

Cambridge University Press

978-0-521-63751-0 - Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet

Corporate Author International Phonetic Association

Excerpt

[More information](#)

PART 1

Introduction to the IPA

Cambridge University Press

978-0-521-63751-0 - Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet

Corporate Author International Phonetic Association

Excerpt

[More information](#)

1 What is the International Phonetic Alphabet?

The aim of the International Phonetic Association is to promote the study of the science of phonetics and the various practical applications of that science. For both these it is desirable 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 set of symbols 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, and here the Alphabet will generally be abbreviated to 'the IPA'. 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 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 generally agreed set of symbols for designating sounds unambiguously, and the IPA aims to fulfil this role. 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 and corrections approved by the Council of the Association. Despite these and earlier changes, the Alphabet today shows striking continuity with the Association's Alphabet as it was at the end of the nineteenth 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

Cambridge University Press

978-0-521-63751-0 - Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet

Corporate Author International Phonetic Association

Excerpt

[More information](#)

4 *Handbook of the IPA*

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’, which is reproduced in its entirety after the foreword, and section by section in appendix 5; readers are encouraged to photocopy and enlarge the Chart for ease of reference. The structure of the Chart reflects the assumptions above. The following subsections provide a 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.

In introducing the IPA, it is necessary to refer to examples from languages. As far as possible, languages are used with which many readers may be acquainted, but of course this is not possible for many sounds. Variation also provides a problem: all languages have different accents and other varieties of pronunciation. When a sound is exemplified by a word in a particular language, this should be taken to mean that the sound can often be heard in that word, not that it will *always* occur in pronunciations of the word.

In the case of English, reference is made here mainly to two varieties, General American and Standard Southern British. These varieties are widely heard, in the United States and Britain respectively, in formal areas of broadcasting such as newscasts. General American is regarded as a variety which transcends regional divides. Standard Southern British (where ‘Standard’ should not be taken as implying a value judgment of ‘correctness’) is the modern equivalent of what has been called ‘Received Pronunciation’ (‘RP’). It is an accent of the south east of England which operates as a prestige norm there and (to varying degrees) in other parts of the British Isles and beyond. Where necessary, reference will be made specifically to one of these varieties, but normally the term ‘English’ will be used, indicating that the sound occurs in both General American and Standard Southern British. In most cases, of course, the sound will also occur in many other varieties of English.

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* ([ˈrɛfjʊz] meaning ‘rubbish’ and [rɪˈfjuːz] 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 3.

2.2 *Segments*

Observation of the movements of the speech organs reveals that they are in almost

Cambridge University Press

978-0-521-63751-0 - Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet

Corporate Author International Phonetic Association

Excerpt

[More information](#)

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 rapid 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*, the pattern ebbs and flows constantly, and there are no boundaries between successive sounds. Nonetheless the word can be segmented as [wɜːɪ] – that is, as [w] + [ɜː] + [ɪ] + [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, [wɜːɪ] *worry*, [hɜːɪ] *hurry*, [hæɪ] *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.

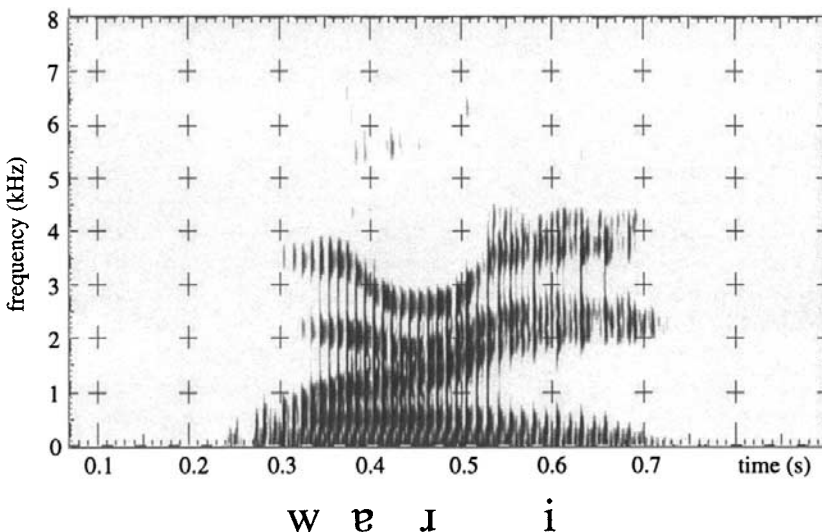


Figure 1 Spectrogram of the word *worry*, spoken in a Southern British accent.

