to our motor, articulatory, habits. If this feedback control is disturbed, e.g. by the imposition of an artificial delay upon our reception of our own speech, disturbance in the production of our utterance is likely to result. Those who are born deaf or who become deaf before the acquisition of speech habits are rarely able to learn normal speech completely; similarly, a severe hearing loss later in life is likely to lead eventually to a deterioration of speech, although not down to the same level as those born deaf.

Notes

1 Fant (1956).
2 Hollien & Shipp (1972); Russell *et al.* (1995).
3 Denes & Pinson (1993: 101ff).
4 For example, *Praat* (Boersma & Weenink, 2011).
5 Abberton & Fourcin (1984).

Chapter 4

The description and classification of speech sounds

4.1 Phonetic description

We have considered briefly both the mechanism which produces speech sounds and also some of the acoustic and auditory characteristics of the sounds themselves. We have seen that a speech sound has at least three stages available for investigations—the production, transmission and reception stages. The most convenient and brief descriptive classification of speech sounds relies either on articulatory criteria or on auditory judgements, or on a combination of both. Those sounds which are commonly known as consonants are most easily described mainly in terms of their articulation, whereas the description of vowels depends more on auditory impressions.

4.2 Vowel and consonant

Two types of meaning are associated with the terms VOWEL and CONSONANT. In one type of definition consonants are those segments which, in a particular language, occur at the edges of syllables, while vowels are those which occur at the centre of syllables. So, in *red*, *wed*, *dead*, *lead*, *said*, the sounds represented by <r,w,d,l,s> are consonants, while in *beat*, *bit*, *bet*, *but*, *bought*, the sounds represented by <ea,i,e,u,ou> are vowels. This reference to the functioning of sounds in syllables in a particular language is a phonological definition. But once any attempt is made to define what sorts of sounds generally occur in these different syllable-positions, then we are moving to a phonetic definition. This type of definition might define vowels as median (air must escape over the middle of the tongue, thus excluding the lateral [l]), oral (air must escape through the mouth, thus excluding nasals like [n]), frictionless (thus excluding fricatives like [s]), and continuant (thus excluding plosives like [p]); all sounds excluded from this definition would be consonants. But difficulties arise in English with this definition (and with others of this sort) because English /j,w,r/, which are consonants phonologically (functioning at the edges of syllables), are vowels phonetically. Because of this these sounds are often called semi-vowels. The reverse type of difficulty is encountered in words like *sudden* and *little* where

the final consonants /n/ and /l/ form syllables on their own and hence must be the centre of such syllables even though they are phonetically consonants, and even though /n/ and /l/ more frequently occur at the edges of syllables, as in *net* and *let*. When occurring in words like *sudden* and *little*, nasals and laterals are called syllabic consonants.

In this chapter we will be describing and classifying speech sounds phonetically (in the next chapter we return to the phonological definitions). We shall find that consonants can be voiced or voiceless and are most easily described wholly in articulatory terms, since we can generally feel the contacts and movements involved. Vowels, on the other hand, are voiced and, depending as they do on subtle adjustments of the body of the tongue, are more easily described in terms of auditory relationships.

4.3 Consonants

For consonantal articulations, a description must provide answers to the following questions:

(1) Is the airstream set in motion by the lungs or by some other means? (pulmonic or non-pulmonic)
(2) Is the airstream forced outwards or sucked inwards? (egressive or ingressive)
(3) Do the vocal cords vibrate or not? (voiced or voiceless)
(4) Is the soft palate raised, directing the airstream wholly through the mouth, or lowered, allowing the passage of air through the nose? (oral, or nasal or nasalised)
(5) At what point or points and between what organs does closure or narrowing take place? (place of articulation)
(6) What is the type of closure or narrowing at the point of articulation? (manner of articulation)

In the case of the sound [z], occurring medially in the word *easy*, the following answers would be given:

(1) pulmonic
(2) egressive
(3) voiced
(4) oral
(5) tongue blade–alveolar ridge
(6) fricative.

These answers provide a reasonably full but concise phonetic label for the sound; a more detailed description would include additional information concerning, for instance, the shape of the remainder of the tongue, the relative position of the jaws and the lip position.

