ɹeɪt] *night-rate,* can be symbolized as shown. The use of suprasegmental symbols is further demonstrated in section 3.

2.8 Diacritics

Diacritics are small letters or other marks which can be added to a vowel or consonant letter to modify or refine its meaning in various ways. A letter and any diacritic or diacritics attached to it are regarded as a single (complex) symbol. The set of diacritics approved by the International Phonetic Association is given in the table at the bottom right of the Chart.

A number of diacritics deal with different aspects of phonation. Two are available to reverse the voicing value otherwise implied by any letter. Voiceless trills or nasals, for instance, for which there are otherwise no symbols, can be notated as [r̥], [n̥] etc. (some diacritics may be placed above a letter when a descender on the letter would interfere with legibility). Vowels which occur without voicing can also be indicated, e.g. [e̬]. More rarely employed is [̬] which indicates voicing in a symbol otherwise implying voicelessness. It sometimes indicates the spreading of voicing from an adjacent segment ('assimilation' of voicing), as in French [ʃaɡ̬ʒuʁ] *chaque jour* 'each day'. It is a moot point whether [ɡ̬] and [g] refer to phonetically identical sounds, and likewise [s] and [z̬]. It is possible that the distinction between [k] and [g] or between [s] and [z] can involve dimensions independent of vocal cord vibration, such as tenseness versus laxness of articulation, so that the possibility of notating voicing separately becomes important; but in any case, it can be convenient to be able to preserve the lexical shape of a word (e.g. French [ʃak] *chaque* 'each') while noting assimilation. The diacritic [ʰ] is used to indicate a voiceless release of air after a consonant, most commonly between a voiceless plosive and a vowel as in [tʰaɪ] *tie*. Two different phonation types which are used contrastively by some languages, creaky voice and breathy voice, can be indicated on vowels or consonants (see examples at the top of the second column of the diacritic table).

The diacritics shown in rows four to nine of the first column of the diacritic table, together with the diacritics for 'raised' and 'lowered' shown to the right in rows nine and ten, can be used to modify the lip or tongue position implied by a vowel symbol. Thus [u̜] indicates a vowel like cardinal [u] but with less close lip rounding, and (as seen in section 2.6 above) [e̞] indicates a vowel centralized and lowered from cardinal [e]. Vowel qualities between [u] and [ʉ] might be symbolized [u̟̽], indicating retraction relative to the central vowel, or (if nearer back than central) [u̟] or [ü] indicating fronting

relative to the back vowel. The diacritic for 'mid-centralized' indicates a quality displaced in the direction of the mid-central vowel [ə], so that [e̽] is generally equivalent to [ë], and [ɑ̽] to [ä].

The diacritics for 'raised' and 'lowered', when applied to a consonant symbol, change its manner category, so that [t̝] could be used to indicate an articulation like that of an alveolar plosive but one in which complete closure is not achieved, yielding a fricative-like sound (but lacking the grooved tongue shape of [s]) as in some Irish English pronunciations of the sound at the end of *right*. The diacritics for 'advanced' and 'retracted' are also commonly used to modify consonant place of articulation. So, for instance, a voiceless fricative at the front of the velar region could be symbolized [x̟], and a specifically postalveolar nasal [n̠].

The 'rhoticity' diacritic [˞] indicates a vowel with a specific auditory effect like that of the vowel in General American [fɝ] *fur*, probably caused by a constriction in the pharynx combined with an expansion of the space in the mouth in front of the tongue, either by curling the tongue tip up and back or lowering it and creating a narrow humped curve with the front of the tongue. In some languages the tongue root functions independently of other determinants of vowel quality, adjusting the width of the pharynx, and at the bottom right of the table there are two diacritics to indicate advancement and retraction of the tongue root. The syllabic diacritic is used to mark consonants which are acting as syllable nuclei, and the non-syllabic diacritic to mark vowels which are not fulfilling their customary syllabic role.

The dental diacritic (third column) modifies those consonant symbols found under 'alveolar' to indicate unambiguously a dental articulation. As noted in section 2, although only one symbol is provided in the consonant table (except in the fricative row), dental / alveolar / postalveolar can be distinguished as [n̪ n n̠] (postalveolar being marked by the 'retracted' diacritic). The linguolabial diacritic, which is used to symbolize an otherwise omitted (and very rare) consonantal type, indicates a sound made with the tip or blade of the tongue against the upper lip. The diacritic is used to modify the relevant alveolar consonant symbol. The diacritics for 'apical' and 'laminal' make specific which part of the frontmost area of the tongue is making an articulation: the tip (apical), or the blade (laminal).

Secondary articulations are narrowings of the vocal tract which are less narrow than the main one producing a consonant. The names palatalization, velarization, and pharyngealization, make explicit where the narrowing is. In one sense a secondary articulation is the superimposition of a close-vowel-like articulation on a consonant – [i] for palatalization, symbolized for instance [tʲ], [ɯ] for velarization ([tˠ]), and [ɑ] for pharyngealization ([tˤ]). The fourth secondary articulation for which a diacritic is provided, labialized (e.g. [tʷ]), is slightly problematic. In principle labialization should mean simply a reduction in the opening of the lips, but the diacritic chosen reflects the fact that such a reduction is often accompanied by a velar constriction. [ʷ] is probably best regarded, then, as a diacritic for labial-velarization. These superscript diacritics which are placed after the letter look rather as if they imply a sequence of events; but strictly the notation means that the secondary articulation is simultaneous with the consonant. This is unlike the case of the aspiration diacritic (e.g. [tʰ]) where the plosive and the aspiration are sequential. The simultaneity of the secondary articulation is clearer from the alternative diacritic for symbolizing velarization or pharyngealization, [˗], which

is placed through the consonant symbol in question (often to the detriment of legibility). Nasalization, despite the similarity of name, is not a secondary articulation in the same sense, but the addition of the resonances of the nasal cavities to a sound. Vowels, and some consonants (e.g. [ɪ̃]), can be nasalized.

Finally, there are three diacritics in the third column dealing with release ('nasal release', 'lateral release', and 'no audible release'). All three show that a stop consonant has not been released into a vowel. Instead the air escape is through the nose (e.g. [bʌtⁿn̩] *button*), round the side of the tongue (e.g. [bɒtˡt̚] *bottle*), or the air is not released until a later sound [ɹæg̚bæg] *ragbag*. The use of diacritics is further exemplified in Section 3.

2.9 Other symbols

These symbols are included in their own section of the Chart for presentational convenience. The section contains several consonant symbols which would not fit easily into the 'place and manner' grid of the main consonant table. In some cases, such as the epiglottals and the alveolo-palatals, no column is provided for the place of articulation because of its rarity and the small number of types of sounds which are found there. In other cases, such as [w], the sound involves two places of articulation simultaneously, which makes it inconvenient to display in the table. If separate columns for all consonants with two places of articulation were provided, the size of the grid would become unmanageable. Most consonants that involve two simultaneous places of articulation are written by combining two symbols with the 'tie-bar' [͡], for example [k͡p] which represents a voiceless labial-velar plosive.

3 Guide to IPA notation

3.1 Exemplification of the symbols

The general value of the symbols in the chart is listed below. In each case a symbol can be regarded as a shorthand equivalent to a phonetic description, and a way of representing the contrasting sounds that occur in a language. Thus [m] is equivalent to 'voiced bilabial nasal', and is also a way of representing one of the contrasting nasal sounds that occur in English and other languages.

