ENGLISH PHONETICS AND PHONOLOGY

GLOSSARY

(A LITTLE ENCYCLOPAEDIA OF PHONETICS)

Peter Roach

- CLICK ON A LETTER FROM THE LIST BELOW TO JUMP TO THE RELEVANT PAGE
- CLICK HERE TO JUMP TO THE INDEX

A B C D E F G H I J K L M N O P R S T U V W X

accent

This word is used (rather confusingly) in two different senses: (1) accent may refer to prominence given to a syllable, usually by the use of pitch. For example, in the word 'potato' the middle syllable is the most prominent; if you say the word on its own you will probably produce a fall in pitch on the middle syllable, making that syllable accented. In this sense, *accent* is distinguished from the more general term stress, which is more often used to refer to all sorts of prominence (including prominence resulting from increased loudness, length or sound quality), or to refer to the effort made by the speaker in producing a stressed syllable. (2) accent also refers to a particular way of pronouncing: for example, you might find a number of English speakers who all share the same grammar and vocabulary, but pronounce what they say with different accents such as Scots or Cockney, or BBC pronunciation. The word

accent in this sense is distinguished from dialect, which usually refers to a variety of a language that differs from other varieties in grammar and/or vocabulary.

START ~ INDEX

acoustic phonetics

An important part of phonetics is the study of the physics of the speech signal: when sound travels through the air from the speaker's mouth to the hearer's ear it does so in the form of vibrations in the air. It is possible to measure and analyse these vibrations by mathematical techniques, usually by using specially-developed computer software to produce spectrograms. Acoustic phonetics also studies the relationship between activity in the speaker's vocal tract and the resulting sounds. Analysis of speech by acoustic phonetics is claimed to be more objective and scientific than the traditional auditory method which depends on the reliability of the trained human ear.

START ~ INDEX

active articulator

See articulator

START ~ INDEX

Adam's Apple

This is an informal term used to refer to the pointed part of the larynx that can be seen at the front of the throat. It is most clearly visible in adult males. Moving the larynx up and down (as in swallowing) causes visible movement of this point, which is in fact the highest point of the thyroid cartilage.

START ~ INDEX

affricate

An affricate is a type of consonant consisting of a plosive followed by a fricative with the same place of articulation: examples are the **tʃ** and **dʒ** sounds at the beginning and end of the English words 'church' **tʃɜːtʃ**, 'judge' **dʒʌdʒ** (the first of these is voiceless, the second voiced). It is often difficult to decide whether any particular combination of a plosive plus a fricative should be classed as a single affricate sound or as two separate sounds, and the question depends on whether these are to be

regarded as separate phonemes or not. It is usual to regard **tʃ**, **dʒ** as affricate phonemes in English (usually symbolised **č**, **ǰ** by American writers); **ts**, **dz**, **tr**, **dr** also occur in English but are not usually regarded as affricates. The two phrases 'why choose' **waɪ tʃuːz** and 'white shoes' **waɪt ʃuːz** are said to show the difference between the **tʃ** affricate (in the first example) and separate **t** and **ʃ** (in the second).

START ~ *INDEX*

airstream

All speech sounds are made by making air move. Usually the air is moved outwards from the body, creating an egressive airstream; more rarely, speech sounds are made by drawing air into the body – an ingressive airstream. The most common way of moving air is by compression of the lungs so that the air is expelled through the vocal tract. This is called a pulmonic airstream (usually an egressive pulmonic one, but occasionally speech is produced while breathing in). Others are the glottalic (produced by the larynx with closed vocal folds; it is moved up and down like the plunger of a bicycle pump) and the velaric (where the back of the tongue is pressed against the soft palate, or velum, making an air-tight seal, and then drawn backwards or forwards to produce an airstream). Ingressive glottalic consonants (often called implosives) and egressive ones (ejectives) are found in many non-European languages; click sounds (ingressive velaric) are much rarer, but occur in a number of southern African languages such as Nàmá, Xhosa and Zulu. Speakers of other languages, including English, use click sounds for non-linguistic communication, as in the case of the "tut-tut" (American "tsk-tsk") sound of disapproval.

START ~ *INDEX*

allophone

Central to the concept of the phoneme is the idea that it may be pronounced in many different ways. In English (BBC pronunciation) we take it for granted that the **r** sounds in 'ray' and 'tray' are "the same sound" (i.e. the same phoneme), but in reality the two sounds are very different – the **r** in 'ray' is voiced and non-fricative, while the **r** sound in 'tray' is voiceless and fricative. In phonemic transcription we use the same symbol **r** for both, but we know that the allophones of **r** include the voiced non-fricative sound **ɹ** and the voiceless fricative one **ʂ**.

In theory a phoneme can have an infinite number of allophones, but in practice for descriptive purposes we tend to concentrate on a small number that occur most regularly.

START ~ *INDEX*

alveolar

Behind the upper front teeth there is a hard, bony ridge called the alveolar ridge; the skin covering it is corrugated with transverse wrinkles. The tongue comes into contact with this in some of the consonants of English and many other languages; sounds such as **t**, **d**, **s**, **z**, **n**, **l** are consonants with alveolar place of articulation.

START ~ *INDEX*

alveolo-palatal

When we look at the places of articulation used by different languages, we find many differences in the region between the upper teeth and the front part of the palate. It has been proposed that there is difference between alveolo-palatal and palato-alveolar that can be reliably distinguished, though others argue that factors other than place of articulation are usually involved, and there is no longer an alveolo-palatal column on the IPA chart. The former place is further forward in the mouth than the latter: the usual example given for a contrast between alveolo-palatal and palato-alveolar consonants is that of Polish **ɕ** and **ʃ** as in 'Kasia' **kaɕa** and 'kasza' **kaʃa**.

START ~ *INDEX*

ambisyllabic

We face various problems in attempting to decide on the division of English syllables: in a word like 'better' **betə** the division could be (using the **.** symbol to mark syllable divisions) *either* **be.tə** *or* **bet.ə**, and we need a principle to base our decision on. Some phonologists have suggested that in such a case we should say that the **t** consonant belongs to *both* syllables, and is therefore ambisyllabic; the analysis of 'better' **betə** is then that it consists of the syllables **bet** and **tə**.

START ~ *INDEX*

anterior

In phonology it is sometimes necessary to distinguish the class of sounds that are articulated in the front part of the mouth (anterior sounds) from those articulated towards the back of the mouth. All sounds forward of palato-alveolar are classed as anterior.

START ~ *INDEX*

apical

Consonantal articulations made with the tip of the tongue are called apical; this term is usually contrasted with laminal, the adjective used to refer to tongue-blade articulations. It is said that English **s** is usually articulated with the tongue blade, but Spanish **s** (when it occurs before a vowel) and Greek **s** are said to be apical, giving a different sound quality.

START ~ *INDEX*

approximant

This is a phonetic term of comparatively recent origin. It is used to denote a consonant which makes very little obstruction to the airflow. Traditionally these have been divided into two groups: "semivowels" such as the **w** in English 'wet' and **j** in English 'yet', which are very similar to close vowels such as [**u**] and [**i**] but are produced as a rapid glide; and "liquids", sounds which have an identifiable constriction of the airflow but not one that is sufficiently obstructive to produce fricative noise, compression or the diversion of airflow through another part of the vocal tract as in nasals. This category includes laterals such as English **l** in 'lead' and non-fricative **r** (phonetically **ɹ**) in 'read'. Approximants therefore are never fricative and never contain interruptions to the flow of air.

START ~ *INDEX*

articulation

See articulator.

START ~ *INDEX*

articulator/ory/ation

The concept of the articulator is a very important one in phonetics. We can only produce speech sound by moving parts of our body, and this is done by the contraction of muscles. Most of the movements relevant to speech take place in the mouth and throat area (though we should not forget the activity in the chest for breath control), and the parts of the mouth and throat area that we move when speaking are called articulators. The principal articulators are the tongue, the lips, the lower jaw and the teeth, the velum or soft palate, the uvula and the larynx. It has been suggested that we should distinguish between active articulators (those which can be moved into contact with other articulators, such as the tongue) and passive articulators which are

fixed in place (such as the teeth, the hard palate and the alveolar ridge). The branch of phonetics that studies articulators and their actions is called articulatory phonetics.

START ~ *INDEX*

articulatory setting

This is an idea that has an immediate appeal to pronunciation teachers, but has never been fully investigated. The idea is that when we pronounce a foreign language, we need to set our whole speech-producing apparatus into an appropriate "posture" or "setting" for speaking that language. English speakers with a good French accent, for example, are said to adjust their lips to a more protruded and rounded shape than they use for speaking English, and people who can speak several languages are claimed to have different "gears" to shift into when they start saying something in one of their languages. (See also voice quality.)

START ~ *INDEX*

arytenoids

Inside the larynx there is a tiny pair of cartilages shaped rather like dogs' ears. They can be moved in many different directions. The rear ends of the vocal folds are attached to them so that if the arytenoids are moved towards each other the folds are brought together, making a glottal closure or constriction, and when they are moved apart the folds are parted to produce an open glottis. The arytenoids contribute to the regulations of pitch: if they are tilted backwards the vocal folds are stretched lengthwise (which raises the pitch if voicing is going on), while tilting them forwards lowers the pitch as the folds become thicker.

START ~ *INDEX*

aspiration

This is noise made when a consonantal constriction is released and air is allowed to escape relatively freely. English **p t k** at the beginning of a syllable are aspirated in most accents so that in words like 'pea', 'tea', 'key' the silent period while the compressed air is prevented from escaping by the articulatory closure is followed by a sound similar to **h** before the voicing of the vowel begins. This is the result of the vocal folds being widely parted at the time of the articulatory release. It is noticeable that when **p t k** are preceded by **s** at the beginning of a syllable they are not aspirated. Pronunciation teachers used to make learners of English practise aspirated plosives by seeing if they could blow out a candle flame with the rush of air after **p t k** – this can, of course, lead to a rather exaggerated pronunciation (and superficial burns). A rather different articulation is used for so-called voiced aspirated plosives found in many

Indian languages (often spelt ‘bh’, ‘dh’, ‘gh’ in the Roman alphabet) where after the release of the constriction the vocal folds vibrate to produce voicing, but are not firmly pressed together; the result is that a large amount of air escapes at the same time, producing a “breathy” quality.

It is not necessarily only plosives that are aspirated: both unaspirated and aspirated affricates are found in Hindi, for example, and unaspirated and aspirated voiceless fricatives are found in Burmese.

See also voice onset time (VOT).