4.3.1 Egressive pulmonic consonants

Most speech sounds are made with egressive lung air. Virtually all English sounds are so made, the exception being [p,t,k], which in some dialects sometimes become ejectives (see §4.3.9).

4.3.2 Voicing

At any place of articulation, a consonantal articulation may involve the vibration of the vocal cords, i.e. may be voiceless or voiced.

4.3.3 Place of articulation

The chief points of articulation are the following (reference is made to the videos on the companion website where the articulations can be seen in spoken phrases):[1]

BILABIAL—The two lips are the primary articulators, e.g. [p,b,m]. **(See videos 1.12, 5.18, 6.0, 6.14, 14.6.)**

LABIODENTAL—The lower lip articulates with the upper teeth, e.g. [f,v]. **(See videos 3.13, 4.14, 10.14, 13.14.)**

DENTAL—The tongue tip and rims articulate with the upper teeth, e.g. [θ,ð], as in *think* and *then*. **(See videos 6.24, 10.21, 12.16.)**

ALVEOLAR—Either the blade, or tip and blade, of the tongue articulates with the alveolar ridge, e.g. English [t,d,l,n,s,z]. **(See videos 3.1, 3.15, 3.23, 5.16, 9.25, 11.10.)**

POST-ALVEOLAR—The tip of the tongue articulates with the rear part of the alveolar ridge, e.g. [ɹ] as at the beginning of English *red*. **(See videos 2.4, 4.2.)**

RETROFLEX—The tip of the tongue is curled back to articulate with the part of the hard palate immediately behind the alveolar ridge, e.g. [ɻ] such as is found in South-West British and some American English pronunciations of *red*.

PALATO-ALVEOLAR—Either the blade, or the tip and blade, of the tongue articulates with the alveolar ridge and there is at the same time a raising of the front of the tongue towards the hard palate, e.g. [ʃ,ʒ,ʃ̺,dʒ] as in English *ship*, *measure*, *beach*, *edge*.[2] **(See videos 1.1, 8.15, 9.1, 11.23.)**

PALATAL—The front of the tongue articulates with the hard palate, e.g. [j] or [ç] as in *queue* [kj̊uː] or [kçuː] or a very advanced type of [k,g] = [c,ɟ], as in French *quitter* or *guide*. **(See videos 5.5, 14.14.)**

VELAR—The back of the tongue articulates with the soft palate, e.g. [k,g,ŋ], the last as in *sing*. **(See videos 4.21, 7.1, 10.4, 11.1, 14.11, 15.22.)**

UVULAR—The back of the tongue articulates with the uvula, e.g. [ʁ] as in French *rouge*.

GLOTTAL—An obstruction, or a narrowing causing friction but not vibration, between the vocal cords, e.g. English [h] as in *ham*.

In the case of some consonantal sounds, there may be a secondary place of articulation in addition to the primary. Thus, in the so-called 'dark' [ɬ], as at the end of *bull* **(see video 1.20)**, in addition to the partial alveolar contact, there is an essential raising of the back of the tongue towards the velum (velarisation); or, again, some post-alveolar articulations of [ɹ] are accompanied by slight lip-rounding (labialisation). The place of PRIMARY ARTICULATION is that of the greatest stricture, that which gives rise to the greatest obstruction to the airflow. The SECONDARY ARTICULATION exhibits a stricture of lesser rank. Where there are two co-extensive strictures of equal rank an example of DOUBLE ARTICULATION results.

4.3.4 Manner of articulation

The obstruction made by the organs may be total, intermittent, partial, or may merely constitute a narrowing sufficient to cause friction. The chief types of articulation, in decreasing degrees of closure, are as follows:

(1) *Complete Closure*

PLOSIVE—A complete closure at some point in the vocal tract, behind which the air pressure builds up and can be released explosively, e.g. [p,b,t,d,k,g,ʔ] as in *pay*, *boot*, *tea*, *down*, *car*, *gate* and in a Cockney pronunciation of *water* as [wɔʊʔə].

AFFRICATE—A complete closure at some point in the mouth, behind which the air pressure builds up but the separation of the organs is slow compared with that of a plosive, so that friction is a characteristic of the second part of the sound, e.g. [ʧ,ʤ] in *cheese, joke.*

NASAL—A complete closure at some point in the mouth but, the soft palate being lowered, the air escapes through the nose, e.g. [m,n,ŋ] as in *modern, name, sing*. These sounds are continuants and, in their (most usual) voiced form, have no noise component; they are, to this extent, vowel-like.