When a symbol is said to be suitable for the representation of sounds in two different languages, it does not necessarily mean that the sounds in the two languages are identical. Thus [p] is shown as being suitable for the transcription of *pea* in English, and also for *pis* in French; similarly [b] is shown as being suitable for the transcription of *bee* in English, and also for *bis* in French; but the corresponding sounds are not the same in the two languages. The IPA has resources for denoting the differences, if it is necessary to do so, as illustrated below; but at a more general level of description the symbols can be used to represent the sounds in either language

. The symbols are exemplified in the order in which they appear on the chart; they are discussed using the terms given as headings for the rows and columns. English (Eng.) and French (Fr.) examples are given when unambiguous. Where practical, an orthographic version of the exemplifying word is provided, in italics. English glosses of words in other languages are given in parentheses. The languages used for exemplification are identified at the end of the list

PLOSIVES

p	Eng. *pea* [pi]; Fr. *pis* [pi] (worst)	b	Eng. *bee* [bi]; Fr. *bis* [bis] (encore)
t	Eng. *tea* [ti]; Fr. *thé* [te] (tea)	d	Eng. *deep* [dip]; Fr. *dix* [dis] (ten)
ʈ	Hindi [ʈal] (postpone)	ɖ	Hindi [ɖal] (branch)
c	Hungarian *tyùk* [cuk] (pen)	ɟ	Hungarian *gyùr* [ɟur] (to knead)
k	Eng. *key* [ki]; Fr. *qui* [ki] (who) K'ekchi [kaʔa] (grindstone)	g	Eng. *geese* [gis]; Fr. *guide* [gid] (guide)
q	K'ekchi [qa] (our)	ɢ	Farsi [ɢar] (cave)
		ʔ	Hawaiian *Hawai'i* [hawaiʔi] (place name), *ha'a* [haʔa] (dance)

The plosives in the lefthand column above are said to be voiceless, and those on the right are said to be voiced. The degree of voicing may vary considerably. The voiceless consonants may be not only voiceless, but also aspirated; and the voiced consonants may be voiced throughout their duration, or may have voicing during only part of that time. Usually the use of a pair of symbols such as [p] and [b] in a given language signifies only that there is a contrast in the degree of voicing within that pair of sounds.

NASALS

m	Eng. *me* [mi]; Fr. *mis* [mi] (put)
ɱ	Eng. *emphasis* [ɛɱfəsɪs]
n	Eng. *knee* [ni]; Fr. *nid* [ni] (nest)
ɳ	Fr. *agneau* [aɲo] (lamb); Malayalam [kɛɳɳi] (link in a chain)
ɲ	Malayalam [kɐɲɲi] (boiled rice and water)
ŋ	Eng. *hang* [hæŋ]
ɴ	Inuit [saaɴɴi] (his bones)

Note that the symbols [t, d, n] listed above, and the symbols [r, ɾ, ɬ, ɮ, ɹ, l] which will be exemplified below, all represent sounds that can be either dental, or alveolar, or postalveolar. If there is a need to represent specifically one of these places of articulation, there are IPA resources for doing so, which will be exemplified below.

TRILLS

ʙ	Kele [mbʙuen] (fruit)
r	Spanish *perro* [pɛro] (dog)
ʀ	Fr. *rouge* [ʀuʒ] (red); Southern Swedish *ras* [ʀas] (breed)

Note: most forms of English, French, German, Swedish do not have trills.

TAPS OR FLAPS

ɾ	Spanish *pero* [pero] (but); Am. Eng. *atom* ['ærəm]
ɽ	Hausa *shaara* [ʃàːɽa] or [ʃàːɻa] (sweeping) (Some speakers of Hausa have [ɽ] and others have [ɻ].)

FRICATIVES

The fricatives in the lefthand column below are voiceless, and those on the right are voiced. To a somewhat lesser extent than in the case of the plosives, the degree of voicing may vary.

ɸ	Ewe *e fa* [é ɸá] (he polished);	β	Ewe *eβẹ* [èβè] (Ewe)
f	Eng. *fee* [fi]; Fr. *fixe* [fiks]; Ewe *e fa* [é fá] (he was cold)	v	Eng. *vat* [væt]; Fr. *vie* [vi] (life); Ewe *ẹvẹ* [èvè] (two)
θ	Eng. *thief* [θif]	ð	Eng. *thee* [ði]
s	Eng. *see* [si]; Fr. *si* [si] (yes)	z	Eng. *zeal* [zil]; Fr. *zéro* [zeʁo] (zero)
ʃ	Eng. *she* [ʃi]; Fr. *chic* [ʃik] (chic)	ʒ	Eng. *vision* [vɪʒn]; Fr. *joue* [ʒu] (cheek)
ʂ	Standard Chinese *sha* [ʂa] (to kill)	ʐ	Standard Chinese *ráng* [ʐaŋ] (to assist)
ç	German *ich* [ɪç] (I);	ʝ	Eng. variant of [j] in *ye* [ʝi]
x	German *bach* [bax] (stream)	ɣ	Greek *λαγα* ['lɑɣɑ] (to purify)
χ	Hebrew [maχar] (he sold)	ʁ	Fr. *riz* [ʁi] (rice)
ħ	Hebrew [ħor] or [ʜor] (hole)	ʕ	Hebrew [ʕor] or [ʢor] (skin)

Hebrew, Arabic and other Afro-Asiatic languages with pharyngeal consonants do not distinguish between [ħ] and [ʜ]; nor do they distinguish between [ʕ] and [ʢ]. Most speakers of these languages use the epiglottal fricatives [ʜ] and [ʢ].

h	Eng. *he* [hi]	Eng. *ahead* [əɦɛd]

[ɦ] represents a breathy voiced sound, rather than an ordinary voiced sound.

LATERAL FRICATIVES

ɬ	Zulu *hlanza* [ɬânzà] (vomit); Welsh *llan* [ɬan] (church)	Zulu *dlala* [ɮálà] (play)

APPROXIMANTS

ʋ	Hindi [nɔʊ̃ʋ̃] (ninth)
ɹ	Eng. *read* [ɹid]
ɻ	Hausa *shaara* [ʃàːɽa] or [ʃàːɻa] (sweeping) (Some speakers of Hausa have [ɽ] and others have [ɻ].)
j	Eng. *yes* [jɛs]; Fr. *yeux* [jø] (eyes)
ɰ	Spanish variant of [ɣ] in *paga* [paɣa] (pay)

LATERAL APPROXIMANTS

l	Eng. *leaf* [lif]; Fr. *lit* [li] (bed)
ɭ	Tamil: [vaɭ] (sword)
ʎ	Italian *figlio* [fiʎʎɔ] (son)
ʟ	Mid-Waghi *aglagle* [aʟaʟe] (dizzy)

NON-PULMONIC CONSONANTS

CLICKS

ʘ	!Xóõ [k͡ʘôõ] (dream)
ǀ	Xhosa *ukucola* [ukúk͡ǀola] (to grind finely)
ǃ	Xhosa *ukuqoba* [ukúk͡ǃoba] (to break stones)
ǂ	!Xóõ [k͡ǂàã] (bone)
ǁ	Xhosa *ukuxhoba* [ukúk͡ǁʰoba] (to arm oneself)

VOICED IMPLOSIVES

ɓ Sindhi [ɓəni] (field)
ɗ Sindhi [ɗɪnu] (festival)
ʄ Sindhi [ʄətu] (illiterate)
ɠ Sindhi [ɠənu] (handle)
ʛ (voiced uvular implosives do not occur in any well known language)

EJECTIVES

p' Amharic [p'ap'as] (bishop [loan word])
t' Amharic [t'ɨl] (fight)
k' Amharic [k'alat] (word)
s' Amharic [s'ahaɪ] (sun)

VOWELS

The symbols on the vowel chart can be regarded as providing reference points in the vowel space. They can also be used to represent vowel qualities generally in the area of the corresponding reference points. With the vowel symbols it is especially important to note that they may represent different sounds in different languages. For example, [e] may be used for the vowel in the English word *hay* or in the French word *été* (summer), despite the fact that the English vowel may be slightly diphthongal and more open than the French vowel.