START ~ *INDEX*

assimilation

If speech is thought of as a string of sounds linked together, assimilation is what happens to a sound when it is influenced by one of its neighbours. For example, the word ‘this’ has the sound **s** at the end if it is pronounced on its own, but when followed by **ʃ** in a word such as ‘shop’ it often changes in rapid speech (through assimilation) to **ʃ**, giving the pronunciation **ðɪʃʃɒp**. Assimilation is said to be *progressive* when a sound influences a following sound, or *regressive* when a sound influences one which precedes it; the most familiar case of regressive assimilation in English is that of alveolar consonants, such as **t**, **d**, **s**, **z**, **n**, which are followed by non-alveolar consonants: assimilation results in a change of place of articulation from alveolar to a different place. The example of ‘this shop’ is of this type; others are ‘football’ (where ‘foot’ **fʊt** and ‘ball’ **bɔːl** combine to produce **fʊpbɔːl**) and ‘fruit-cake’(**fruːt** + **keɪk** → **frʊːkkeɪk**). Progressive assimilation is exemplified by the behaviour of the ‘s’ plural ending in English, which is pronounced with a voiced **z** after a voiced consonant (e.g. ‘dogs’ **dɒgz**) but with a voiceless **s** after a voiceless consonant (e.g. ‘cats’ **kæts**).

The notion of assimilation is full of problems: it is often unhelpful to think of it in terms of one sound being the cause of the assimilation and the other the victim of it, when in many cases sounds appear to influence each other mutually; it is often not clear whether the result of assimilation is supposed to be a different allophone or a different phoneme; and we find many cases where instances of assimilation seem to spread over many sounds instead of being restricted to two adjacent sounds as the conventional examples suggest. Research on such phenomena in experimental phonetics does not usually use the notion of assimilation, preferring the more neutral concept of coarticulation.

START ~ *INDEX*

attitude/inal

Intonation is often said to have an attitudinal function. What this means is that intonation is used to indicate to the hearer a particular attitude on the part of the speaker (e.g. friendly, doubtful, enthusiastic). Considerable importance has been given by some language teaching experts to learning to express the right attitudes through intonation, but it has proved extremely difficult to state usable rules for foreigners to learn and results have often been disappointing. It has also proved very difficult to design and carry out scientific studies of the way intonation conveys attitudes and emotions in normal speech.

START ~ *INDEX*

auditory

When the analysis of speech is carried out by the listener's ear, the analysis is said to be an auditory one, and when the listener's brain receives information from the ears it is said to be receiving auditory information. In practical phonetics, great importance has been given to auditory training: this is sometimes known as ear-training, but in fact it is the brain and not the ear that is trained. With expert teaching and regular practice, it is possible to learn to make much more precise and reliable discriminations among speech sounds than untrained people are capable of. Although the analysis of speech sounds by the trained expert can be carried out entirely auditorily, in most cases the analyst also tries to make the sound (particularly when working face to face with a native speaker of the language or dialect), and the proper name for this analysis is then *auditory-kinaesthetic*.

START ~ *INDEX*

autosegmental phonology

One fairly recent development in phonology is one which attempts to separate out the phonological material of an utterance into components on different levels. For example, if we give a fall–rise intonation pattern to the following two utterances:

\/some and \/some of them

the pitch movement is phonologically the same object in both cases, but stretches over a longer sequences of syllables in the second case. We can make up similar examples in terms of rhythm, using the unit of the foot, and autosegmental phonology is closely linked to metrical phonology.

Although this is an approach that was mainly developed in the 1990s in America, it is very similar to the Prosodic Phonology proposed by J. R. Firth and his associates at the School of Oriental and African Studies of London University in the 1940s and 50s.

START ~ INDEX

BBC pronunciation

The British Broadcasting Corporation is looked up to by many people in Britain and abroad as a custodian of good English; this attitude is normally only in respect of certain broadcasters who represent the formal style of the Corporation, such as newsreaders and announcers, and does not apply to the more informal voices of people such as disc-jockeys and chat-show presenters (who may speak as they please). The high status given to the BBC's voices relates both to pronunciation and to grammar, and there are listeners who write angry letters to the BBC or the newspapers to complain about "incorrect" pronunciations such as "loranorder" for "law and order". Although the attitude that the BBC has a responsibility to preserve some imaginary pure form of English for posterity is extreme, there is much to be said for using the "formal" BBC accent as a model for foreign learners wishing to acquire an English accent. The old standard "Received Pronunciation (RP)" is based on a very old-fashioned view of the language; the present-day BBC accent is easily accessible and easy to record and examine. It is relatively free from class-based associations and it is available throughout the world where BBC broadcasts can be received; however, in recent years, the Overseas Service of the BBC has taken to using a number of newsreaders and announcers who are not native speakers of English and have what is, by British standards, a foreign accent. The BBC nowadays uses quite a large number of speakers from Celtic countries (particularly Ireland, Scotland and Wales), and the description of "BBC Pronunciation" should not be treated as including such speakers.

The Corporation has its own Pronunciation Research Unit, but contrary to some people's belief its function is to advise on the pronunciation of foreign words and of obscure British names and not to monitor pronunciation standards. Broadcasters are not under any obligation to consult the Unit.

START ~ INDEX

bilabial

A sound made with both lips. See labial, place of articulation.

START ~ INDEX

binary

Phonologists like to make clear-cut divisions between groups of sounds, and usually this involves "either-or" choices: a sound is either voiced or voiceless, consonantal or

non-consonantal, rounded or unrounded. Such choices are binary choices. In the study of phonetics, however, it is acknowledged that sounds differ from each other in "more or less" fashion rather than "either-or": features like voicing, nasality or rounding are scalar or multi-valued, and a sound can be, for example, fully voiced, partly voiced, just a little bit voiced or not voiced at all.

When distinctive features of sounds are given binary values, they are usually marked with the plus and minus signs + and - , so a voiced consonant is classed as +voice and a voiceless one as -voice.

START ~ *INDEX*

boundary

The notion of the boundary is very important in phonetics and phonology. At the segmental level, we need to know where one segment ends and another begins, and this can be a difficult matter: in a word like 'hairier' **heəriə**, which contains no plosives or fricatives, each sound seems to merge gradually into the next. In dividing words into syllables we have many difficulties, resulting in ideas like juncture and ambisyllabicity to help us solve them. In intonation we have many different units at different levels, and dividing continuous speech into tone-units separated by boundaries is one of the most difficult problems.

START ~ *INDEX*

brackets

When we write in phonetic or phonemic transcription it is conventional to use brackets at the beginning and end of the item or passage to indicate the nature of the symbols. Generally, slant brackets (also known as "obliques") are used to indicate phonemic transcription and square brackets for phonetic transcription. For example, for the word 'phonetics' we would write / **fənetɪks** / (phonemic transcription) and [**fənetʰɪʔks**] (phonetic transcription). However, in writing *English Phonetics and Phonology* I decided not to use brackets in this way, apart from using square brackets when representing cardinal vowels, because I thought that this would make the transcriptions easier to read, and that it would almost always be obvious which type of transcription was being used in a given place.

START ~ *INDEX*

breath-group

In order to carry out detailed analysis, linguists need to divide continuous speech into small, identifiable units. In the present-day written forms of European languages, the

sentence is an easy unit to work with, and the full stop ("period" in American English) clearly marks its boundaries. It would be helpful if we could identify something similar in spoken language and one possible candidate is a unit whose boundaries are marked by the places where we pause to breathe: the breath-group. Unfortunately, although in the production of isolated sentences and in very careful speech the places where a speaker will breathe may be quite predictable, in natural speech such regularity disappears, so that the breath-group can vary very greatly in terms of its length and its relationship to linguistic structure. It is, consequently, little used in modern phonetics and linguistics.

START ~ *INDEX*

breathing

This is the movement of air into and out of the lungs. Speech is something which is imposed on normal breathing, resulting in a reduced rate of air-flow out of the body. Mostly the air pressure that pushes air out and allows us to produce speech sounds is caused by the chest walls pressing down on the lungs, but we can give the air an extra push with the diaphragm, a large sheet of muscle lying between the lungs and the stomach.

START ~ *INDEX*

breathy

This is one of the adjectives used to describe voice quality or phonation type. In breathy voice, the vocal folds vibrate but allow a considerable amount of air to escape at the same time; this adds "noise" (similar to loud breathing) to the sound produced by the vocal folds. It is conventionally thought that breathy voice makes women's voices sound attractive, and it is used by speakers in television advertisements for "soft" products like toilet paper and baby powder.

START ~ *INDEX*

burst

When a plosive (such as English **p**, **t**, **k**, **b**, **d**, **g**) is released while air is still compressed within the vocal tract, the air rushes out with some force. The resulting sound is usually referred to as plosion in general phonetic terminology, but in acoustic phonetics it is more common to refer to this as a burst. It is usually very brief – somewhere around a hundredth of a second.

START ~ *INDEX*

cardinal vowel

Phoneticians have always needed some way of classifying vowels which is independent of the vowel system of a particular language. With most consonants it is quite easy to observe how their articulation is organised, and to specify the place and manner of the constriction formed; vowels, however, are much less easy to observe. Early in the 20th century, the English phonetician Daniel Jones worked out a set of "cardinal vowels" that students learning phonetics could be taught to make and which would serve as reference points that other vowels could be related to, rather like the corners and sides of a map. Jones was strongly influenced by the French phonetician Paul Passy, and it has been claimed that the set of cardinal vowels is rather similar to the vowels of educated Parisian French of the time.

From the beginning it was important to locate the vowels on a chart or four-sided figure (the exact shape of which has changed from time to time), as can be seen on the IPA chart. The cardinal vowel diagram is used both for rounded and unrounded vowels, and Jones proposed that there should be a primary set of cardinal vowels and a secondary set. The primary set includes the front unrounded vowels [**ɪ**, **e**, **ɛ**, **a**], the back unrounded vowel [**ɑ**] and the rounded back vowels [**ɔ**, **o**, **u**], while the secondary set comprises the front rounded vowels [**y**, **ø**, **œ**, **ɶ**], the back rounded [**ɒ**] and the back unrounded [**ʌ**, **ɤ**, **ɯ**]. For the sake of consistency, I believe it would be better to abandon the "primary/secondary" division and simply give a "rounded" or "unrounded" label (as appropriate) to each vowel on the quadrilateral.

Phonetic "ear-training" makes much use of the cardinal cowel system, and students can learn to identify and discriminate a very large number of different vowels in relation to the cardinal vowels.

START ~ *INDEX*

cartilage

Many parts of the body used in speech are made of cartilage, which is less hard than bone. In particular, the structure of the larynx is largely made of cartilage, though as we get older some of this turns to bone.

START ~ *INDEX*

centre/al

A vowel is central if it is produced with the central part of the tongue raised (i.e. it is neither front like [**i**] nor back like [**u**]). All descriptions of vowel quality recognise a vowel that is both central (i.e. between front and back) and mid (i.e. half-way between close and open), usually named schwa (for which the symbol is [**ə**]). Phonetic symbols exist also for central vowels which are close - either rounded [**ʉ**] or

unrounded [ɨ] – and for open-mid to open unrounded [ɐ], as well as close-mid and open-mid (see the IPA Chart). Apart from the symbol used for the English vowel in 'fur' [ɜ] these are little used.