(2) *Intermittent Closure*

TRILL (OR ROLL)—A series of rapid intermittent closures made by a flexible organ on a firmer surface, e.g. [r], where the tongue tip trills against the alveolar ridge as in Spanish *perro*, or [ʀ] where the uvula trills against the back of the tongue, as in a stage pronunciation of French *rouge*.

TAP—A single tap made by a flexible organ on a firmer surface, e.g. [ɾ] where the tongue tip taps once against the teeth ridge, as in many Scottish pronunciations of English /r/.

(3) *Partial Closure*

LATERAL—A partial (but firm) closure is made at some point in the mouth, the airstream being allowed to escape on one or both sides of the contact. These sounds may be continuant and frictionless and therefore vowel-like (i.e. approximants in (5) below), as in [l,ɬ] as pronounced in the south of England *little* [lɪtɬ] or they may be accompanied by a little friction [l̝] as in *fling* or by considerable friction [ɬ] as in *please*.

(4) *Narrowing*

FRICATIVE—Two organs approximate to such an extent that the airstream passes between them with friction, e.g. [ɸ,β,f,v,ʍ,θ,ð,s,z,ʃ,ʒ,ç,x,h]. In the bilabial region, a distinction is to be made between those purely bilabial such as [ɸ,β] where the friction occurs between spread lips, and a labial-velar sound like [ʍ] where the friction occurs between rounded lips and is accompanied by a characteristic modification of the mouth cavity brought about by the raising of the back of the tongue towards the velum. [ç] may occur at the beginning of *huge*, [x] and [ʍ] in Scottish pronunciations of *loch* and *which*, and [β] in Spanish *haber*.

(5) *Narrowing without Friction*

APPROXIMANT (or FRICTIONLESS CONTINUANT)—A narrowing is made in the mouth but the narrowing is not quite sufficient enough to cause friction. In being frictionless and continuant, approximants are vowel-like; however, they function phonologically as consonants, i.e. they appear at the edges of syllables. They also differ phonetically from such sounds functioning as vowels in either of two ways. First, the articulation may not involve the body of the tongue, e.g. post-alveolar [ɹ] and labiodental [ʋ], the former the usual pronunciation in GB at the beginning of *red*, the latter a regional or idiosyncratic pronunciation of the same sound, as well as a regular consonant in some languages, e.g. Hindi. Second, where they do involve the body of the tongue, the articulations represent only brief glides to a following vowel: thus [j] in *yet* is a glide starting from the [i] region and [w] in *wet* is a glide starting from the [u] region.

4.3.5 Obstruents and sonorants

It is sometimes found useful to classify categories of sounds according to their noise component. Those in whose production the constriction impeding the airflow through the vocal tract is sufficient to cause noise are known as OBSTRUENTS. This category comprises plosives, fricatives and affricates. SONORANTS are those voiced sounds in which there is no noise component i.e. voiced nasals, approximants and vowels.

4.3.6 Fortis and lenis

A voiceless/voiced pair such as English /s,z/ are distinguished not only by the presence or absence of voice but also by the degree of breath and muscular effort involved in their articulation. Those English consonants which are usually voiced tend to be articulated with relatively weak energy (they are LENIS), whereas those which are always voiceless are relatively strong (they are FORTIS). Indeed, we shall see that in certain situations, the so-called voiced consonants may have very little voicing, so that the energy of articulation becomes a significant factor in distinguishing the voiced and voiceless series.

4.3.7 *Classification of consonants*

The chart of the International Phonetic Alphabet (IPA)[3] (see Table 1) shows manner of articulation on the vertical axis; place of articulation on the horizontal axis; and a pairing within each box thus created shows voiceless consonants on the left and voiced consonants on the right.