Because of their status as reference points, it is difficult to illustrate some of the vowel symbols appropriately in terms of particular languages; this is particularly true of the mid-central vowels [ə, ɵ, ɜ, ɞ]. The symbols [ə] and [ɐ] are available for representing vowels in the mid-central and lower central regions. [a] is often used for an open central vowel. No language has been found that uses the open front reference quality [œ].

The symbols in the lefthand column below specify vowels with more rounded lips than the corresponding symbols in the righthand column.

i	Eng. *heed* [hid]; Fr. *lit* [li] (bed)	y	Fr. *lu* [ly] (read)
ɪ	Eng. *hid* [hɪd]	ʏ	German *Fusse* ['fʏsə] (feet); Swedish *nytta* [nʏtta] (use, n.)
e	Eng. *hay* [he]; Fr. *les* [le] (the, pl.)	ø	Fr. *peu* [pø] (few)
ɛ	Eng. *head* [hɛd]; Fr.*lait* [lɛ] (milk)	œ	Fr. *peur* [pœʁ] (fear)
æ	Eng. *had* [hæd]		
a	Fr. *patte* [pat] (paw)		
ɑ	Eng. *father* [fɑðə(ɹ)]; Fr. *pâte* [pɑt] (pastry). (Most French speakers do not distinguish [a] and [ɑ].)	ɒ	British Eng. *bother* [bɒðə]
ʌ	Vietnamese [ʌŋ] (favor). (This symbol is sometimes used for a different vowel, the central vowel in Eng. *hut* [hʌt]).	ɔ	British Eng. *caught* [kɔt]; Fr. *botte* [bɔt] (bootie); Vietnamese [tɔ] (large)
ɤ	Vietnamese [tɤ] (silk)	o	Fr. *lot* [lo] (share); Vietnamese [to] (soup bowl)
		ʊ	Eng. *book* [bʊk]
ɯ	Vietnamese [tɯ] (fourth)	u	Eng. *who* [hu]; Fr. *loup* [lu] (wolf); Vietnamese [tu] (to drink)
ɨ	Korean [kʰɨn] (large)	ʉ	Norwegian [bʉt] (shack)

OTHER SYMBOLS

ʍ Scottish Eng. *whether* [ʍɛðəɹ]
w Eng. *weather* [wɛðə(ɹ)]; Fr. *oui* [wi] (yes)
ɥ Fr. *huit* [ɥit] (eight)
ʜ, ʢ See discussion of [ħ] and [ʕ] in the fricative section above.
ʡ Some speakers of Hebrew use [ʡ] instead of [ʜ] in [ʡor] (hole)
ɕ Polish *Basia* [baɕa] (Barbara, dim.)
ɺ KiChaka [iɺaa] (to dress oneself)
ʃ͡ʒ Some dialects of Swedish *schal* [ʃ͡ʒal] (scarf) (Note: it is not clear that any Swedish speakers actually have simultaneous [ʃ] and [x].)

Affricates and double articulations

k͡p, t͡ʃ, etc Eng. *chief* [t͡ʃif], Yoruba *apa* [ak͡pá] (arm), Tswana *tsetse.* [t͡sét͡sé] (tsetse fly)

SUPRASEGEMENTALS

In general, only one or two degrees of stress are marked

ˈˈ may be used to indicate extra strong stress.
ˈ Eng. *phonetics* [fəˈnɛtɪks]
ˌ Eng. *phonetician* [ˌfoʊnəˈtɪʃən]

Length is not contrastive (at least, without concomitant changes in quality) in English and French, but may be used to show allophonic differences.

ː Eng. *bead* [biːd]; Fr. *vie* [viː] (life);
ˑ Eng. *beat* [biːt]; Fr. *vite* [viːt] (quickly)
◡ Eng. *police* [pə̆lis]

In other languages length may be contrastive for vowels and/or consonants.

ː Finnish [tule] (come, imper.), [tulːe] (may be coming, neg.), [tuleː] (comes), [tulːeː] (may be comes),

White spaces can be used to indicate word boundaries. Syllable breaks can be marked when required. The other two boundary symbols are used to mark the domain of larger prosodic units. There is also a linking symbol that can be used for explicitly indicating the lack of a boundary.

. Eng. *lamb prepared* [ˈlæm.pɹə.ˈpɛəd] *lamp repaired* [ˈlæmp.ɹə.ˈpɛəd]

| Eng. *Jack, preparing the way, went ahead* [ˈdʒæk | prəˈpɛəɹɪŋ ðə ˈweɪ | wɛnt əˈhɛd ‖]
‖ Fr. *Jacque, préparant le sol, tomba* [ʒak | pʁepaʁɑ̃ lə sol | tɔ̃ba ‖]
(Jack, preparing the soil, fell down)
‿ Fr. *petit ami* [pətit‿ami] (boyfriend)

As explained in the previous section, there are two alternative systems of tone transcription. The chart shows these two systems as if there were direct equivalencies between them. However, they are usually used in different ways.

̋ Bariba [nɛ̋ ná nā kɔ̀] (I am the one who came)
́ Yoruba *o bá* [ō bá] (he/she met)
̄ Yoruba *o ba* [ó bā] (he/she hid)
̀ Yoruba *o bà* [ō bà] (it perched)
̏ Trique [ɛ̏ʔ] (bitter)

˙ It is also possible to combine these symbols so that, for example, [ˆ] represents a high tone followed by a low tone, i.e. a falling tone. Similarly [ˇ] represents a rising tone, and [˝] and [ˋ] represent mid-rising and mid-falling tones.

There are two symbols for showing that subsequent tones may be a step lower or higher. The introduction of a downstep is phonologically contrastive in the Igbo example below, but the Hausa upstep indicates only a predictable allophone.

ꜜ Igbo *ụlọ anyi* [ú ꜜlɔ́ ꜜáɲí] (our house)

ꜛ Hausa [túránꜛtʃí nè] (it is English)

The use of the other set of symbols is illustrated below.

	CHINESE (STANDARD)	CANTONESE	THAI
˥	[ma˥] (mother)	[ɕik˥] (to know)	
˦			
˧		[ɕi˧] (to try); [ɕit˧] (to release)	
˨		[ɕi˨] (matter); [ɕik˧] (to eat)	
˩		[ɕi˩] (time)	
˥˩	[ma˥˩] (scold)		[n̪a:˥˩] (face)
˩˥		[ɕi˩˥] (poem)	
˧˥		[ɕi˧˥] (city)	
˦˥	[ma˦˥] (hemp)	[ɕi˦˥] (to cause)	[n̪a:˦˥] (aunt)
˧˦˥	[ma˧˦˥] (horse)		[n̪a:˧˦˥] (thick)
˨˩			[n̪a:˨˩] (name)
˩˦			[n̪a:˩˦] (field)

The symbols for general upglide and downglide are appropriate for use in many languages.

↗ Eng. *No?* [↗ noʊ]

↘ Eng. *No.* [↘ noʊ], *When he came, I went,* ['wɛn hi ↗ k'em aɪ ↘ 'wɛnt]

DIACRITICS

The diacritics allow symbols to be created to represent many additional types of sounds. In the representation of many languages (including English) diacritics are necessary only when making detailed transcriptions.

◌̥ Burmese [n̥á] (nose).

The voiceless diacritic can also be used to show that a symbol that usually represents a voiced sound in a particular language on some occasions represents a voiceless sound, as in a detailed transcription of conversational English *Please say* ... as [pl̥iz̥ se ...]