START ~ *INDEX*

chart

It is usual to display sets of phonetic symbols on a diagram made of a rectangle divided into squares, usually called a chart, but sometimes called a matrix or a grid. The best-known phonetic chart is that of the alphabet of the International Phonetic Association – the IPA chart. On this chart the vertical axis represents the manner of articulation of a sound (e.g. plosive, nasal) and the horizontal axis represents the place of articulation (e.g. bilabial, velar). Within each box on the chart it is possible to have two symbols, of which the left hand one will be voiceless and the right hand voiced. A number of charts are given in *English Phonetics and Phonology*; the IPA chart is printed on page xii.

START ~ *INDEX*

chest-pulse

This is a notion used in the theory of syllable production. Early in the twentieth century it was believed by some phoneticians that there was a physiological basis to the production of syllables: experimental work was claimed to show that for each syllable produced, there was a distinct effort, or pulse, from the chest muscles which regulate breathing. It is now known that chest-pulses are not found for every syllable in normal speech, though there is some evidence that there may be chest-pulses for stressed syllables.

START ~ *INDEX*

clear l

This is a type of lateral sound (such as the English l in 'lily'), in which the air escapes past the sides of the tongue. In the case of an alveolar lateral (e.g. English l) the blade of the tongue is in contact with the alveolar ridge, but the rest of the tongue is free to take up different shapes. One possibility is for the front of the tongue (the part behind the blade) to be raised in the same shape as that for a close front vowel [i]. This gives the l an [i]-like sound, and the result is a "clear l". It is found in BBC English only before vowels, but in some other accents, notably Irish and Welsh ones, it is found in all positions.

See also dark l.

START ~ INDEX

click

Clicks are sounds that are made within the mouth and are found as consonantal speech sounds in some languages of Southern Africa, such as Xhosa (the name of which itself begins with a click) and Zulu. Clicks are more familiar to English speakers as non-speech sounds such as the "tut-tut"or "tsk-tsk" sound of disapproval. A different type of click sound (a lateral click) is (or was) used to make a horse move on, and also for some social purposes such as expressing satisfaction. The way in which these sounds are made is for the back of the tongue to make an air-tight closure against the back of the palate (see velaric airstream); an articulatory closure is then made further forward in the mouth and this results in a completely sealed air chamber within the mouth. The back of the tongue is then drawn backwards, which has the effect of lowering the air pressure within the chamber so that if the forward articulatory closure is released quickly a plosive sound is heard. There are many variations on this mechanism, including voicing, affricated release, and simultaneous nasal consonant.

START ~ INDEX

clipped

The term "clipped speech" has two meanings in the context of speech: in non-technical usage it refers to a style of speaking often associated with military men and "horsey" people, characterised by unusually short vowels; the term is also used in the study of speech acoustics to refer to a speech signal that has been distorted in a particular way, usually through overloading.

START ~ INDEX

close vowel

In a close vowel the tongue is raised as close to the palate as is possible without producing fricative noise. Close vowels may be front (when the front of the tongue is raised), either unrounded [i] or rounded [y], or they may be back (when the back of the tongue is raised), either rounded [u] or unrounded [ɯ]. There are also close central vowels: rounded [ʉ] and unrounded [ɨ]. English i and u are often described as close vowels, but are rarely fully close in English accents.

See also open.

START ~ INDEX

closure

This word is one of the unfortunate cases where different meanings are given by different phoneticians: it is generally used in relation to the production of plosive consonants, which require a total obstruction to the flow of air. To produce this obstruction, the articulators must first move towards each other, and must then be held together to prevent the escape of air. Some writers use the term closure to refer to the coming together of the articulators, while others use it to refer to the period when the compressed air is held in.

START ~ *INDEX*

cluster

In some languages (including English) we can find several consonant phonemes in a sequence, with no vowel sound between them: for example, the word 'stray' **streɪ** begins with three consonants, and 'sixths' **sɪksθs** ends with four. Sequences of two or more consonants within the same syllable are often called consonant clusters. It is not usual to refer to sequences of vowels as vowel clusters.

START ~ *INDEX*

coalescence

Speech sounds rarely have clear-cut boundaries that mark them off from their neighbours. It sometimes happens that adjacent phonemes slide together (*coalesce*) so that they seem to happen simultaneously. An example is what is sometimes called *yod-coalescence*, where a sound preceding a **j** ("yod") becomes palatalised: thus the **s** at the end of 'this' can merge with the **j** of 'year' to give a pronunciation **ðɪʃʃɪə** or **ðɪʃɪə**.

START ~ *INDEX*

coarticulation

Experimental phonetics studies coarticulation as a way of finding out how the brain controls the production of speech. When we speak, many muscles are active at the same time and sometimes the brain tries to make them do things that they are not capable of. For example, in the word 'Mum' **mʌm** the vowel phoneme is one that is normally pronounced with the soft palate raised to prevent the escape of air through the nose, while the two **m** phonemes must have the soft palate lowered. The soft palate cannot be raised very quickly, so the vowel is likely to be pronounced with the soft palate still lowered, giving a nasalised quality to the vowel. The nasalization is a

coarticulation effect caused by the nasal consonant environment. Another example is the lip-rounding of a consonant in the environment of rounded vowels: in the phrase 'you too', the **t** occurs between two rounded vowels, and there is not enough time in normal speech for the lips to move from rounded to unrounded and back again in a few hundredths of a second; consequently the **t** is pronounced with lip-rounding.

Coarticulation is a phenomenon closely related to assimilation; the major difference is that assimilation is used as a name for the process whereby one sound becomes like another neighbouring sound, while coarticulation, though it refers to a similar process, is concerned with articulatory explanations for why the assimilation occurs, and considers cases where the changes may occur over a number of segments.

START ~ *INDEX*

cocktail party phenomenon

If you are at a noisy party with a lot of people talking close to you, it is a striking fact that you are able to choose to listen to one person's voice and to "shut out" what others are saying equally loudly. The importance of this effect was first highlighted by the communications engineer Colin Cherry, and has led to many interesting experiments by psychologists and psycholinguists. Cocktail parties are hard to find nowadays, but you can simulate the effect by making someone wear headphones and playing simultaneous voices to them, one in each ear, and asking them to concentrate on just one voice. The voices may be presented separately to each ear (dichotic listening) or mixed together and played to both ears (binaural listening).

START ~ *INDEX*

coda

This term refers to the end of a syllable. The central part of a syllable is almost always a vowel, and if the syllable contains nothing after the vowel it is said to have no coda ("zero coda"). Some languages have no codas in any syllables. English allows up to four consonants to occur in the coda, so the total number of possible codas in English is very large – several hundred, in fact.

START ~ *INDEX*

commutation

When we want to demonstrate that two sounds are in phonemic opposition, we normally do this with the commutation test; this means substituting one sound for another in a particular phonological context. For example, to prove that the sounds **p**,

b, **t**, **d** are different contrasting phonemes we can try them one at a time in a suitable context which is kept constant; using the context **-n** we get 'pin', 'bin', 'tin' and 'din', all of which are different words.

There are serious theoretical problems with this test. One of them is the widespread assumption that if you substitute one allophone of a phoneme for another allophone of the same phoneme, the meaning will not change; this is sometimes true (substituting a "dark l" where a "clear l" is appropriate in BBC pronunciation, for example, is unlikely to change a perceived meaning) but in other cases it is at least dubious: for example, the unaspirated allophones of **p**, **t**, **k** found after **s** at the beginning of syllables such as **sp**, **st**, **sk** are phonetically very similar to **b**, **d**, **g** , and pronouncing one of these unaspirated allophones followed by **-ɪl**, for example, would be likely to result in the listener hearing 'bill', 'dill', 'gill' rather than 'pill', 'till', 'kill'.

START ~ *INDEX*

complementary distribution

Two sounds are in complementary distribution if they never occur in the same context. A good example is provided by the allophones of the **l** phoneme in BBC pronunciation: there is a voiceless allophone **l̥** when **l** occurs after **p**, **t**, **k** at the beginning of a syllable, "clear l" which occurs before vowels and "dark l" which occurs elsewhere (i.e. before consonants or a pause). Leaving aside less noticeable allophonic variation, these three allophones together account for practically all the different ways in which the **l** phoneme is realised; since each of them has its own specific context in which it occurs, and does not occur in the contexts in which the others occur, we can say that each is in complementary distribution with the others.

In conventional phoneme theory, sounds which are in complementary distribution are likely to belong to the same phoneme; thus "voiceless l", "clear l" and "dark l" in the example given above will be classed as members of the same phoneme. There are problems in the argument, however: we can find quite a lot of sounds in English, for example, which are in complementary distribution with each other but are still not considered members of the same phoneme, a frequently quoted case being that of **h** (which cannot occur at the end of a syllable) and **ŋ** (which cannot occur at the beginning of a syllable) – this forces us to say that sounds which are in complementary distribution and are to be considered as allophones of the same phoneme must be phonetically similar to each other (which **h** and **ŋ** clearly are not). But measuring phonetic similarity is itself a very problematical area.

START ~ *INDEX*

consonant

There are many types of consonant, but what all have in common is that they obstruct the flow of air through the vocal tract. Some do this a lot, some not very much: those which make the maximum obstruction (i.e. plosives, which form a complete stoppage of the airstream) are the most consonantal. Nasal consonants result in complete stoppage of the oral cavity but are less obstructive than plosives since air is allowed to escape through the nose. Fricatives make a considerable obstruction to the flow of air, but not a total closure. Laterals obstruct the flow of air only in the centre of the mouth, not at the sides, so obstruction is slight. Other sounds classed as approximants make so little obstruction to the flow of air that they could almost be thought to be vowels if they were in a different context (e.g. English **w** or **r**).

The above explanation is based on phonetic criteria. An alternative approach is to look at the phonological characteristics of consonants: for example, consonants are typically found at the beginning and end of syllables while vowels are typically found in the middle. See also contoid.

START ~ *INDEX*

constriction

All speech sounds apart from fully-open vowels involve some narrowing (constriction) of the vocal tract, and one of the most important ways in which speech sounds differ from each other is the position of the constriction and the degree of narrowing of the constriction. In addition to the main constriction there is often also a secondary constriction: for example, the ʃ sound in English has a primary constriction in the post-alveolar region (where the fricative noise is produced), but many English speakers produce the sound with lip-rounding and this creates a secondary constriction at the lips.

START ~ *INDEX*

continuant

It is sometimes useful to have a word for speech sounds which can be produced as a continuous sound. A vowel is thus a continuant, while a plosive is not. A vowel, or other continuant sounds such as nasals and fricatives, can be continued for as long as the speaker has enough breath.