4.3.8 *Ingressive pulmonic consonants*

Consonants of this type, made as we are breathing in, sometimes occur in languages as variants of their egressive pulmonic equivalents. We may use such sounds when we are out of breath, but have not got time to pause, either because the need for communication is pressing, or because we do not wish someone else to have a chance to speak. The use of such an ingressive pulmonic airstream is, however, variable between languages and is not especially common in English. Individual sounds made on an ingressive pulmonic airstream may occur as speech defects.[4] Some sounds may also occur extralinguistically, so in English a common way of expressing surprise or pain involves the energetic inspiration of air accompanied by bilabial or labiodental friction.

4.3.9 *Egressive glottalic consonants*

In the production of these sounds, known as EJECTIVES, the glottis is closed, so that lung air is contained beneath it. A closure or narrowing is made at some point above the glottis (the soft palate being raised) and the air between this point and the glottis is compressed by a general muscular constriction of the chamber and a raising of the larynx. Thus plosive ejectives [p',t',k'] may be made by compressing the air in this way. But it is not only plosives which may be ejective; affricates and fricatives have this type of compression in a number of languages, e.g. [ts',tl',s',x']. Since the glottis is tightly closed, it follows that this type of articulation applies only to voiceless sounds.

[p',t',k'] occur sometimes in final positions in some dialects of English (e.g. in south-east Lancashire, where *stop* may be [stɒp']). This is not to be confused with other variants of final [p,t,k], which may be reinforced or replaced by a glottal stop, e.g. *stop* [stɒʔp] as in GB, or [stɒʔ] as in popular London speech.

4.3.10 *Ingressive glottalic consonants*

For these sounds a complete closure is made in the mouth but, instead of air pressure from the lungs being compressed behind the closure as in §4.3.9, the almost completely closed larynx is lowered so that the air in the mouth and pharyngeal cavities is rarefied. The result is that outside air is sucked in once the mouth closure is released; at the same time, there is sufficient leakage of lung air through the glottis to produce voice. It will be seen that the resulting sound is made by means of a combined airstream mechanism, namely an egressive

Table 1 The International Phonetic Alphabet.

THE INTERNATIONAL PHONETIC ALPHABET (revised to 2005)

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		ⱱ		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

OTHER SYMBOLS

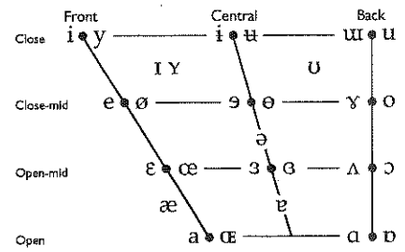
ʍ	Voiceless labial-velar fricative	ɕ ʑ	Alveolo-palatal fricatives
w	Voiced labial-velar approximant	ɺ	Voiced alveolar lateral flap
ɥ	Voiced labial-palatal approximant	ʢ	Simultaneous ʃ and X
ʜ	Voiceless epiglottal fricative		
ʡ	Voiced epiglottal fricative		
ʔ	Epiglottal plosive		

Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary. k͡p t͡s

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 (ɪ̝ = voiced alveolar fricative)
̩	Syllabic	n̩	̞	Lowered	e̞ (β̞ = voiced bilabial approximant)
̯	Non-syllabic	e̯	̘	Advanced Tongue Root	e̘
˞	Rhoticity	ə˞ a˞	̙	Retracted Tongue Root	e̙

VOWELS



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

SUPRASEGMENTALS

ˈ	Primary stress
ˌ	Secondary stress ˌfoʊnəˈtɪʃən
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	
ȅ or ˥	Extra high	ě or ˦˥	Rising
é ˦	High	ê	˥˩	Falling
ē ˧	Mid	ě	˦˥	High rising
è ˨	Low	ě	˨˦	Low rising
ȅ ˩	Extra low	ẽ	˥˦˩	Rising-falling
↓	Downstep	↗	Global rise
↑	Upstep	↘	Global fall

pulmonic airstream in combination with ingressive glottalic air. Such ingressive stops (generally voiced) are known as IMPLOSIVES and occur as bilabial [ɓ], dental or alveolar [ɗ], or velar [ɠ]. Though such sounds occur in a number of languages, they are not usually found in English, although they are sometimes heard in the speech of the deaf and in types of stammering. Much more rarely voiceless implosives may occur in some languages, which means that in these cases the larynx must be completely closed and the ingressive glottalic airstream occurs on its own without the egressive pulmonic one.