◌̬ The voiced diacritic can be used to show that a symbol that usually represents a voiceless sound in a particular language on some occasions represents a voiced sound, as in a detailed transcription of conversational English *back of* as [bæk̬ əv].

◌ʰ Hindi [kʰan] (mine).

Detailed transcription of English *pea, tea, key* [pʰi, tʰi, kʰi]

◌̼ Assamese [pɔ̼t] (to bury).

◌̹ In some forms of British English (RP) [ɔ] may be over-rounded, e.g., *caught* [kɔ̹t].

◌̜ In Californian English [ʊ] is under-rounded, e.g., *good* [gʊ̜d].

◌̰ Eng. [k̰] in *key* [k̰i].

◌̠ Eng. [t̠] in *tree* [t̠ɹi].

◌̈	Eng. [ë] in *ale* [ëɫ].
◌̽	Eng. [o̽] in *November* [no̽vɛmbə(ɹ)].
◌̩	Eng. [l̩] in *fiddle* [fɪdl̩]
◌̯	Spanish *poeta* [po̯'eta] (poet)
◌˞	Am. Eng. [ɚ] in *bird* [bɚd]. This sound can also be written [ɹ̩]
◌̤	Hindi [kʊm̤ar] (potter)
◌̰	Mazatec *nda'* [ndæ̰] (buttocks)
◌̼	Tangoa [t̼et̼e] (butterfly)
ʷ	Eng. [t] in *twin* [tʷɪn]; Cantonese [kʷɔk] (family name)
ʲ	Russian [matʲ] (mother)
ˠ	Russian [lˠɨsij] (bald)
ˤ	Arabic [sˤad] (name of the letter)
◌̴	Eng. [l] in *hill* [hɪɫ]
◌̝	Some forms of South African Eng. [ɹ] in *dry* [dɹ̝aɪ]
◌̞	Danish [ð̞] in *lade* [læð̞ə] (barn)
◌̘	Igbo *óbì* [ó̘bì̘] (heart)
◌̙	Igbo *ụ̀bị̀* [ù̙bì̙] (poverty of ability)
◌̪	Eng. [t̪] in *width* [wɪt̪θ]
◌̺	Ewe *e da* [é d̺à] (he throws)
◌̻	Ewe *e ḍa* [é d̻à] (he cooks)
◌̃	Fr. *fin* [fɛ̃] (end)
ⁿ	Russian [dⁿno] (bottom)
ˡ	Navajo [dˡőőʔ] (prairie dog)
◌̚	Eng. [k̚] in *act* [æk̚t]

3.2 Languages used for exemplification

The principal country in which a language is spoken is given only when it is not apparent from the name.

Amharic, Afro-Asiatic, spoken in Ethiopia.
Arabic, Afro-Asiatic, spoken in many North African and Middle Eastern countries.
Assamese, Indo-European, spoken in India
Bariba, Niger-Congo, spoken in Nigeria.
Burmese, Sino-Tibetan, spoken in Myanmar.
Cantonese, Sino-Tibetan, spoken in China.
Chinese (Standard), Sino-Tibetan.
Danish, Indo-European.
English, Indo-European.
Ewe, Niger-Congo, spoken in Ghana and Toga.
Farsi, Indo-European, spoken in Iran.
Finnish, Finno-Ugric.
French, Indo-European.
German, Indo-European.
Greek, Indo-European.
Hausa, Afro-Asiatic, spoken in Nigeria.
Hawaiian, Austronesian.
Hebrew, Afro-Asiatic, spoken in Israel.
Hindi, Indo-European, spoken in India.
Hungarian, Finno-Ugric.

Igbo, Niger-Congo, spoken in Nigeria.
Inuit, Eskimo-Aleut.
Italian, Indo-European.
K'ekchi, Mayan, spoken in Guatemala.
Kele, Austronesian, spoken in Papua New Guinea.
KiChaka, Niger-Congo, spoken in Tanzania.
Korean, Ural-Altaic.
Malayalam, Dravidian, spoken in India.
Mazatec, Oto-Manguean, spoken in Mexico.
Mid-Waghi, Papuan, spoken in Papua New Guinea.
Navajo, Na-Dene, spoken in United States.
Norwegian, Indo-European.
Polish, Indo-European.
Russian, Indo-European.
Sindhi, Indo-European, spoken in Pakistan.
Spanish, Indo-European.
Swedish, Indo-European.
Tamil, Dravidian, spoken in India.
Tangoa, Austronesian, spoken in Vanuatu.
Thai, Tai-Kadai.
Trique, Oto-Manguean, spoken in Mexico.
Tswana, Niger-Congo, spoken in Botswana.
Vietnamese, Austro-Asiatic.
Welsh, Indo-European.
Xhosa, Niger-Congo, spoken in South Africa.
!Xóõ, Khoisan, spoken in Botswana.
Yoruba, Niger-Congo, spoken in Nigeria.
Zulu, Niger-Congo, spoken in South Africa.

4 The phonemic principle

From its earliest days (see Appendix 4) the International Phonetic Association has aimed to provide 'a separate sign for each distinctive sound; that is, for each sound which, being used instead of another, in the same language, can change the meaning of a word'. This notion of a 'distinctive sound' is what became widely known in the twentieth century as the phoneme. Its history is far longer, though. For instance, the phonemic principle is implicit in the invention of alphabetic writing. However a lot of languages, such as English, have spelling systems in which the relation between phonemes and letters of the alphabet have become obscured. This very fact was a motivation for the creation of a universally agreed system of phonetic notation. So, in English, the IPA provides a symbol /k/ which stands unambiguously for the phoneme which is variously written as c (*car*), k (*kettle*), ck (*back*), ch (*monarch*), q (*quick*), and in other ways.

Each language can be analyzed as having an inventory of phonemes. This inventory may range in size from around a dozen phonemes to nearer a hundred depending on the language. Conventionally, as in the English example above, symbols for the phonemes of a language are placed within oblique lines.

In general, the symbol for a phoneme will be an unmodified letter of the IPA, but letters may also be combined to make a phoneme symbol (for instance /tʃ/, as at the beginning and end of English *church*; if necessary the phonological unity of the two segments can be shown by a tie bar: /t͡ʃ/). Diacritics may also be employed to create symbols for phonemes, thus reducing the need to create new letter shapes. This may be convenient in particular when a subset of the phonemic system of a language shares a phonetic property, as in the case of the nasalized vowel phonemes of French /ɛ̃ œ̃ ɑ̃ ɔ̃/ (which when they stand alone represent the French words *hein* 'huh', *un* 'a, one', *an* 'year', and *on* 'one, impersonal pn.').

The use of the phrase 'distinctive sound' above implies that there are other sounds which do not change the identity of a word, sounds which are not 'distinctive' in this technical sense. Central to the notion of the phoneme is the recognition that many finely distinct sounds can be phonetically identified which do not have the word-distinguishing role of, say, English /k/ and /t/ (as in /ki/ *key* versus /ti/ *tea*). For instance the English /k/ phoneme is made with a tongue closure further forward in the mouth before a front vowel (such as the /i/ of *key*) than before a back vowel (such as the /ɔ/ of *caw*). But crucially it is not possible, in English, to swap these two varieties of /k/ to make two new words, so the two varieties of /k/ are not 'distinctive' in English.

A phoneme can be regarded as an element in an abstract linguistic system, an element which has to be realized in the physical world by an acoustic signal produced by vocal activity. Variation arises in the process of realization. Some of this variation can be attributed to the influence of adjacent sounds affecting the articulation, so for instance the /k/ of *key* may be thought of as being further forward to facilitate integration with the following /i/; while in other cases the variation seems to be merely a language-specific but arbitrary habit. Variant realizations of a phoneme are known as its allophones.