START ~ *INDEX*

contoid

For most practical purposes a contoid is the same thing as a consonant; however, there are reasons for having a distinction between sounds which function phonologically as consonants and sounds (contoids) which have the phonetic characteristics that we look on as consonantal. As an example, let us look at English **w** (as in 'wet') and **j** (as in 'yet'). If you pronounce these two sounds very slowly you will hear that they are closely similar to the vowels [**i**] and [**u**] – yet English speakers treat them as consonants. How do we know this? Consider the pronunciation of the indefinite article: the rule is to use 'a' before consonants and 'an' before vowels, and it is the former version which we find before **w** and **j**; similarly, the definite article is pronounced **ði** before a vowel but **ðə** before a consonant, and we find the **ðə** form before **j** and **w**.

Another interesting case is the normal pronunciation of the **r** phoneme in the BBC accent – in many ways this sound is more like a vowel than a consonant, and in some languages it actually is found as one of the vowels, yet we always treat it as a consonant.

The conclusion that has been drawn is that since the word 'consonant' as used in describing the phonology of a language can include sounds which could be classed phonetically as vowels, we ought also to have a different word which covers just those sounds which are phonetically of the type that produces a significant obstruction to the flow of air through the vocal tract (see consonant above): the term proposed is contoid.

START ~ *INDEX*

contour

It is usual to describe a movement of the pitch of the voice in speech as a contour. In the intonation of a language like English many syllables are said with a fairly level tone, but the most prominent syllables are said with a tonal contour (which may be continued on following syllables). In the study of tone languages it is usual to make a distinction between register languages which generally use only phonologically level tones (e.g. many West African languages) and those which also use contour tones such as rises, falls, fall–rises and rise–falls (e.g. many East Asian languages, such as Chinese).

START ~ *INDEX*

contraction

English speech has a number of cases where pairs of words are closely combined into a contracted form that is almost like a single word. For example, 'that' and 'is' are

often contracted to ‘that’s’. These forms are so well established in spoken English that they have their own representation in the spelling. There is a brief list of these in *English Phonetics and Phonology*, page 114.

START ~ *INDEX*

contrast

A notion of central importance in traditional phoneme theory is that of contrast: while it is important to know what a phoneme is (in terms of its sound quality, articulation and so on), it is vital to know what it is not – i.e. what other sounds it is in contrast with. For example, English **t** contrasts with **p** and **k** in place of articulation, with **d** (in the matter of voicing or force of articulation), **n** (by being plosive rather than nasal), and so on. Phonologists have claimed that the English **n** sound is different from the phonetically similar sound **n̪** in the Indian language Malayalam, since in English the only other voiceless plosive consonants that **n** contrasts with are **m** and **ŋ**, whereas in Malayalam **n̪** contrasts not only with **m** and **ŋ** but also with the nasal consonants **n̪** and **ɳ**.

Some phonologists state that a theoretical distinction must be made between contrast and opposition. In their use of the terms, ‘opposition’ is used for the “substitutability” relationship described above, while ‘contrast’ is reserved to refer to the relationship between a sound and those adjacent to it.

START ~ *INDEX*

conversation

The interest in conversation for the phonetics specialist lies in the differences between conversational speech and monologue. Much linguistic analysis in the past has concentrated on monologue or on pieces of conversational speech taken out of context. Specialised studies of verbal interaction between speakers look at factors such as turn-taking, the way in which interruptions are managed, the use of intonation to control the course of the conversation and variations in rhythm.

START ~ *INDEX*

coronal

A coronal sound is one in which the blade of the tongue is raised from its rest position (that is, the position for normal breathing). Examples are **t̻**, **d̻**. This term is used in phonology to refer to a distinctive feature.

START ~ INDEX

creak

Creak is a special type of vocal fold vibration that has proved very difficult to define though easy to recognise. In English it is most commonly found in adult male voices when the pitch of the voice is very low, and the resulting sound has been likened to the sound of a stick being run along railings. However, creak is also found in female voices, and it has been claimed that among female speakers creak is typical of upper-class English women. It appears to be possible to produce creak at any pitch, and a number of languages in different parts of the world make use of it contrastively (i.e. to change meanings). Some languages have creaky-voiced (or 'laryngealised') consonants (e.g. the Hausa language of West Africa), while some tone languages (e.g. Vietnamese) have creaky tones that contrast with normally-voiced ones.

It is clear that some form of extreme laryngeal constriction is involved in the production of creak, but the large number of experimental studies of the phenomenon seem to indicate that different speakers have very different ways of producing it.

START ~ INDEX

dark l

In the description of "clear l" it is explained that while the blade and tip of the tongue are fixed in contact with the alveolar ridge, the rest of the tongue is free to adopt different positions. If the back of the tongue is raised as for an [**u**] vowel, the quality is [**u**]-like and "dark"; this effect is even more noticeable if the lips are rounded at the same time. This sound is typically found in English (BBC and similar accents) when **l** occurs before a consonant (e.g. 'help') or before a pause (e.g. 'hill'). In several accents of English, particularly in the London area, the dark l has given way to a **w** sound, so that 'help' and 'hill' might be transcribed **hewp** and **hɪw**; this process (sometimes referred to as "l-vocalisation") took place in Polish some time ago, and the sound represented in Polish writing with the letter ł is almost always pronounced as **w**, though foreigners usually try to pronounce it as an **l**.

START ~ INDEX

declination

It can be claimed that there is a universal tendency in all languages to start speaking at a higher pitch than is used at the end of the utterance. Of course, it cannot be denied that pitch sometimes rises through an utterance, but this would be regarded as a special "marked" case produced for a particular reason such as signalling a question.

In tone languages the phenomenon is usually referred to as ‘downdrift’, but the term ‘declination’ has been introduced in recent work on English intonation to predict the normal pitch pattern of utterances. However, there are in English (and probably many other languages) accents where rising pitch in statements is by no means unusual or special – this is the case in accents of Northern Ireland, for example; consequently the notion of declination cannot be taken as showing that (in a literal, phonetic way) pitch always declines except in special marked cases.

START ~ *INDEX*

dental

A dental sound is one in which there is approximation or contact between the teeth and some other articulator. The articulation may be of several different sorts. The tip of the tongue may be pressed against the inner surface of the top teeth (as is usual in the **t** and **d** of Spanish and most other Romance languages); the tongue tip may be protruded between the upper and lower teeth (as in a careful pronunciation of English **θ** and **ð**); the tongue tip may be pressed against the inside of the lower teeth, with the tongue blade touching the inside of the upper front teeth, as is said to be usual for French **s** and **z**. If there is contact between lip and teeth the articulation is labelled labiodental.

START ~ *INDEX*

devoicing

A devoiced sound is one which would normally be expected to be voiced but which is pronounced without voice in a particular context: for example, the **l** in ‘blade’ **bleɪd** is usually voiced, but in ‘played’ **pleɪd** the **l** is usually voiceless because of the preceding voiceless plosive. The notion of devoicing leads to a rather confusing use of phonetic symbols in cases where there are separate symbols for voiced and voiceless pairs of sounds: a devoiced **d** can be symbolised by adding a diacritic that indicates lack of voice – **d̥** but one is then left in doubt as to what the difference is between this sound and **t**. The usual reason for doing this is to leave the symbol looking like the phoneme it represents.

START ~ *INDEX*

diacritic

A problem in the use of phonetic symbols is to know how to limit their number: it is always tempting to invent a new symbol when there is no existing symbol for a sound that one encounters. However, since it is undesirable to allow the number of symbols

to grow without limit, it is often better to add some modifying mark to an existing symbol, and these marks are called diacritics. The International Phonetic Association recognises a wide range of diacritics: for vowels, these can indicate differences in frontness, backness, closeness or openness, as well as lip-rounding or unrounding, nasalisation and centralisation. In the case of consonants, diacritics exist for voicing or voicelessness, for advanced or retracted place of articulation, aspiration and many other aspects. See the chart of the International Phonetic Alphabet.

START ~ *INDEX*

dialect

It is usual to distinguish between dialect and accent. Both terms are used to identify different varieties of a particular language, but the word 'accent' is used for varieties which differ from each other only in matters of pronunciation while 'dialect' also covers differences in such things as vocabulary and grammar.

START ~ *INDEX*

diglossia

This word is used to refer to the case where speakers of a language regularly use (or at least understand) more than one variety of that language. In one sense this situation is found in all languages: it would always be strange to talk to one's boss in the same way as one spoke to one's children. But in some languages the differences between varieties are much more sharply defined, and many societies have evolved exclusive varieties which may only be used by one sex, or in conversation between people of a particular status or relationship relative to the speaker.

START ~ *INDEX*

digraph

It has sometimes been found necessary to combine two symbols together to represent a single sound. This can happen with alphabetic writing – the term seems mainly to be used for letter pairs in words where in Roman inscriptions the letters were regularly written (or carved) joined together (e.g. spellings such as 'oe' in 'foetid' or 'ae' in 'mediaeval'), though the writing of Anglo-Saxon also involves extra symbols. It seems unlikely that anyone would call the 'ae' in 'sundae' a digraph. In the development of printed symbols some digraphs have been created, notably the combination of 'a' and 'e' in **æ** and 'o' and 'e' in **œ**; the resulting symbol when used in phonetics for vowels is supposed to signify an "intermediate" or "combined"

quality. In the case of ʧ the two symbols simply represent the phonetic sequence of events.

START ~ INDEX

diphthong

The most important feature of a diphthong is that it contains a glide from one vowel quality to another one. BBC English contains a large number of diphthongs: there are three ending in ɪ (eɪ, aɪ, ɔɪ), two ending in ʊ (əʊ, aʊ) and three ending in ə (ɪə, eʊ, ʊə). Opinions differ as to whether these should be treated as phonemes in their own right, or as combinations of two phonemes.

START ~ INDEX

discourse, discourse analysis

Although the word *discourse* has a general meaning that refers usually to speaking, in linguistics the field of *discourse analysis* has been a source of much interest for the last thirty years or so. It concentrates on language and speech as related to real-life interaction between speakers and hearers, looking at the different roles they play and the ways in which they interact. Discourse analysis has become relevant to phonetics and phonology because of what it has to say about intonation; this is explained in *English Phonetics and Phonology*, Chapter 19, Section 3.

START ~ INDEX

distinctive feature

In any language it seems that the sounds used will only differ from each other in a small number of ways. If for example a language had 40 phonemes, then in theory each of those 40 could be utterly different from the other 39. However, in practice there will usually be just a small set of important differences: some of the sounds will be vowels and some consonants; some of the consonants will be plosives and affricates, and the rest will be continuants; some of the continuants will be nasal and some not, and so on. These differences are identified by phonologists, and are known as distinctive features.

There is disagreement about how to define the features (e.g. whether they should be labelled according to articulatory characteristics or acoustic ones), and about how many features are needed in order to be able to classify the sounds of all the languages in the world. See the entry for feature.