4.3.11 Ingressive velaric consonants

Another set of sounds involving an ingressive airstream is produced entirely by means of closures within the mouth cavity; normal breathing through the nose may continue quite independently if the soft palate is lowered and may even produce accompanying nasalisation. Thus, the sound made to indicate irritation or sympathy (often written as 'tut-tut') is articulated by means of a double closure, the back of the tongue against the velum and the tip, blade and sides against the teeth and teeth ridge. The cavity contained within these closures is then enlarged mainly by tongue movement, so that the air is rarefied. The release of the forward closure causes the outer air to be sucked in; the release may be crisp in which case a sound of a plosive type is heard, or relatively slow, in which case an affricated sound is produced. These sounds are known as CLICKS, the one referred to above ('tut-tut') being a dental click [ǀ]. The sound made to encourage horses is a lateral click, i.e. the air is sucked in by releasing one side of the tongue [ǁ]. These clicks and several others occur as significant sounds in a small number of languages in Africa (e.g. Zulu) and paralinguistically in most languages (as in English).

4.4 Vowels

Vowels are normally made with a voiced egressive airstream, without any closure or narrowing such as would result in the noise component characteristic of many consonantal sounds; moreover, the escape of the air is characteristically accomplished in an unimpeded way over the middle line of the tongue. We are now concerned with a glottal tone modified by the action of the upper resonators of the mouth, pharyngeal and nasal cavities. As we have seen (Chapters 2 and 3), the movable organs mainly responsible for shaping these resonators are the soft palate, lips and tongue. A description of vowel-like sounds must, therefore, note:

(1) The position of the soft palate—raised for oral vowels, lowered for nasalised vowels.
(2) The kind of aperture formed by the lips—neutral, spread, close-rounded, or open-rounded.
(3) The part of tongue which is raised and the degree of raising.

Of these three factors, only the second—the lip position—can be easily described by visual or tactile means. Our judgement of the action of the soft palate depends less on our feeling for its position than on our perception of the presence or absence of nasality in the sound produced. The movements of the tongue, which so largely determine the shape of the mouth and pharyngeal cavities, may be so minute that it is impossible to assess them by any simple means; moreover, there being normally no contact of the tongue with the roof of the mouth, no help is given by any tactile sensation. A vowel description will usually, therefore, be based mainly on auditory judgements of sound relationships, together with some articulatory information, especially as regards the position of the lips. In addition, an acoustic description can be given in terms of the disposition of the characteristic formants of the sound (see §3.1).

4.4.1 Difficulties of description

The verbal description of vowel sounds has always presented considerable difficulty. Certain positions and gross movements of the tongue can be felt. We are, for instance, aware that when we pronounce most vowel sounds the tongue tip lies behind the lower teeth; moreover, in comparing two such vowels as /iː/ (*key*) **(see videos 3.6, 11.1)** and /ɑː/ (*car*) **(see videos 10.18, 12.7)** (Fig. 4), we can feel that, in the case of the former, the front of the tongue is the part which is mainly raised, whereas, in the case of the latter, such raising as there is is accomplished by the back part of the tongue. Therefore, it can be stated in articulatory terms that some vowel sounds require the raising of the front of the tongue, while others are articulated with a typical 'hump' at the back; and these statements have been confirmed by means of X-ray photography and by MRI scans. But the actual point and degree of raising is more difficult to judge. It is not, for instance, helpful to say that a certain vowel is articulated with the front part of the tongue raised to within 5 mm of the hard palate. Moreover this may be a statement of fact for one person's pronunciation, but an identical sound may be produced by another speaker with a different relationship between the tongue and palate. It is no more helpful to relate the vowel quality to a value used in

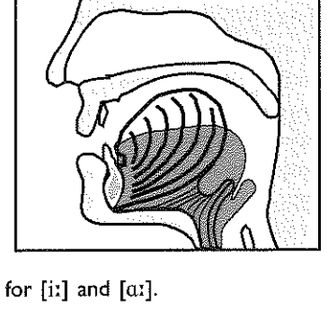


Figure 4 Tongue positions for [iː] and [ɑː].

a particular language, as is still so often done. A statement such as 'a vowel quality similar to that in the English word *cat*' is not very precise, since the vowel in *cat* may have a wide range of values in English. The statement becomes more useful if the accent of English is specified, but even then a number of variant interpretations will always be possible.