The IPA aims not only to provide symbols which can unambiguously represent phonemic inventories, but also to be able to represent details of phonetic realization. The above example could be represented as [k̟ʰiː] *key* and [k̠ʰɔː] *caw*, where the 'Subscript plus' and 'Under-bar' indicate advanced and retracted articulation respectively (see Appendix 3 for diacritic names). A further detail of realization is also indicated here – the 'Superscript H' indicates aspiration, a delay in the onset of voicing after the voiceless plosive, characteristic of such plosives at the beginning of stressed syllables in many varieties of English. Square brackets are used conventionally to make clear that a symbol or sequence of symbols represents phonetic realizations rather than phonemes.

In providing the means to show the detail of phonetic realization in a given language, the IPA also achieves the delicacy of notation needed to compare the phonetic detail of different languages. For instance, although a phonemic representation /tru/ might be suitable for the English word *true* or the French word *trou*, the difference in pronunciation of the two words is reflected in phonetically more detailed representations such as [t̪ɹʉ] (*true*) and [t̪ʁu] (*trou*). These show the dental realization of /t/ in French compared to the alveolar realization in English, here retracted under the influence of the following postalveolar; the uvular realization of /r/ in French compared to the postalveolar realization in English, here, furthermore, devoiced after the voiceless plosive; and the fully back realization of /u/ in French compared to the central realization in (many varieties of) English.

THE INTERNATIONAL PHONETIC ALPHABET (revised to 1993, corrected 1996)

CONSONANTS (PULMONIC)

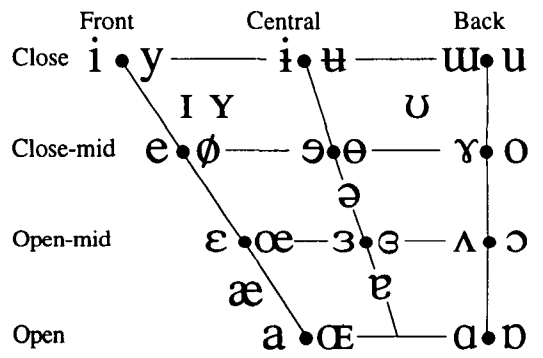
	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap				ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

CONSONANTS (NON-PULMONIC)

Clicks		Voiced implosives		Ejectives	
ʘ	Bilabial	ɓ	Bilabial	’	Examples:
ǀ	Dental	ɗ	Dental/alveolar	pʼ	Bilabial
ǃ	(Post)alveolar	ʄ	Palatal	tʼ	Dental/alveolar
ǂ	Palatoalveolar	ɠ	Velar	kʼ	Velar
ǁ	Alveolar lateral	ʛ	Uvular	sʼ	Alveolar fricative

VOWELS



Where symbols appear in pairs, the one to the right represents a rounded vowel.

OTHER SYMBOLS

ʍ	Voiceless labial-velar fricative	ɕ ʑ	Alveolo-palatal fricatives
w	Voiced labial-velar approximant	ɺ	Alveolar lateral flap
ɥ	Voiced labial-palatal approximant	ʜ	Simultaneous ʃ and x
ʜ	Voiceless epiglottal fricative		Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary.
ʢ	Voiced epiglottal fricative		k͡p t͡s
ʡ	Epiglottal plosive		

DIACRITICS

Diacritics may be placed above a symbol with a descender, e.g. n̥

◌̥	Voiceless	n̥ d̥	◌̤	Breathy voiced	b̤ a̤
◌̬	Voiced	s̬ t̬	◌̰	Creaky voiced	b̰ a̰
◌ʰ	Aspirated	tʰ dʰ	◌̼	Linguolabial	t̼ d̼
◌̹	More rounded	ɔ̹	◌ʷ	Labialized	tʷ dʷ
◌̜	Less rounded	ɔ̜	◌ʲ	Palatalized	tʲ dʲ
◌̟	Advanced	u̟	◌ˠ	Velarized	tˠ dˠ
◌̠	Retracted	e̠	◌ˤ	Pharyngealized	tˤ dˤ
◌̈	Centralized	ë	◌̴	Velarized or pharyngealized	ɫ
◌̽	Mid-centralized	e̽	◌̝	Raised	e̝ (ɹ̝ = voiced alveolar fricative)
◌̩	Syllabic	n̩	◌̞	Lowered	e̞ (β̞ = voiced bilabial approximant)
◌̯	Non-syllabic	e̯	◌̪	Advanced Tongue Root	e̪
◌˞	Rhoticity	ɚ a˞	◌̺	Retracted Tongue Root	e̺

◌̃	Nasalized	ẽ	
◌ⁿ	Nasal release	dⁿ	
◌ˡ	Lateral release	dˡ	
◌̚	No audible release	d̚	

SUPRASEGMENTALS

ˈ	Primary stress	ˌfoʊnəˈtɪʃən	
ˌ	Secondary stress		
ː	Long	eː	
ˑ	Half-long	eˑ	
◌̆	Extra-short	ĕ	
|	Minor (foot) group		
‖	Major (intonation) group		
.	Syllable break	ɹi.ækt	
‿	Linking (absence of a break)		

TONES AND WORD ACCENTS

LEVEL			CONTOUR	
é̋ or ˥	Extra high		ě or ˩˥	Rising
é ˦	High		ê ˥˩	Falling
ē ˧	Mid		ě ˩˦	High rising
è ˨	Low		ě ˨˩	Low rising
ȅ ˩	Extra low		ê ˦˥	Rising-falling
↓	Downstep		↗	Global rise
↑	Upstep		↘	Global fall

5 Broad and narrow transcriptions

A connected text represented in terms of phonemes is known as a phonemic transcription, or, almost equivalently, a broad transcription. The term broad sometimes carries the extra implication that, as far as possible, unmodified letters of the roman alphabet have been used. This restriction may facilitate printing, and might be considered particularly if a phonemic transcription is to form the basis of a writing system. Under this definition a transcription of English *hideout* as /haidaut/ would be broad, while /haɪdaʊt/ would not be because it introduces letters to the symbol for the phoneme /aɪ/ and the phoneme /aʊ/ which are not absolutely necessary for the unambiguous representation of the phonemes of English, but which may be desirable to remind the reader of the phonetic realization of these phonemes. Frequently, though, 'broad' is used merely as a way of referring to transcriptions which are phonemic, regardless of the letter shapes used to represent the phonemes. Phonemic transcriptions are one type of 'systematic' transcription, meaning they require the phonological patterns or 'system' of a language to be known.

The term narrow transcription most commonly implies a transcription which contains details of the realization of phonemes. There are two ways in which such a transcription may come about. If a transcription is made in circumstances where nothing can be assumed about the phonological system, it is necessary to include all phonetic details because it is not clear which phonetic properties will turn out to be important. The transcription would be made taking into account only the phonetic properties of the speech. This type of narrow transcription, as might be made in the first stages of fieldwork, or when transcribing disordered speech, is sometimes called an impressionistic transcription or a general phonetic transcription. If an impressionistic transcription were made of an utterance of the English phrase *check the lens well* it might be [tʃe̝ʔ͡kð̞əlɛ̃nzwæ̠lˤ]. This includes a glottalized velar stop, a dental approximant (the lowering diacritic indicating that the stricture was not close enough to cause frication), a pharyngealized lateral (often described more generally as 'dark', or alternatively as velarized ([lˠ], but likely to be pharyngealized in the case of many speakers), and three different vowel qualities in the stressed syllables, even though these vowels are the same in phonemic terms.