START ~ INDEX

distribution

A very important aspect of the study of the phonology of a language is examining the contexts and positions in which each particular phoneme can occur: this is its distribution. In looking at the distribution of the **r** phoneme, for example, we can see that there is a major difference between BBC pronunciation and General American: in the former, **r** can only occur before a vowel, whereas in the latter it may occur in all positions like other consonants. It is possible to define the concepts of 'vowel' and 'consonant' purely in terms of the distributions of the two groups of sounds: as a simple example, one could list all the sounds that may begin a word in English – this would result in a list containing all the consonants except **ŋ** and all the vowels except **ʊ**. Next we would look at all the sounds that could come in second place in a word, noting which initial sound each could combine with. After the sound **æ**, for example, only consonants can follow, whereas after **ʃ**, with the exception of a few words beginning **ʃr**, such as 'shrew', only a vowel can follow. If we work carefully through all the combinatory possibilities we find that the phonemes of English separate out into two distinct groups (which we know to be vowels and consonants) without any reference to phonetic characteristics – the analysis is entirely distributional.

START ~ INDEX

dorsal

For the purposes of phonetic classification, the different regions of the surface of the tongue are given different names. Each of these names has a noun form and a corresponding adjective. The back of the tongue is involved in the production of consonants such as velar and uvular, and the adjective for the type of tongue contact used is dorsal.

START ~ INDEX

drawl

This term is quite widely used in everyday language but does not have a scientific meaning in phonetics. From the way it is used one can guess at its likely meaning: it seems to be different from speaking slowly, and probably involves the extreme lengthening of the vowels of stressed syllables. This is used to indicate a relaxed or "laid-back" attitude.

START ~ INDEX

duration

The amount of time that a sound lasts for is a very important feature of that sound. In the study of speech it is usual to use the term length for the listener's impression of how long a sounds lasts for, and duration for the physical, objectively measurable time. For example, I might listen to a recording of the following syllables and judge that the first two contained short vowels while the vowels in the second two were long: bɪt, bet, biːt, bɔːt; that is a judgement of length. But if I use a laboratory instrument to measure those recordings and find that the vowels last for 100, 110, 170 and 180 milliseconds respectively, I have made a measurement of duration.

START ~ *INDEX*

dysphonia

This is a general term used for disorders of the voice; the word 'voice' here should be taken to refer to the way in which the vocal folds vibrate. Dysphonia may result from infection (laryngitis), from a growth on the vocal fold (e.g. a polyp), from over-use (hoarseness) or from surgery.

START ~ *INDEX*

ear-training

An essential component of practical phonetic training, ear-training is used to develop the student's ability to hear very small differences between sounds (discrimination), and to identify particular sounds (identification). Although it is possible for a highly-motivated student to make considerable progress in ear-training by working from recorded material in isolation, in general it is necessary to receive training from a skilled phonetician. The "British tradition" of ear-training has grown up through the pioneering teaching of Daniel Jones, his colleagues and his former pupils, working mainly in British universities, and is maintained today by teachers trained in the same tradition.

START ~ *INDEX*

ejective

This is one of the types of speech sound that are made without the use of air pressure from the lungs – they are non-pulmonic consonants. Such sounds are much easier to demonstrate than to describe: in an ejective the vocal folds are closed, and a closure or obstruction is made somewhere in the vocal tract; then the larynx is brought upwards, raising the air pressure in the vocal tract. This air pressure is used in the same way as

pulmonic pressure to produce consonants; the mechanism is surprisingly powerful, and the intensity of the noise produced by ejectives tends to be stronger than one finds in pulmonic consonants. The IPA phonetic symbols for ejectives are made by adding an apostrophe to the corresponding pulmonic symbol, so an ejective bilabial plosive is symbolised as **pʼ**, ejective velar plosive is **kʼ** and so on. Ejective plosives are found contrasting with pulmonic plosives in many languages in different parts of the world. Much less frequently we find ejective fricatives (e.g. Amharic **sʼ**). In English we find ejective allophones of **p**, **t**, **k** in some accents of the Midlands and North of England, usually at the end of a word preceding a pause: in utterances like 'On the top', 'That's right' or 'On your bike', it is often possible to hear a glottal closure just before the final consonant begins, followed by a sharp plosive release.

START ~ *INDEX*

elision

Some of the sounds that are heard if words are pronounced slowly and clearly appear not to be pronounced when the same words are produced in a rapid, colloquial style, or when the words occur in a different context; these "missing sounds" are said to have been elided. It is easy to find examples of elision, but very difficult to state rules that govern which sounds may be elided and which may not. Elision of vowels in English usually happens when a short, unstressed vowel occurs between voiceless consonants, e.g. in the first syllable of 'perhaps', 'potato', the second syllable of 'bicycle', or the third syllable of 'philosophy'. In some cases we find a weak voiceless sound in place of the normally voiced vowel that would have been expected. Elision also occurs when a vowel occurs between an obstruent consonant and a sonorant consonant such as a nasal or a lateral: this process leads to syllabic consonants, as in 'sudden' **sʌdn̩**, 'awful' **ɔːfl̩** (where a vowel is only heard in the second syllable in slow, careful speech).

Elision of consonants in English happens most commonly when a speaker "simplifies" a complex consonant cluster: 'acts' becomes **æks** rather than **ækts**, 'twelfth night' becomes **twelθnaɪt** or **twelfnaɪt** rather than **twelfθnaɪt**. It seems much less likely that any of the other consonants could be left out: the **l** and the **n** seem to be unelidable.

It is very important to note that sounds do not simply "disappear" like a light being switched off. A transcription such as **æks** for 'acts' implies that the **t** phoneme has dropped out altogether, but detailed examination of speech shows that such effects are more gradual: in slow speech the **t** may be fully pronounced, with an audible transition from the preceding **k** and to the following **s**, while in a more rapid style it may be articulated but not given any audible realisation, and in very rapid speech it may be observable, if at all, only as a rather early movement of the tongue blade towards the **s** position. Much more research in this area is needed (not only on English) for us to understand what processes are involved when speech is "reduced" in rapid articulation.

elocution

This is the traditional name for teaching "correct speech" to native speakers. It is rather surprising that phoneticians generally have no hesitation in telling foreign learners how they should pronounce the language they are learning, but are reluctant to advise native speakers on how to acquire a different accent or speaking style (apart, perhaps, from the "dialect coaching" given to actors). The training given by Professor Higgins to Eliza in *Pygmalion* and *My Fair Lady* is an example of elocution. Though this is nowadays scorned as something that belongs only in expensive private schools for upper-class girls, it has a respectable ancestry that goes back to the Greek teachers of rhetoric over two thousand years ago. It does not seem sensible to assume that everyone knows how to speak their native language with full clarity and intelligibility.

There has been considerable controversy in recent years over whether children should be taught in school how to speak with a "better" accent; while most people would agree that this sounds like an unwelcome attempt to level out accent differences in the community and to make most children feel that their version of the language is inferior to some arbitrary standard, it is also true that some of the more extreme statements on the subject have claimed that children's speech should be left untouched even if as a result the child will have problems in communicating outside its local environment, and may experience difficulty in getting a job on leaving school.

epenthesis

When a speaker inserts a redundant sound in a sequence of phonemes, that process is known as epenthesis; redundant in this context means that the additional sound is unnecessary, in that it adds nothing to the information contained in the other sounds. It happens most often when a word of one language is adopted into another language whose rules of phonotactics do not allow a particular sequence of sounds, or when a speaker is speaking a foreign language which is phonotactically different.

As an example of the first, we can look at examples where English words (which often have clusters of several consonants) are adopted by languages with a much simpler syllable structure: Japanese, for example, with a basic consonant-vowel syllable structure, tends to change the English word 'biscuit' to something like **bisuketo**.

Consonant epenthesis is also possible, and in BBC Pronunciation it quite frequently happens that in final nasal plus voiceless fricative clusters an epenthetic voiceless plosive is pronounced, so that the word 'French', phonemically **frenʃ**, is pronounced

as **frentʃ**. Such speakers lose the distinction between minimal pairs such as 'mince' **mɪns** and 'mints' **mɪnts**, pronouncing both words as **mɪnts**.

START ~ *INDEX*

Estuary English

Many learners of English have been given the impression that Estuary English is a new accent of English. In reality, there is no such accent, and the term should be used with care. The idea originates from the sociolinguistic observation that some people in public life who would previously have been expected to speak with a BBC (or RP) accent now find it acceptable to speak with some characteristics of the accents of the London area (the Estuary referred to is the Thames estuary), such as glottal stops, which would in earlier times have caused comment or disapproval.

START ~ *INDEX*

experimental phonetics

Quite a lot of the work done in phonetics is *descriptive* (providing an account of how different languages and accents are pronounced), and some is *prescriptive* (stating how they ought to be pronounced). But an increasing amount of phonetic research is *experimental*, aimed at the development and scientific testing of hypotheses. Experimental phonetics is quantitative (based on numerical measurement). It makes use of controlled experiments, which means that the experimenter has to make sure that the results could only be caused by the factor being investigated and not by some other. For example, in an experimental test of listeners' responses to intonation patterns produced by a speaker, if the listeners could see the speaker's face as the items were being produced it would be likely that their judgements of the intonation would be influenced by the facial expressions produced by the speaker rather than (or as well as) by the pitch variations. This would therefore not be a properly controlled experiment.

Experimental research is carried out in all fields of phonetics: in the articulatory field, we measure and study how speech is produced, in the acoustic field we examine the relationship between articulation and the resulting acoustic signal, and look at physical properties of speech sounds in general, while in the auditory field we do perceptual tests to discover how the listener's ear and brain interpret the information in the speech signal.

The great majority of experimental research makes use of instrumental phonetic techniques and laboratory facilities, though in principle it is possible to carry out reasonably well controlled experiments with no instruments. A classic example is Labov's study of the pronunciation of **r** in the words 'fourth floor' in New York department stores of different levels of prestige, a piece of low-cost research that

required only a notebook and pencil. This should be compulsory reading for anyone applying for a large research grant.

START ~ *INDEX*

falsetto

Many terms to do with speech prosody are taken from musical terminology, and falsetto is a singing term for a particular voice quality. It is almost always attributed to adult male voices, and is usually associated with very high pitch and a rather "thin" quality; it is sometimes encountered when a man tries to speak like a boy, or like a woman. Yodelling is a rapid alternation between falsetto and normal voice. Its linguistic role seems to be slight: an excursion into falsetto can be an indication of surprise or disbelief.

START ~ *INDEX*

feature

When the idea of the phoneme was new it was felt that phonemes were the ultimate constituents of language, the smallest element that it could be broken down into. But at roughly the same time as the atom was being split, phonologists pointed out that phonemes could be broken down into smaller constituents called features. All consonants, for example, share the feature Consonantal, which is not possessed by vowels. Some consonants have the feature Voice, while voiceless consonants do not. It is conventional to treat feature labels as being capable of having differing values – usually they are either "plus" (+) or "minus" (-), so we can say that a voiceless consonant is +Consonantal and -Voice while a vowel is -Consonantal and +Voice. The features are the things that distinguish each phoneme of a language from every other phoneme of that language; it follows that there will be a minimum number of features needed to distinguish them in this way, and that each phoneme must have a set of + and - values that is different from that of any other phoneme. For most languages, around twelve features are said to be sufficient (though in mathematical terms the theoretical minimum number can be calculated as follows: a set of n features will produce 2^n distinctions, so twelve features potentially allow for 2^{12} – i.e. 4096 – distinctions).