4.4.2 *Cardinal Vowels*

It is clear that a finer and more independent system of description is needed, on both the auditory and articulatory levels. The most satisfactory scheme is that devised by Daniel Jones and known as the CARDINAL VOWEL system. The basis of the system is physiological, i.e. the two qualities, upon which all the others were 'hinged', were produced with the tongue in certain easily felt positions: the front of the tongue raised as close as possible to the palate without friction being produced, for the Cardinal Vowel [i]; and the whole of the tongue as low as possible in the mouth, with very slight raising at the extreme back, for the Cardinal Vowel [ɑ]. Starting from the [i] position, the front of the tongue was lowered gradually, the lips remaining spread or neutrally open and the soft palate raised. The lowering of the tongue was halted at three points at which the vowel qualities seemed, from an auditory standpoint, to be equidistant. The tongue positions of these qualities were X-rayed and were indeed found to be fairly equidistant from a spatial point of view. The symbols [e,ɛ,a] were assigned to these vowel values. The same procedure was applied to vowel qualities depending on the height of the back of the tongue: thus the back of the tongue was raised in stages from the [ɑ] position and with the soft palate again raised; additionally the lips were changed progressively from a wide open shape for [ɑ] to a closely rounded one for [u] (reflecting the most usual lip positions for these tongue positions in the world's languages). As with the front of the tongue, three auditorily equidistant points were established from the lowest to the highest position; the corresponding tongue positions were photographed and the spatial relationships again confirmed. These values were given the symbols [ɔ,o,u]. Thus, a scale of eight primary Cardinal Vowels was set up, denoted by the following numbers and symbols: 1, [i]; 2, [e]; 3, [ɛ]; 4, [a]; 5, [ɑ]; 6, [ɔ]; 7, [o]; 8, [u].

The front series [i,e,ɛ,a] and [ɑ] of the back series are pronounced with spread or open lips, whereas the remaining three members of the back series have varying degrees of lip-rounding. These combinations of tongue and lip positions in the primary Cardinal Vowels are the most frequent in languages, i.e. front and open vowels are most commonly unrounded while back vowels other than in the open position are most commonly rounded. A secondary series can be obtained by reversing the lip positions, e.g. lip-rounding applied to the [i] tongue position, or lip-spreading applied to the [u] position. Such a secondary series is denoted by the following numbers and symbols: 9, [y]; 10, [ø]; 11, [œ]; 12, [ɶ]; 13, [ɒ]; 14, [ʌ]; 15, [ɤ]; 16, [ɯ].

This complete series of sixteen Cardinal Vowel values may be divided into two lip shape categories, with corresponding tongue positions:

Unrounded—[i,e,ɛ,a,ɑ,ʌ,ɤ,ɯ].
Rounded—[y,ø,œ,ɶ,ɒ,ɔ,o,u].

Such a scale is useful because (a) the vowel qualities are unrelated to particular values in languages, though many may occur in various languages and (b) the set is recorded, so that reference may always be made to a standard, invariable scale.[5] Thus a vowel quality can be described as being, for instance, similar to that of Cardinal 2 ([e]), or another as being a type halfway between Cardinal 6 ([ɔ]) and Cardinal 7 ([o]), but somewhat centralised. Diacritics are available in the IPA alphabet to show modifications of Cardinal values: subscript diacritics indicate more open, e.g. [o̞], and more close, e.g. [o̝], while a pair of dots over a symbol, e.g. [ö], indicates centralisation. The vowel quality mentioned above, that between C.[ɔ] and C.[o] and centralised, can in this way be symbolised as [ö̝] or [ö̞].

It is, moreover, possible to give a visual representation of these vowel relationships on a chart which is based on the Cardinal Vowel tongue positions. The simplified diagram shown in Fig. 5 is obtained by plotting the highest point of tongue raising for each of the primary Cardinal Vowels and joining the points together. The internal triangle, corresponding to the region of central or [ə]-type vowel sounds, is made by dividing the top line into three approximately equal sections and drawing lines parallel to the two sides, so that they meet near the base of the figure. On such a figure, the sound symbolised by [ö̝] or [ö̞] may have its relationship to the Cardinal scale shown visually (see the black circle on Fig. 5).