The other kind of narrow transcription containing realizational information is termed allophonic. If the relevant phonological system is known, a transcription can be devised which includes any number of additional symbols to indicate the phonetic realizations of the phonemes, i.e. their allophones. An allophonic transcription is also known as a systematic narrow transcription. In the knowledge that a possible phonemic analysis of *check the lens well* is /tʃɛkðəlɛnzwɛl/, one allophonic or systematic narrow transcription would, perhaps surprisingly, be [tʃe̝ʔ͡kð̞əlɛ̃nzwæ̠lˤ], that is, one which is identical to the impressionistic transcription in the previous paragraph, incorporating all the phonetic detail which can be heard. The difference is that now it would be possible to express, in conventions, the relation between the allophones transcribed and the phonemes which they realize. Alternatively, it is possible (and customary) to be selective about the information which is explicitly incorporated into the allophonic transcription. The choice might be made, for instance, to leave out the information about vowel height (the closer vowel in *check* is triggered by the high tongue body position of the following velar, and the more open and retracted quality in *well* caused mainly by the pharyngealization of the following lateral), and about vowel nasalization (which is very general before a following

nasal), giving a transcription which focuses on consonant realization: [tʃɛʔ͡kð̪əlɛnzwɛlˤ]. Minimally, if the focus of interest were glottalization of plosives, the allophonic transcription could be [tʃɛʔ͡kðəlɛnzwɛl], or if the focus were the 'dark' lateral, [tʃɛkðəlɛnzwɛlˤ]. These last two transcriptions look superficially very like a phonemic transcription, but they are in principle different because information has been included (albeit sparingly) which is not required for the unambiguous representation of the words. Narrowness is regarded as a continuum, so that [tʃɛkðəlɛnzwɛlˤ] might be regarded as a slightly narrow (or 'narrowed') transcription, and [tʃe̞ʔ͡kð̪əlɛ̃nzwæ̠lˤ] as very narrow. (In all these transcriptions, no spaces between words have been included. This is inevitable in an impressionistic transcription where it is not yet known how the utterance divides into words. In phonemic and allophonic transcriptions it is common to include spaces to aid legibility, but their theoretical validity is problematic.)

Any transcription is connected to a speech event by a set of conventions. In the case of an impressionistic ('general phonetic') transcription, the conventions are precisely those lying behind the IPA Chart, indicating for instance that the phonetic value of [ʔ͡k] is a simultaneous velar and glottal closure. In the case of a phonemic transcription, the conventions also include the 'phonological rules' of the particular language which determine the realization of its phonemes, such as the fact that for some varieties of English the lateral phoneme /l/ is realized with accompanying pharyngealization ([lˤ]) or velarization ([lˠ]) when not followed directly by a vowel or /j/ in the same word. Likewise, the realizational information which is not explicit in a particular allophonic transcription is, in principle, provided by conventions.

6 IPA transcriptions for a language

There can be many systems of phonemic transcription for the same variety of a language, all of which conform fully to the principles of the IPA. Sometimes the differences between the systems results from the fact that more than one phonetic symbol may be appropriate for a phoneme. For instance the vowel phoneme of *get* in standard southern British English has allophones, according to phonetic environment, which mostly lie between the cardinal vowels [e] and [ɛ], some realizations being closer to one and some to the other. It is therefore permissible to choose either symbol as the one to represent the phoneme.

In other cases the differences between competing transcriptions result from alternative ways of representing the phonological contrast between sounds. In English, for example, the contrast between the words *bead* and *bid* has phonetic correlates in both vowel quality and vowel duration. A phonemic representation which explicitly notes this might use the symbols /iː/ and /ɪ/, where the difference in letter shape reflects the difference in vowel quality, and the length mark on the first letter reflects the difference in duration. But it is equally possible unambiguously to represent these phonemes as /iː/ and /i/ (where the phonemic symbol only explicitly shows the length difference), or as /i/ and /ɪ/ (where only quality is shown explicitly). All three pairs of symbols are in accord with the principles of the IPA (as long as the principle chosen for this pair of vowels is applied consistently throughout the vowels of the language).

Other differences may stem not from alternative representations of what is essentially the same phonemic analysis, as above, but from alternative phonemic analyses. For

instance, English long vowels and diphthongs are often analyzed as unitary phonemes such as /iː/ (as in *heed*) and /aʊ/ (as in *how*) (The fact that the phoneme symbol is made up of two letters or a letter and diacritic does not affect the status of a sound as a single phoneme in the analysis.) It is also possible to analyze them as combinations of a short vowel and an approximant: /i/ + /j/ and /a/ + /w/. Again, the representation resulting from such an analysis is perfectly in keeping with the principles behind the IPA. The important point is that the IPA does not provide a phonological analysis for a particular language, let alone a single 'correct' transcription, but rather the resources to express any analysis so that it is widely understood.

7 Working with the IPA

There are a number of practical issues that may arise when using the IPA. Some of these involve problems of how to refer to symbols. In what follows, reference will be made to Appendix 3. This contains a comprehensive listing of symbols used in phonetics, including those of the IPA, but also many which are not recommended by the International Phonetic Association but which may be encountered. The listing indicates which symbols are not IPA usage, or which were once recommended but are no longer recommended. The listing was produced by the IPA Working Group on Computer Coding, set up at the 1989 International Phonetic Association Congress in Kiel.

7.1 Symbol names

It is often useful to be able to refer to symbols by an agreed name. If it is a question of substituting [ɐ] by [ɒ], it is easier to say 'not "Turned A" but "Turned Cursive A"' than to attempt a verbal description of the relevant symbols. Although the International Phonetic Association has never officially approved a set of names, many symbols have informally one or more names, and a greater degree of consensus has arisen as the result of the use of names in Pullum and Ladusaw (1986)[1]. Appendix 3 therefore includes with each symbol a convenient and systematic name, most of which are those used by Pullum and Ladusaw.

7.2 Using the IPA in handwriting

There are cursive forms of IPA symbols, but it is doubtful if these are much in use today. They may have been of greater use when transcription by hand was the only way of recording speech, and so speed was essential. The cursive forms are harder for most people to decipher, and it is preferable to use handwritten versions which closely copy the printed form of the symbols.

7.3 Using the IPA in print

Printers should normally have a font including IPA symbols. Even if they do, however, there will be a danger of superficially similar symbols being mixed up (for instance [ɵ] with [θ] or [ɣ] with [ɤ]). Some publishers have tables in which unusual symbols can be identified by index numbers and letters, but practice is variable. It should therefore be noted that the Chart in Appendix 3 provides for each symbol a unique identifying number, its 'IPA Number'. It may therefore be helpful to identify symbols

[1] Pullum, G.K. & Ladusaw, W.A. (1986) *Phonetic Symbol Guide.* Chicago: University of Chicago Press.

which might prove difficult by that number, and to supply to printers and publishers a copy of the table.

7.4 Using the IPA on computers

Character sets including most or all of the IPA are available for several computing environments. Most straightforwardly a number of commercial and free fonts are available for Macintosh and Microsoft Windows. The situation in other computing environments may be less straightforward. One problem for those devising IPA character sets which has hindered the interchangeability of data containing phonetic symbols was the lack of an agreed standard coding for the symbols. The International Phonetic Association, through its Working Group on Computer Coding, has worked with the International Standards Organization in its project to set up a universal character set (UCS) for all alphabets. An agreed set of UCS 16-bit codes is included in the list in Appendix 3.

8 Going beyond the IPA

As noted in section 2, the descriptive resources of the IPA were developed principally for the linguistically relevant aspects of speech. This was because the whole tradition of phonetic description was concerned with the properties which realize the phonological systems of languages. A phonological system can be seen as the conventions which speakers of a language share about its sounds. Many aspects of individual utterances such as personal voice quality, emotive modifications of speech, accidental mis-articulations, disfluencies, and speech pathologies, are not relevant to the phonological system, and so phoneticians have tended to ignore such aspects when working on the phonology of a language. The IPA reflects this orientation, being, in essence, a system for describing the linguistic-phonetic properties of error-free utterances not specific to a particular individual. There are, however, many circumstances in which it is essential to be able to transcribe other properties of speech.