Features are used more in phonology than in phonetics, and in this use are normally called distinctive features; features are also used in some phonetic descriptions of the sounds of languages, and for these purposes the features have to indicate much more precise phonetic detail. For phonological purposes it is generally felt that the phonetic aspect of the labels needs to be only roughly right. A full feature-based analysis of a sound system is a long and complex task, and many theoretical problems arise in carrying it out.

START ~ *INDEX*

feedback

The process of speech production is controlled by the brain, and the brain seems to require information in the form of feedback about how the process is going. This can be in the form of tactile feedback, where the brain receives information about surfaces in the mouth being touched (e.g. contact between tongue and palate, or lip against lip): a pain-killing injection at the dentist's disables this feedback temporarily, often with adverse effects on speech production. There is also kinaesthetic feedback, where the brain receives information about movements in muscles and joints. Finally, there is auditory feedback, where information about the sounds produced is picked up either from sound waves outside the head, or from inside the head through "bone conduction"; experiments have shown that if this feedback is interfered with in some way, serious problems can result. In a noisy environment speakers adjust the level of their speech to compensate for the diminished feedback (this is known as the Lombard effect), while if the auditory feedback is experimentally delayed by a small fraction of a second it can have a devastating effect on speech, reducing many speakers to acute stuttering (this is known as the Delayed Auditory Feedback, or DAF, effect).

In a rather different sense, feedback also plays a vital role in dialogue: speakers do not usually like to speak without getting some idea of whether their audience is taking in what is being said (talking for an hour in a lecture without any response from those present is very daunting). In dialogue it is normal for the listener to respond helpfully.

START ~ *INDEX*

final lengthening

Instrumental studies of duration in speech show that there is a strong tendency in speakers of all languages to lengthen the last syllable or two before a pause or break in the rhythm, to such an extent that final syllables have to be excluded from the calculation of average syllable durations in order to avoid distorting the figures. Presumably this lengthening is noticeable perceptually and plays a role in helping the listener to anticipate the end of an utterance.

START ~ *INDEX*

flap

This is a type of consonant sound that is closely similar to the tap; it is usually voiced, and is produced by slightly curling back the tip of the tongue, then throwing it forward and allowing it to strike the alveolar ridge as it descends. The phonetic symbol for this sound is ɽ; it is most commonly heard in languages which have retroflex consonants, such as languages of the Indian sub-continent; it is also heard in the English of native speakers of such languages, often as a realisation of **r**. In

American English a flap is sometimes heard in words like ‘party’, ‘birdie’ , where the **r** consonant causes retroflexion of the tongue and the stress pattern favours a flap-type articulation.

START ~ *INDEX*

foot

The foot is a unit of rhythm. It has been used for a long time in the study of verse metre, where lines may be divided into sections based on patterns on strong and weak syllables. It is rather more controversial to suggest that normal speech is also structured in terms of regularly repeated patterns of syllables, but this is a claim that has been quite widely accepted for English. The suggested form of the English foot is that each foot consists of one stressed syllable plus any unstressed syllables that follow it; the next foot begins when another stressed syllable is produced. The sentence ‘Here is the news at nine o’clock’ could be analysed into feet in the following way (stressed syllables underlined, foot divisions marked with vertical lines):

|here is the |news at |nine o |clock

It is claimed that English feet tend to be of equal length, or isochronous, so that in feet consisting of several syllables there has to be compression of the syllables in order to maintain the stress-timed rhythm. There are many problems with this theory, as one discovers in trying to apply it to natural conversational speech, but the foot has been adopted as a central part of metrical phonology.

START ~ *INDEX*

formant

When speech is analysed acoustically we examine the spectrum of individual speech sounds by seeing how much energy is present at different frequencies. Most sounds (particularly voiced ones like vowels) exhibit peaks of energy in their spectrum at particular frequencies which contribute to the perceived quality of the sound rather as the notes in a musical chord contribute to the quality of that chord. These peaks are called formants, and it is usual to number them from the lowest to the highest; their frequency is usually specified in Hertz (meaning cycles per second, and abbreviated Hz). For example, typical values for the first two formants of the **ɜː** vowel in English ‘bird’ would be 650 Hz for Formant 1 and 1593 Hz for Formant 2. These are values for an adult female voice; typical adult male values are 513 Hz for F1 and 1377 Hz for F2.

START ~ *INDEX*

fortis

It is claimed that in some languages (including English) there are pairs of consonants whose members can be distinguished from each other in terms of whether they are "strong" (fortis) or "weak" (lenis). These terms refer to the amount of energy used in their production, and are similar to the terms tense and lax more usually used in relation to vowels. The fortis/lenis distinction does not (in English, at least) cut across any other distinction, but rather it duplicates the voiceless/voiced distinction. It is argued that English b, d, g, v, ð, z, ʒ often have little or no voicing in normal speech, and it is therefore a misnomer to call them voiced; since they seem to be more weakly articulated than p, t, k, f, θ, s, ʃ it would be appropriate to use the term lenis (meaning "weak") instead. Counter-arguments to this include the following: the term voiced could be used with the understood meaning that sounds with this label have the *potential* to receive voicing in appropriate contexts even if they sometimes do not receive it; no-one has yet provided a satisfactory way of measuring strength of articulation that could be used to establish that there is actually such a physical distinction in English; and it is, in any case, confusing and unnecessary to use Latin adjectives when there are so many suitable English ones.

START ~ *INDEX*

free variation

If two sounds that are different from each other can occur in the same phonological context and one of those sounds may be substituted for the other, they are said to be in free variation. A good example in English is that of the various possible realisations of the r phoneme: in different accents and styles of speaking we find the post-alveolar approximant ɹ which is the most common pronunciation in contemporary BBC pronunciation and General American, the tap ɾ which was typical of carefully spoken BBC pronunciation of fifty years ago, the labiodental approximant ʋ used by speakers who have difficulty in articulating tongue-tip versions of the r phoneme and by some older upper-class English speakers, the trilled r found in carefully-pronounced Scots accents and the uvular ʁ of the old traditional form of the Geordie accent on Tyneside. Although each of these is instantly recognisable as different from the others, the substitution of one of these for another would be most unlikely to cause an English listener to hear a sound other than the r phoneme. These different allophones of r are, then, in free variation. However, it is important to remember that the word "free" does not mean "random" in this context – it is very hard to find examples where a speaker will pronounce alternative allophones in an unpredictable way, since even if that speaker always uses the same accent, she or he will be monitoring the appropriateness of their style of speaking for the social context.

START ~ *INDEX*

fricative

This type of consonant is made by forcing air though a narrow gap so that a hissing noise is generated. This may be accompanied by voicing (in which case the sound is a voiced fricative, such as **z**, or it may be voiceless (e.g. **s**). The quality and intensity of fricative sounds varies greatly, but all are acoustically composed of energy at relatively high frequency – an indication of this is that much of the fricative sound is too high to be transmitted over a phone (which usually cuts out the highest and lowest frequencies in order to reduce the cost), giving rise to the confusions that often arise over sets of words like English 'fin', 'thin', 'sin' and 'shin'. In order for the sound quality to be produced accurately the size and direction of the jet of air has to be very precisely controlled; while this is normally something we do without thinking about it, it is noticeable that fricatives are what cause most difficulty to speakers who are getting used to wearing false teeth.

A distinction is sometimes made between sibilant or strident fricatives (such as **s**, **ʃ**) which are strong and clearly audible and others which are weak and less audible (such as **θ**, **f**). BBC pronunciation has nine fricative phonemes: **f**, **θ**, **s**, **ʃ**, **h** (voiceless) and **v**, **ð**, **z**, **ʒ** (voiced).

START ~ *INDEX*

front

One of the most important articulatory features of a vowel is determined by which part of the tongue is raised nearest to the palate. If it is the front of the tongue the vowel is classed as a front vowel: front vowels include **i**, **e**, **ɛ**, **a** (unrounded) and **y**, **ø**, **œ**, **ɶ** (rounded).

START ~ *INDEX*

function word

The notion of the function word belongs to grammar, not to phonetics, but it is a vital one in the description of English pronunciation. This class of words is distinguished from "lexical words" such as verbs, nouns, adjectives and adverbs, though it is difficult to be precise about how the distinction is to be defined. Function words include such types as conjunctions (e.g. 'and', 'but'), articles ('a/an', 'the') and prepositions (e.g. 'to', 'from', 'for', 'on'. Many function words have the characteristic that they are pronounced sometimes in a strong form (as when the word is pronounced in isolation) and at other times in a weak form (when pronounced in context, without stress); for example, the word 'and' is pronounced **ænd** in isolation (strong form) but as **ən** or **n̩** (weak form) in a context such as 'come and see', 'fish and chips'.

START ~ *INDEX*

fundamental frequency (F0)

When voicing is produced, the vocal folds vibrate; since vibration is an activity in which a movement happens repeatedly, it is possible in principle to count how many times per second (or other unit of time) one cycle of vibration occurs; if we do this, we can state the frequency of the vibration. In adult female voices the frequency of vibration tends to be around 200 or 250 cycles per second, and in adult males the frequency is about half of this. It is usual to express the number of cycles per second as Hertz (abbreviated Hz), so a frequency of 100 cycles per second is a frequency of 100 Hz.

Why "fundamental"? The answer is that all speech sounds are complex sounds made up of energy at many different component frequencies (unlike a "pure tone" such as an electronic whistling sound); when a sound is voiced, the lowest frequency component is always that of the vocal fold vibration – all other components are higher. So the vocal fold vibration produces the fundamental frequency. See also pitch.

START ~ *INDEX*

geminate

When two identical sounds are pronounced next to each other (e.g. the sequence of two **n** sounds in English 'unknown' **ʌnnəʊn**) they are referred to as geminate. Many languages have geminates occurring regularly. The problem with the notion of gemination is that there is often no way of discerning a physical boundary between the two paired sounds – more often, one simple hears a sound with greater length than the usual single consonant. In the case of long affricates (as found, for example, in Hindi), the gemination involves only the silent interval of the plosive part, and the fricative part is the same as the single consonant. Long vowels are not always treated as geminates: in the case of English (BBC accent) it is more common to describe the phonemic system as having phonemically long and phonemically short single vowels.