It must be understood that this diagram is a highly conventionalised one which shows, above all, quality relationships. Some attempt is, however, made to relate the shape of the figure to actual tongue positions: thus, the range of movement

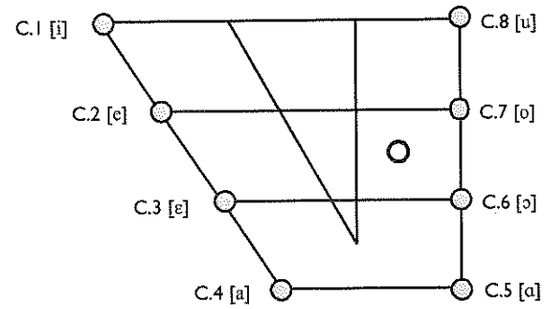


Figure 5 The primary Cardinal Vowels; the area symbolised by [ö̝] or by [ö̞] shown by a circle.

is greater at the top of the figure and the tongue raising of front vowels becomes more retracted as the tongue position lowers. Nevertheless, it has been shown that it is possible to articulate vowel qualities without the exact tongue and lip positions which this diagram seems to postulate as necessary. It is, for instance, possible to produce a sound of the Cardinal 7 ([o]) type without the lip–tongue relationship suggested. But, on the whole, it may be assumed that a certain auditorily identified vowel quality will be produced by an articulation of the kind suggested by the Cardinal Vowel diagram. Moreover, it is a remarkable fact that the auditory judgements as to vowel relationships made by Daniel Jones have been largely supported by acoustic analysis; in fact, charts based on acoustic analyses of GB vowels correspond very well with the traditional Cardinal Vowel figure (see Figs 10 and 11 in Chapter 8).

4.4.3 Nasality

Besides the information concerning lip and tongue positions which the above chart and symbolisation denote, a vowel description must also indicate whether the vowel is purely oral or whether it is nasalised. The sixteen Cardinal Vowels mentioned may all be transformed into their nasalised counterparts if the soft palate is lowered. It is unusual, however, to find such an extensive series of nasalised vowels, since it is unusual (though not unknown) for languages to make such fine, significant distinctions of nasalised qualities as are common in the case of the purely oral values.

4.4.4 Relatively pure vowels vs gliding vowels

It is clearly not possible for the quality of a vowel to remain absolutely constant (or, in other words, for the organs of speech to function for any length of time in an unchanging way). Nevertheless, we may distinguish between those vowels which are relatively pure (or unchanging), such as the vowel in *learn*, and those which have a considerable and deliberate glide, such as the gliding vowel in *line*. The so-called pure vowels are marked on a diagram as a dot, showing the highest point of the tongue, or, better, as a ring, since it is inadvisable to attempt to be over-precise in the matter of these auditory judgements; a gliding vowel sound (or DIPHTHONG) is shown as an arrow, which indicates the quality of the starting-point and the direction in which the quality change is made (corresponding to a movement of the tongue). Fig. 6 shows the way in which the vowels of *learn* **(see video 7.7)** and *line* **(see videos 4.4, 8.3ff)** will be marked.

We can now give a practical and comprehensive description of any vowel sound, partly in articulatory terms, partly in auditory terms. The vowel which we have symbolised in Fig. 5 as [ɵ̟] or [ö] might be described in this way: 'A vowel quality between Cardinal Vowels 6 and 7, but somewhat centralised'. Such a written description will have a meaning in terms of sound for anyone who is familiar with the Cardinal Vowel scale. The position of the lips and the soft

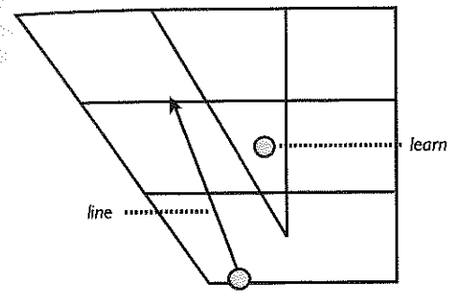


Figure 6 The vowels of *learn* and *line*.

palate is subsumed in this description. There may, of course, be other features of the sound which may be worth mentioning in a full phonetic description, e.g. a breathy or creaky voice quality.