One important set of such properties constitutes a conventionalized system of communication beyond the verbal component of speech, and which is often referred to as paralanguage. This includes the use of phenomena such as voice quality, pitch range, and rate of utterance variously to convey aspects of the speaker's emotional state and attitude to other conversational participants, to indicate the status (e.g. confidentiality) of the information being conveyed, and to regulate the course of a conversation by encouraging or discouraging others from speaking. Researchers involved in the analysis of spoken interaction, for instance, clearly need resources for the description of such speech phenomena.

In other situations the phonetic properties of interest may be those which realize the phonological system, but in speakers who for one reason or another do not achieve normal realizations of the system. Most obviously, clinical practice and research in the field of speech pathology require a system of phonetic notation which will cope with sounds and combinations of sounds which lie outside the usual range. Research, too, on children's utterances during language acquisition, which contain many sounds that do not occur systematically in the languages of the world, also requires notational devices outside those provided by the IPA.

Researchers in these fields have, of course, developed their own notational devices as required. But clearly it would be preferable to have a widely agreed standard set of conventions for these additional applications, comparable to the standard provided by the IPA. With this aim the International Phonetic Association's Clinical Phonetics and Linguistics Group has proposed a set of 'Extensions to the IPA' for transcribing non-linguistic speech events, and other aspects of speech such as deviant or pathological speech. These are listed and explained in Appendix 2.

9 Some problematic issues

9.1 Segmentation

In making an impressionistic transcription of a language whose phonological system is not known, uncertainties over the division of an utterance into segments may arise. Some articulatory sequences produce a speech signal which different languages may interpret as made up of a different number of segments. This is sometimes the case, for instance, when secondary articulations are added to primary articulations. An articulatory sequence such as would be represented by [lɔ] is relatively unproblematic. Here the syllable begins with a lateral, which involves a closure by the tongue tip or blade against the alveolar ridge, with the tongue body left free to anticipate the position required for the following vowel. The acoustic signal, as shown in the left half of the spectrogram in Figure 5, clearly falls into two distinct parts corresponding to the lateral and the vowel. But if the tongue body is high during the lateral, there will be an [i]-like transition or palatal approximant between the consonant and the vowel – see the right half of Figure 5. Given only the phonetic event, it is not clear whether to transcribe this sequence as three segments ([ljɔ]) or two segments ([lʲɔ], where the superscript 'j' indicates a modification of the lateral by palatalization, and not an additional segment).

There may be some evidence in the phonetic signal to help resolve the issue. For instance, if there is a noticeable [i]-like transition into the lateral from a preceding vowel as well as into the following vowel, it is more likely that the high tongue position is associated with the lateral. But ultimately the answer will lie in the phonological patterns of the language. If the language has sequences like [jɔ], where the palatal approximant appears independently of another consonant articulation, this points in the direction of treating it as a separate segment elsewhere (e.g. [ljɔ]). Pointing in the other direction would be the discovery that the language contrasted consonants extensively in terms of secondary articulation, even where an approximant would be unlikely to occur, for instance after a word-final voiceless fricative. Sequences such as [asj] and [asw] are not likely, since approximants normally occur adjacent to vowel nuclei, and such a contrast would normally be attributed to the consonant and phonemicized as /asʲ/ and /asʷ/, indicating contrastive secondary articulations of palatalization and labial-velarization. With this knowledge, then, [lʲɔ] would be seen as the segmentation more appropriate to the phonological patterning of the language.

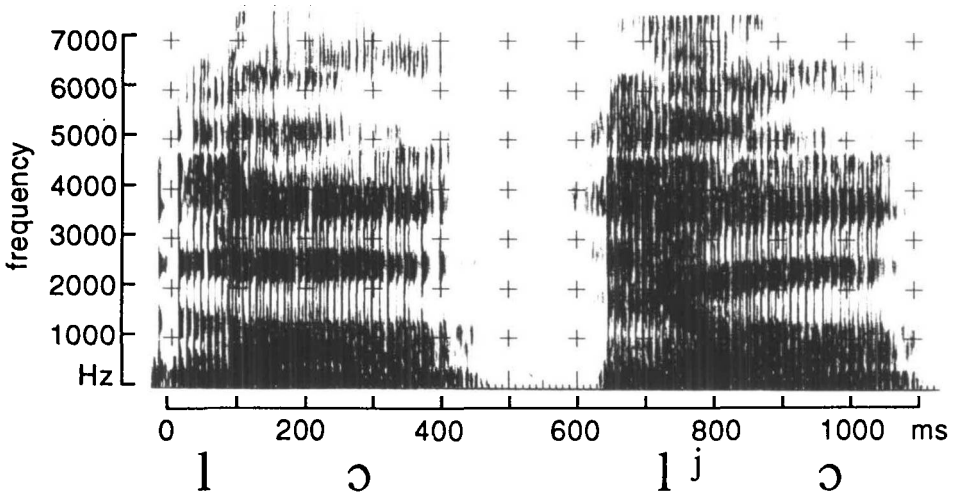


Figure 5. Spectrogram of [lɔ] and an utterance whose segmentation is ambiguously [lʲ]+[ɔ] or [l]+[j]+[ɔ].

The point of considering such examples is to underline the fact that the segmentation which phonetic description requires is not always transparently available in the phonetic event, and impressionistic transcriptions may have to contain unresolved ambiguities until sufficient is known about the structure of the particular language. Moreover, such uncertainties of segmentation will often form the basis of alternative proposals for phonemic interpretation.

9.2 Aligning transcriptions and speech

Even where the structure of the language is known, the alignment of a phonetic transcription with records of the physical speech event will sometimes be problematic. This is because the effects of a particular segment overlap with those of others, or, to put it another way, because the changes in the various parameters which make up speech (voicing, nasality, and so on) are neither instantaneous nor aligned simultaneously. The more closely the physical speech event is observed, the greater the tension between the segmentation derivable from the phonological structure of a language and that suggested by the structure of the physical signal.

The English word *sleeting*, for instance, is phonemically /slitɪŋ/, a sequence of six phonemes. The acoustic signal of this word as spoken by a speaker of a variety of English without voicing of word-medial /t/ is represented by the spectrogram in Figure 6. This shows considerably more than six identifiably different successive aspects. A possible acoustic segmentation is indicated by the vertical lines drawn below the spectro-

gram. A narrow transcription can to an extent capture this acoustic segmentation, and might be suitable for some applications, for instance in speech technology research, where a close annotation of the acoustic signal is required. The representation [s̥l̥lit$^{\text{sh}}$ɪ̃ŋ] given below the spectrogram reflects the fact that the voicelessness of the [s] persists into the first part of the lateral articulation, so there is no single acoustic phase corresponding to the 'voiced lateral approximant' phoneme; that the [t] is released first with a phase of affrication ([s]-like friction locatable to the alveolar region) and then aspiration (non-localizable [h]-like friction); and that the nasality of the final nasal stop is anticipated in the preceding vowel. In the case of the nasality, the narrow transcription captures the way in which cues to a particular phoneme are distributed beyond what would normally be considered the boundaries of the sound; the nasality on the vowel is an early indication that a nasal consonant is imminent. But other instances of this distribution of cues to a sound cannot be captured in an IPA transcription; for example the changing acoustic pattern corresponding to [ɪ̃] is caused by the movement of the tongue body, during the vowel, from the alveolar [t] towards the velar closure of [ŋ]. This changing pattern is an important cue to the place of articulation of both consonants.