START ~ *INDEX*

General American

Often abbreviated as GA, this accent is usually held to be the "standard" accent of American English; it is interesting to note that the standard that was for a long time used in the description of British English pronunciation (Received Pronunciation, or RP) is only spoken by a small minority of the British population, whereas GA is the accent of the majority of Americans. It is traditionally identified as the accent spoken throughout the USA except in the north-east (roughly the Boston and New England area) and the south-eastern states. Since it is widely used in broadcasting it is also known as "Network English".

generative phonology

A major change in the theory of phonology came about in the 1960s when many people became convinced that important facts about the sound systems of languages were being missed by phonologists who concentrated solely on the identification of phonemes and the analysis of relationships between them. Work by Morris Halle, later joined by Noam Chomsky, showed that there were many sound processes which, while they are observable in the phonology, are actually regulated by grammar and morphology. For example, the following pairs of English diphthongs and vowels had previously been regarded as unrelated: **aɪ** and **ɪ**; **iː** and **e**; **eɪ** and **æ**; however, in word-pairs such as 'divine' **dɪvaɪn** and 'divinity' **dɪvɪnəti**, 'serene' **səriːn** and 'serenity' **sərenəti** and 'profane' **prəfeɪn** and 'profanity' **prəfænəti** there are "alternations" that form part of what native speakers know about their language. Similarly, traditional phoneme theory would see no relationship between **k** and **s**, yet there is a regular alternation between the two in pairs such as 'electric' **ɪlektrɪk** – 'electricity' **ɪlektrɪsəti** or 'toxic' **tɒksɪk** – 'toxicity' **tɒksɪsəti**. It was claimed that beneath the physically observable ("surface") string of sounds that we hear there is a more abstract, unobservable "underlying" phonological form.

If such alternations are accepted as a proper part of phonology, it becomes necessary to write rules that state how they work: these rules must regulate such changes as substitutions, deletions and insertions of sounds in specific contexts, and an elaborate method of writing these rules in an algebra-like style was evolved: this can be seen in the best known generative phonological treatment of English, *The Sound Pattern of English* (Chomsky and Halle, 1968). This type of phonology became extremely complex; it has now been largely replaced by newer approaches to phonology, many of which, despite rejecting the theory of the *Sound Pattern of English*, are still classed as generative since they are based on the principle of an abstract, underlying phonological representation of speech which needs rules to convert it into phonetic realisations.

glide

We think of speech in terms of individual speech sounds such as phonemes, and it is all too easy to assume that they have clear boundaries between them like letters on a printed page. Sometimes in speech we can find clear boundaries between sounds, and in others we can make intelligent guesses at the boundaries though these are difficult to identify; in other cases, however, it is clear that a more or less gradual glide from one quality to another is an essential part of a particular sound. An obvious case is that of diphthongs: in their case the glide is comparatively slow. Some sounds which are usually classed as consonants also involve glides: these include "semivowels";

some modern works on phonetics and phonology also class the glottal fricative **h** and the glottal stop **ʔ** as glides. This is a perplexing and almost contradictory use of the word "glide", especially in the latter case.

START ~ *INDEX*

glottalic

This adjective could be used to refer to anything pertaining to the glottis, but it is generally used to name a type of airstream. A glottalic airstream is produced by making a tight closure of the vocal folds and then moving the larynx up or down: raising the larynx pushes air outwards causing an egressive glottalic airstream while lowering the larynx pulls air into the vocal tract and is called an ingressive glottalic airstream. Sounds of this type found in human language are called ejective or implosive respectively.

START ~ *INDEX*

glottal stop, glottalisation

One of the functions of a closure of the vocal folds is to produce a consonant. In a true glottal stop there is complete obstruction to the passage of air, and the result is a period of silence. The phonetic symbol for a glottal stop is **ʔ**. In casual speech it often happens that a speaker aims to produce a complete glottal stop but instead makes a low-pitched creak-like sound. Glottal stops are found as consonant phonemes in some languages (e.g. Arabic); elsewhere they are used to mark the beginning of a word if the first phoneme in that word is a vowel (this is found in German). Glottal stops are found in many accents of English: sometimes a glottal stop is pronounced in front of a **p**, **t** or **k** if there is not a vowel immediately following (e.g. 'captive' **kæʔptɪv**, 'catkin' **kæʔtkɪn**, 'arctic' **ɑːʔktɪk**); a similar case is that of **tʃ** when following a stressed vowel (or when syllable-final), as in 'butcher' **bʊtʃə**. This addition of a glottal stop is sometimes called glottalisation or glottal reinforcement. In some accents, the glottal stop actually replaces the voiceless alveolar plosive **t** as the realisation of the **t** phoneme when it follows a stressed vowel, so that 'getting better' is pronounced **geʔɪŋ beʔə** – this is found in many urban accents, notably London (Cockney), Leeds, Glasgow, Edinburgh and others, and is increasingly accepted among relatively highly-educated young people.

START ~ *INDEX*

glottis

The glottis is the opening between the vocal folds. Like the child who asked "where does your lap go when you stand up?", one may imagine that the glottis disappears when the vocal folds are pressed together, but in fact it is usual to refer to the "closed glottis" in this case. Apart from the fully closed state, the vocal folds may be put in the position appropriate for voicing, with narrowed glottis; the glottis may be narrowed but less so than for voicing – this is appropriate for whisper and for the production of the glottal fricative **h**, while it tends to be more open for voiceless consonants. For normal breathing the glottis is quite wide, usually being wider for breathing in than for breathing out. When producing aspirated voiceless plosive consonants, it is usual to find a momentary very wide opening of the glottis just before the release of the plosive.

For more information and diagrams, see *English Phonetics and Phonology*, Section 4.1.

START ~ *INDEX*

groove

The tongue may make contact with the upper surface of the mouth in a number of different places, and we also know that it may adopt a number of different shapes as viewed from the side. However, we tend to neglect another aspect of tongue control: its shape as viewed from the front. Variation of this sort is most clearly observed in fricatives: it is claimed that in the production of the English **s** sound, the tongue has a deep but narrow groove running from front to back, while **ʃ** has a wide, shallow slit. Experimental support for this claim is, however, not very strong.

START ~ *INDEX*

guttural

This adjective is little used in phonetics these days, though it was included among the "places of articulation" on the IPA chart until 1912, after which it was replaced by the modern term uvular. The word "guttural" tends to be used by English-speaking non-specialists to characterise languages which have noticeable "back-of-the-mouth" consonants (e.g. German, Arabic); used in this way the word has a rather pejorative feel about it.

START ~ *INDEX*

head

In the standard British treatment of intonation, the head is one of the components of the tone-unit; if one or more stressed syllables precedes the tonic syllable (nucleus), the head comprises all syllables from the first stressed syllable up to (but not including) the tonic. Here are some examples:

'here is the 'six oclock \news

¦-------------------------¦

HEAD

'passengers are re'quested to 'fasten their \seat belt

¦---¦

HEAD

If there are unstressed syllables preceding the head, or if there are no stressed syllables before the head but there are some unstressed ones, these unstressed syllables constitute a pre-head.

START ~ *INDEX*

hesitation

We pause in speaking for many reasons, and pauses have been studied intensively by psycholinguists. Some pauses are intentional, either to create an effect or to signal a major syntactic or semantic boundary; but hesitation is generally understood to be involuntary, and often due to the need to plan what the speaker is going to say next. Hesitations are also often the result of difficulty in recalling a word or expression. Phonetically, hesitations and pauses may be silent or may be filled by voiced sound: different languages and cultures have very different hesitation sounds. BBC pronunciation tends to use **ɜː** or **ɜːm**.

START ~ *INDEX*

Higgins, Henry

Henry Higgins is the best-known fictional phonetician, the central male character of Shaw's *Pygmalion* and of the musical *My Fair Lady*. Higgins is given more extreme views about the importance of correct pronunciation in the latter, and most phoneticians are rather embarrassed at the idea that the general public might think of their subject as being capable of being used in the way Higgins used it. Phoneticians like to guess at who the real-life original of Higgins was: it used to be widely thought that this was the great phonetician Henry Sweet, but there is evidence to suggest that Shaw probably had his own contemporary, Daniel Jones, in mind. There is, of course, no reason why Shaw should not have had *both* men in mind.

You can read about the question of Jones being the model for Higgins in *The Real Professor Higgins*, by B. Collins and I. Mees (Mouton, 1999).

START ~ *INDEX*

hoarse(ness)

In informal usage, hoarseness is generally used to refer to phonation (voicing) that is irregular because of illness or extreme emotion.

START ~ *INDEX*

homophone

If two different words are pronounced identically, they are homophones. In many cases they will be spelt differently (e.g. 'saw' – 'sore' – 'soar' in BBC pronunciation), but homophony is possible also in the case of pairs like 'bear' (verb) and 'bear' (noun) which are spelt the same.

START ~ *INDEX*

homorganic

When two sounds have the same place of articulation they are said to be homorganic. This notion is rather a relative one: it is clear that **p** and **b** are homorganic, and most people would agree that **t** and **s** are too. But **t** and **ʃ** in the affricate **tʃ** are usually also said to be homorganic despite the fact that the latter sound is usually described as post-alveolar; the **t** is often articulated nearer to the palatal region than its usual place, but it is not certain to be in the same place of articulation as the **ʃ**.

START ~ *INDEX*

implosive

Several different types of speech sound can be made by drawing air into the body rather than by expelling it in the usual way. In an implosive this is done by bringing the vocal folds together and then drawing the larynx downwards to suck air in; this is usually done in combination with the plosive manner of articulation. Most of the implosives found functioning as speech sounds are voiced, which seems surprising since if the glottis is closed it should not be possible for the vocal folds to vibrate: it appears that while the vocal folds are mostly pressed together firmly, a part of their length is allowed to vibrate as a result of a small amount of air passing between the folds while the larynx is lowered. This produces a surprisingly strong voicing sound. Implosive consonant phonemes are found in a number of languages, in Africa (e.g. Igbo) and also in India (e.g. Sindhi). The phonetic symbols for implosives are ɓ, ɗ, ɠ.

START ~ *INDEX*

ingressive

All speech sounds require some movement of air; almost always when we speak, the air is moving outwards – there is an egressive airflow. In rare cases, however, the airflow is inwards (ingressive). It is possible to speak while drawing air into the lungs: we may do this when out of breath, or coughing badly; children do it to be silly. It has been reported that some societies regularly use this style of speaking when it is customary to disguise the speaker's identity. We also find ingressive airflow created by the larynx (see glottalic, implosive) or by the tongue (see click).

START ~ *INDEX*

instrumental phonetics

The field of phonetics can be divided up into a number of sub-fields, and the term 'instrumental' is used to refer to the analysis of speech by means of instruments; this may be acoustic (the study of the vibration in the air caused by speech sounds) or articulatory (the study of the movements of the articulators which produce speech sounds). Instrumental phonetics is a quantitative approach – it attempts to characterise speech in terms of measurements and numbers, rather than by relying on listeners' impressions.