4.4.5 Articulatory classification of vowels

Although precise descriptions of vowels are better done auditorily, nevertheless it is convenient to have available a rough scheme of articulatory classification. Such a scheme is represented by the vowel diagram on the chart of the International Phonetic Alphabet (IPA) as shown in Table 1. It will be noticed that this is of similar shape to the Cardinal Vowel diagram although a single line is used centrally rather than a triangle. Labels are provided to distinguish between FRONT, CENTRAL and BACK, and between four degrees of opening: CLOSE, CLOSE-MID, OPEN-MID and OPEN (see Fig. 7). At each intersection point on the periphery of the diagram on the IPA chart (Table 1) two symbols are supplied; these symbols are the same as those used for the Cardinal Vowels. However, on the IPA chart the UNROUNDED vowel is always the first of the pair and the ROUNDED the second;

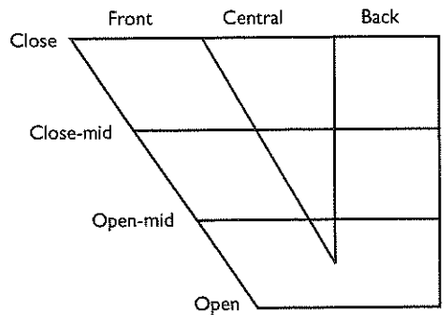


Figure 7 Articulatory labels combined with the Cardinal Vowel diagram.

this means that we cannot say that the first corresponds to the primary cardinal and the second to the secondary cardinal. (It will be remembered that primary cardinals involve the most frequent lip positions, back vowels being more usually rounded). The IPA diagram also supplies us with a number of additional symbols for vowels in certain positions, [ɪ,a,i̞,ə,ɐ] being used for unrounded vowels and [ʉ,ʊ,ɵ] for rounded vowels.

Notes

1 See Cruttenden (2013) for discussion of the use of these videos in teaching.
2 These are called post-alveolar on the chart of the International Phonetic Alphabet (Table 1).
3 Go to http://ipa.group.shef.ac.uk to see and hear recordings of each sound in the International Phonetic Alphabet.
4 See Duckworth *et al.* (1990).
5 Copies of the original recording of the Cardinal Vowels by Daniel Jones are on the companion website.

Chapter 5

Sounds in language

5.1 Speech sounds and linguistic units

We now have a way of classifying the sounds which can be produced by the speech organs. A speech sound produced in isolation can be described in purely phonetic terms; but any purely phonetic approach to the sounds of language encounters difficulties because speech is normally a continuum of sound. Two initial problems concern, first, the identification and delimitation of the sound unit (or segment) to be described and, second, the way in which different sounds are treated, for the purpose of linguistic analysis, as if they were the same.

As we have seen, in any investigation of speech, it is on the physiological and acoustic levels that most information is available to us. But in any articulation, as revealed by MRI, an utterance consists of apparently continuous movements by a very large number of organs; it is almost impossible to say, simply from a video of the speech organs at work, how many speech sounds have been uttered. A display of acoustic information is slightly easier to handle (see Fig. 3), but even here it is not always possible to delimit exactly the beginning and end of sound segments because of the way in which many sounds merge into one another. Moreover, even if we were able to delimit and identify certain sounds, it would not follow that all the individual units would fit into a useful linguistic description of the language being investigated. Thus, the word *tot* is frequently pronounced in the London region in such a way that it is possible to identify five sound segments: [t], [s], [h], [ɒ], [t]. Yet much of this phonetic reality may be discarded as irrelevant when it is a question of the structure of the word *tot* in terms of the sound system of English. Indeed, the speaker himself will probably feel that the utterance *tot* consists of only three 'sounds', such a judgement on his part being a highly sophisticated one which results from his experience in hearing and speaking English (and not only because of influence from the spelling). In other words, the [s] and [h] segments are to be treated as part of the phonological, or linguistic, unit /t/.[1] The phonetic sequence [tsh] does not, in an initial position in this type of English, consist of three meaningful units; in other languages, on the other hand, such a sequence might well constitute three linguistic units as well as three phonetic segments.