Cambridge University Press

978-0-521-63751-0 - Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet

Corporate Author International Phonetic Association

Excerpt

[More information](#)

6 *Handbook of the IPA*

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 front 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 between 0.4 and 0.5 s. 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* ([bənænə] or [bənənə]), 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 centre, 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 centre. So in a common pronunciation of the English word *button* as [bʌtʌn] there are two syllables, but the nucleus of the second is a consonant, as judged from the way it is produced. Conversely in the word [jet] *yet*, the first sound, if prolonged, is very similar to the vowel of [hid] *heed*, and does not involve a narrowing extreme enough to produce friction. However because [j] plays the same role in the syllable as sounds which are by definition consonants (e.g. [b] in [bet] *bet*), it is often included in the class of consonants and described accordingly.

On the IPA Chart, there are separate sections for vowels and for consonants, reflecting

Cambridge University Press

978-0-521-63751-0 - Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet

Corporate Author International Phonetic Association

Excerpt

[More information](#)

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 vocal tract, the mid-sagittal plane as it is known, 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 labelled. 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 or velum). Other places of articulation are exemplified in section 3.

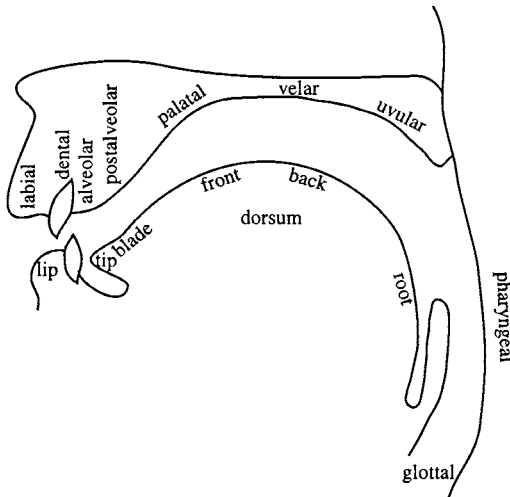


Figure 2 Mid-sagittal section of the vocal tract with labels for place of articulation

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 or vocal folds) at the right. The terms ‘bilabial’ and ‘labiodental’ indicate that the consonant is made by the lower lip against the upper lip and the upper front teeth respectively;

Cambridge University Press

978-0-521-63751-0 - Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet

Corporate Author International Phonetic Association

Excerpt

[More information](#)8 *Handbook of the IPA*

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. Usually alveolar [ɹ] shares some degree of this curling back of the tongue tip, which distinguishes it from other alveolars. 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 [ɲ n ɳ] respectively.

The rows of the consonant table, labelled 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 [v] 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 (place-name element)’) 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. A symbol on the left of a cell in the table is for a voiceless consonant, e.g. [p] and [tʰ], and one on the right is for a voiced consonant, e.g. [b] (the voiced counterpart of [p]) and [m]. 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 in which the vocal cords are, respectively, vibrating during the plosive

Cambridge University Press

978-0-521-63751-0 - Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet

Corporate Author International Phonetic Association

Excerpt

[More information](#)

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 [ɡ].) Or conversely, ‘what sort of a sound is [j]?’ (The answer is one which is voiced, and in which frication can be heard resulting from a narrowing between the tongue front and the hard palate.)

Not all cells or halves of cells in the consonant table contain symbols. The gaps are of three kinds. Shaded cells occur where the intersection of a manner and a place of articulation define a sound which is thought not to be possible, either by definition (a nasal requires an oral occlusion combined with lowering of the velum, and so a pharyngeal or glottal nasal is ruled out), or because the sound is impossible or too difficult to produce, such as a velar trill or a bilabial lateral fricative. Unless phoneticians are mistaken in their view of the latter category of sound, 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 [ɭ], 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 symbol. 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 symbol [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

Cambridge University Press

978-0-521-63751-0 - Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet

Corporate Author International Phonetic Association

Excerpt

[More information](#)10 *Handbook of the IPA*

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. [pʰ], [sʰ]. If instead the air between the glottis and a closure further forward is expanded, reducing its pressure, air will flow into 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 [ɘ] 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: [kɘ̤ ɡɘ̤ ɴɘ̤], [qɘ̤].

2.6 Vowels

Vowels are sounds which occur at syllable centres, and which, because they involve a less extreme 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, middle right). 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 labelled [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’.

Cambridge University Press

978-0-521-63751-0 - Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet

Corporate Author International Phonetic Association

Excerpt

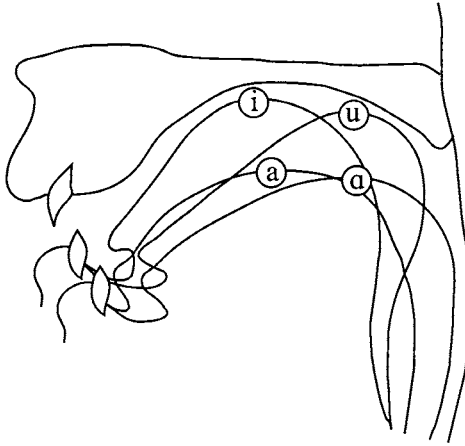
[More information](#)

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 labelled [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 qualities 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 circles 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 differences between each vowel and the next in the series are auditorily equal; 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