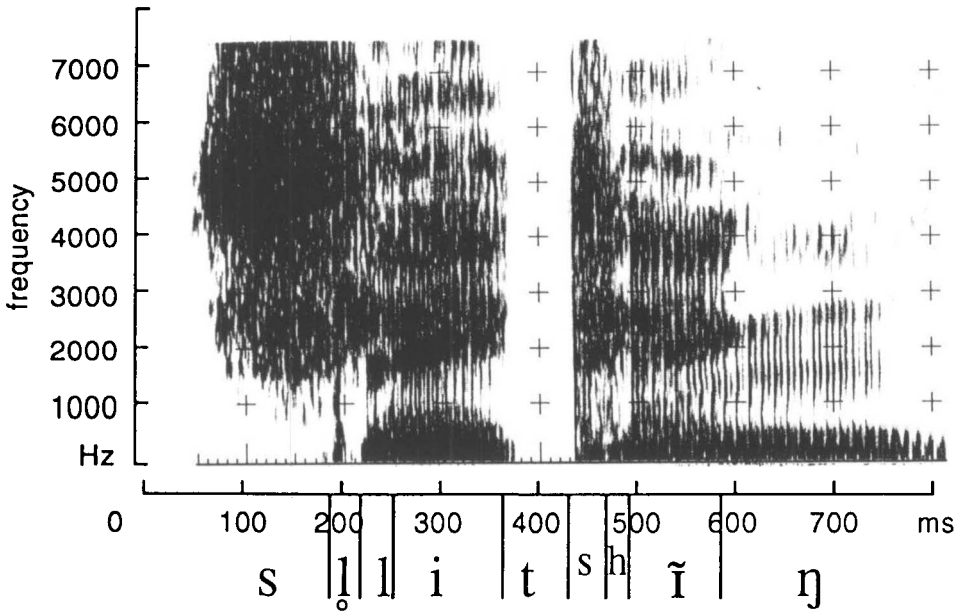


Figure 6. Spectrogram of the word *sleeting*, illustrating the complex relationship between acoustic patterns and phonemic segmentation.

Problems of segmentation and alignment provide a challenge to one of the theoretical assumptions behind the IPA mentioned at the start of section 2, namely that 'speech can be represented partly as a sequence of discrete sounds or "segments"'. The word 'partly' acknowledges the fact that (section 2) 'in addition to the segments, a number of "suprasegmental" aspects of speech, such as stress and tone, need to be represented independently of the segments.' But it turns out that even the 'segmental' aspects of

speech can prove harder to allocate unambiguously to a sequence of discrete segments than might be anticipated. This does not mean that the segmental assumption should be rejected. It is the foundation of phonetic description, and hence of the IPA. What it does mean is that users of the IPA should be aware that the analysis of speech in terms of segments does involve an analytic assumption, and that tensions between the analysis and the data will arise from time to time.

9.3 Transcribing the speaker or the hearer

The relation between a sequence of words and its phonetic realization, far from being unique, is highly variable. A speaker may choose to pronounce carefully, that is with a high degree of 'phonetic explicitness', or to take short cuts. Articulatory short cuts are sometimes known as phonetic reduction. There are tendencies, by no means absolute, for more phonetic reduction to happen the faster someone speaks and the more predictable the content of the speech is.

Many of the differences between explicit and reduced forms can be captured in IPA transcription. For instance a careful utterance of the word *educated* in Standard Southern British English might be transcribed narrowly as [ɛdʲʊkʰeɪtɪd̥], and a phonetically reduced utterance as [ɛd̠zəχeɾəd̥], where unstressed vowels are mid-centralized, the first [d] and the following palatal are assimilated to the alveolo-palatal place of articulation, and the velar and alveolar voiceless plosives of the careful form are instead fricatives.

In other cases the transcriber is faced with theoretically problematic forms as a result of reduction. For instance in the phrase *mad cow*, a careful utterance of which would be [mædˈkʰaʊ], the alveolar at the end of *mad* is susceptible in less careful pronunciation to assimilation to the place of articulation of the following velar. Traditionally, this might be transcribed as [mægˈkʰaʊ], indicating complete loss of the alveolar. Instrumental records of articulation, however, show that sometimes in forms where the alveolar sound cannot be heard the speaker is nonetheless making a reduced tongue movement towards the alveolar ridge. There is then a discrepancy for such an utterance between the transcription which is 'right' for the hearer ([mægˈkʰaʊ] as above) and one which would better reflect the speaker's behavior – perhaps [mæd̠ˈkʰaʊ]. Such a discrepancy violates an assumption implicit in phonetic description, namely that the form to be transcribed is common to speaker and hearer.

10 The IPA and phonological theory

The IPA is intended as a commonly agreed tool for analyzing and representing the phonetic properties of any language. Often, such phonetic analysis will be done in tandem with phonological analysis, that is, the discovery of ways in which sounds pattern in a language and interact with other levels of linguistic structure, particularly morphology (word-building). Views on how best to carry out phonological analysis are constantly evolving as new theories and their associated representational devices are developed.

Although it might be thought ideal if the IPA provided a means of representing phonetic facts independent of theoretical premises, it is inevitable that any means of representation which goes beyond simple replication (as by a tape recorder) must be

shaped by hypotheses about the object being analyzed. Historically, the IPA has its roots in a tradition of phonology in which the notion of the phoneme, as a contrastive sound unit, and allophones, as its variant phonetic realizations, are primary; and in which utterances are seen as the concatenation of the realizations of phonemes. The use of an alphabetic notation underlines the conceptualization of speech as a sequence of sounds.

That conceptualization was shown in the previous section to be sometimes at odds with the physical speech event. It has also been departed from several times in the phonological theories of the last hundred years. Distinctive Feature Theory stressed the importance not of the 'sound' or 'segment', but of the phonetic properties which co-occur in different combinations in sounds. Autosegmental Phonology, and before it Firthian Prosodic Analysis, broke free from the 'slicing' of speech into a single linear sequence of phoneme-sized slots, and allowed some phonetic properties to have larger domains (such as the syllable or word) where this seems in accord with the patterns of a language. Other developments have emphasized the importance of structures such as the mora, the syllable, the foot, and the phonological word in the organization of the phonetic properties of utterances.

These developments in theoretical phonology have had relatively little effect on the IPA. Distinctive Feature Theory has been indirectly acknowledged in the 1989 reformulation of the 'Principles' of the IPA (see Appendix 1); Principle 2 now includes the following:

> 'The representation of [...] sounds uses a set of phonetic categories which describe how each sound is made. These categories define a number of natural classes of sounds that operate in phonological rules and historical sound changes. The symbols of the IPA are shorthand ways of indicating certain intersections of these categories. Thus [p] is a shorthand way of designating the intersection of the categories voiceless, bilabial, and plosive; [m] is the intersection of the categories voiced, bilabial, and nasal; and so on.'

But there has been no loosening of the segmental 'slicing' of a traditional phonemic view. The IPA Chart, in its fundamental conception, remains much as it has been for over a century. Only in the case of those properties explicitly recognized as suprasegmental and in the 'Extensions to the IPA' (Appendix 2) are devices provided for properties extending over domains larger than a segment.

The conservatism inherent in the IPA tradition has advantages. Phonemic analysis is still the most widely understood and practiced form of phonological analysis, at least outside the ranks of theoretical phonologists, and its principles are fairly accessible to all those familiar with alphabetic writing systems. This favors a system of general phonetic description such as the IPA which is closely compatible with a phonemic view. Secondly, the inertia of the IPA protects it from the shorter-lived of the winds of phonological change, and provides an element of continuity which is particularly important to those who use the IPA as a tool for practical purposes. Nonetheless the IPA should not be regarded as immutable, even in its fundamental assumptions, and there needs to be a continuing appraisal of their appropriateness.