Many different instruments have been devised for the study of speech sounds. The best known technique for acoustic analysis is spectrography, in which a computer produces a "picture" of speech sounds. Such computer systems can usually also carry out the analysis of fundamental frequency for producing "pitch displays". For analysis of articulatory activity there are many instrumental techniques in use, including radiography (X-rays) for examining activity inside the vocal tract, laryngoscopy for inspecting the inside of the larynx, palatography for recording patterns of contact between tongue and palate, glottography for studying the vibration of the vocal folds

and many others. Measurement of airflow from the vocal tract and of air pressure within it also give us a valuable indirect picture of other aspects of articulation.

Instrumental techniques are usually used in experimental phonetics, but this does not mean that all instrumental studies are experimental: when a theory or hypothesis is being tested under controlled conditions the research is experimental, but if one simply makes a collection of measurements using instruments this is not the case.

START ~ *INDEX*

intensity

Intensity is a physical property of sounds, and is dependent on the amount of energy present. Perceptually, there is a fairly close relationship between physical intensity and perceived loudness. The intensity of a sound depends both on the amplitude of the sound wave and on its frequency.

START ~ *INDEX*

interdental

For most purposes in general phonetics it is felt sufficient to describe articulations involving contact between the tongue and the front teeth as 'dental'; however, in some cases it is necessary to be more precise in one's labelling and indicate that the tip of the tongue is protruded between the teeth (interdental articulation). It is common to teach this articulation for θ and ð to learners of English who do not have a dental fricative in their native language, but it is comparatively rare to find interdental fricatives in native speakers of English (it is said to be typical of the Californian accent of American English, though I have never observed this myself); most English speakers produce θ and ð by placing the tip of the tongue against the back of the front teeth.

START ~ *INDEX*

International Phonetic Association and Alphabet

The International Phonetic Association was established in 1886 as a forum for teachers who were inspired by the idea of using phonetics to improve the teaching of the spoken language to foreign learners. As well as laying the foundations for the modern science of phonetics, the Association had a revolutionary impact on the language classroom in the early decades of its existence, where previously the concentration had been on proficiency in the written form of the language being learned. The Association is still a major international learned society, though the crusading spirit of the pronunciation teachers of the early part of the century is not so

evident nowadays. The Association only rarely holds official meetings, but contact among the members is maintained by the Association's Journal, which has been in publication more or less continuously since the foundation of the Association, with occasional changes of name.

Since its beginning, the Association has taken the responsibility for maintaining a standard set of phonetic symbols for use in practical phonetics, presented in the form of a chart (see the chart on p. xi of *English Phonetics and Phonology*, or find it on the IPA website referred to below). The set of symbols is usually known as the International Phonetic Alphabet (and the initials IPA are therefore ambiguous). The alphabet is revised from time to time to take account of new discoveries and changes in phonetic theory.

The website of the IPA is http://www2.arts.gla.ac.uk/IPA/ipa.html

START ~ *INDEX*

intonation

There is confusion about intonation caused by the fact that the word is used with two different meanings: in its more restricted sense, 'intonation' refers simply to the variations in the pitch of a speaker's voice used to convey or alter meaning, but in its broader and more popular sense it is used to cover much the same field as 'prosody', where variations in such things as voice quality, tempo and loudness are included. It is, regrettably, common to find in pronunciation teaching materials accounts of intonation that describe only pitch movements and levels, and then claim that a wide range of emotions and attitudes are signalled by means of these pitch phenomena. There is in fact very little evidence that pitch movements alone are effective in doing signalling of this type.

It is certainly possible to analyse pitch movements (or their acoustic counterpart, fundamental frequency) and find regular patterns that can be described and tabulated. Many attempts have been made at establishing descriptive frameworks for stating these regularities. Some analysts look for an underlying basic pitch melody (or for a small number of them) and then describe the factors that cause deviations from these basic melodies; others have tried to break down pitch patterns into small constituent units such as "pitch phonemes" and "pitch morphemes", while the approach most widely used in Britain takes the tone unit as its basic unit and looks at the different pitch possibilities of the various components of the tone unit (the pre-head, head, tonic syllable/nucleus and tail).

As mentioned above, intonation is said to convey emotions and attitudes. Other linguistic functions have also been claimed: interesting relationships exist in English between intonation and grammar, for example: in a few extreme cases a perceived difference in grammatical meaning may depend on the pitch movement, as in the following example:

She 'didnt 'go be'cause of her \/timetable (meaning “she did go, but it was not because of her timetable”)

and

She 'didnt /go ¦ be,cause of her \timetable (meaning “she didn’t go, the reason being her timetable”).

Other “meanings” of intonation include things like the difference between statement and question; the contrast between “open” and “closed” lists, where

'would you like /wine, /sherry or /beer

is “open”, implying that other things are also on offer, while

'would you like /wine, /sherry or \beer

is “closed”, no further choices being available); and the indication of whether a relative clause is restrictive or non-restrictive, as in, for example,

the 'car which had 'bad 'brakes \crashed

compared with

the \/car ¦ which had 'bad \/brakes ¦ \crashed

Another approach to intonation is to concentrate on its role in conversational discourse: this involves such aspects as indicating whether the particular thing being said constitutes new information or old, the regulation of turn-taking in conversation, the establishment of dominance and the elicitation of co-operative responses. As with the signalling of attitudes, it seems that though analysts concentrate on pitch movements there are many other prosodic factors being used to create these effects.

Much less work has been done on the intonation of languages other than English. It seems that all languages have something that can be identified as intonation; there appear to be many differences between languages, but one suspects, on reading the literature, that this is due more to the different descriptive frameworks used by different analysts than to inter-language differences. It is claimed that tone languages also have intonation, which is superimposed upon the tones themselves, and this creates especially difficult problems of analysis.

Chapters 15–19 of *English Phonetics and Phonology* deal with intonation.

START ~ *INDEX*

isochrony

Isochrony is the property of being equally spaced in time, and is usually used in connection with the description of the rhythm of languages. English rhythm is said to exhibit isochrony because it is believed that it tends to preserve equal intervals of time between stressed syllables irrespective of the number of syllables that come between them. For example, if the following sentence were said with isochronous stresses, the four syllables 'both of them are' would take the same amount of time as 'new' and 'here':

'both of them are 'new 'here

This kind of timing is also known as stress-timed rhythm and is based on the notion of the foot. Experimental research suggests that isochrony is rarely found in natural speech, and that (at least in the case of English speakers) the brain judges sequences of stresses to be more nearly isochronous than they really are: the effect is to some extent an illusion.

The notion of isochrony does not necessarily have to be restricted to the intervals between stressed syllables. It is possible to claim that some languages tend to preserve a constant quantity for all syllables in an utterance: this is said to result in a syllable-timed rhythm. French, Spanish and Japanese have been claimed to be of this type, though laboratory studies do not give this claim much support.

It seems that in languages characterised as stress-timed there is a tendency for unstressed syllables to become weak, and to contain short, centralised vowels, whereas in languages described as syllable-timed unstressed vowels tend to retain the quality and quantity found in their stressed counterparts.

See Section 14.1 of *English Phonetics and Phonology*.

START ~ *INDEX*

Jones, Daniel

Jones was, with the possible exception of Henry Sweet, the most influential figure in the development of present-day phonetics in Britain. He was born in 1881 and died in 1967; he was for many years Professor of Phonetics at University College London. He worked on many of the world's languages and on the theory of the phoneme and of phonetics, but is probably best remembered internationally for his works on the phonetics of English, particularly his *Outline of English Phonetics* and *English Pronouncing Dictionary*. It has been suggested that he was the model for Shaw's Professor Henry Higgins.

START ~ *INDEX*

juncture

It is often necessary in describing pronunciation to specify how closely attached one sound is to its neighbours: for example, **k** and **t** are more closely linked in the word 'acting' than in 'black tie', and **t** and **r** are more closely linked in 'nitrate' than in 'night rate'. Sometimes there are clearly observable phonetic differences in such examples: in comparing 'cart rack' with 'car track' we notice that the vowel in 'cart' is short (being shortened by the **t** that follows it) while the same phoneme in 'car' is longer, and the **r** in 'track' is devoiced (because it closely follows **t**) while **r** in 'rack' is voiced.

It seems natural to explain these relationships in terms of the placement of word boundaries, and in modern phonetics and phonology this is what is done; studies have also been made of the effects of sentence and clause boundaries. However, it used to be widely believed that phonological descriptions should not be based on a prior grammatical analysis, and the notion of juncture was established to overcome this restriction: where one found in continuous speech phonetic effects that would usually be found preceding or following a pause, the phonological element of juncture would be postulated. Using the symbol + to indicate this juncture, the transcription of 'car track' and 'cart rack' would be **kɑː + træk** and **kɑː t + ræk**. There was at one time discussion of whether spaces between words should be abolished in the phonetic transcription of connected speech except where there was an observable silence; juncture symbols could have replaced spaces where there was phonetic evidence for them.

Since the position of juncture (or word boundary) can cause a perceptual difference, and therefore potential misunderstanding, it is usually recommended that learners of English should practise making and recognising such differences, using pairs like 'pea stalks/peace talks' and 'great ape/grey tape'.

START ~ *INDEX*

key

Many analogies have been drawn between music and speech, and many concepts from musical theory have been adopted for the analysis of speech prosody; the use of the word "key" is perhaps one of the less appropriate adoptions. In studying the use of pitch it is necessary to assume that each speaker has a range from the highest to the lowest pitch that they use in speaking: it is observable that these extremes are only rarely used and that in general we tend to speak well within the range defined by these extremes. It has, however, also been observed that we sometimes make more use of the higher or lower part of our pitch range than in normal speaking, usually as a result of the emotional content of what we are saying or because of a particular effect we wish to create for the listener; the terms "high key" and "low key" have been used to describe this. But whereas in music "key" refers to a specific configuration of notes based on one particular note within the octave, in the description of speech the word has generally been used simply to indicate a rough location within the pitch range,

while in one recent approach to intonation it has been used to specify the starting and ending points of pitch patterns whose range extends outside the most commonly used part of the pitch range.

START ~ *INDEX*

kinaesthetic/esia

(Sometimes spelt kinaesthetic.) When the brain instructs the body to produce some action or movement, it usually checks to see that the movement is carried out correctly. It is able to do this through receiving feedback through the nervous system. One form of feedback is auditory: we listen to the sounds we make, and if we are prevented from doing this (for example as a result of loud noise going on near us), our speech will not sound normal. But we also receive feedback about the movements themselves, from the muscles and the joints that are moved. This is kinaesthetic feedback, and normally we are not aware of it. However, a phonetics specialist must become conscious of kinaesthetic information: if you are learning to produce the sounds of an unfamiliar language, you must be aware of what you are doing with your articulators, and practical phonetic training aims to raise the learner's sensitivity to this feedback.

START ~ *INDEX*

labial(-ised)

This is a general label for articulations in which one or both of the lips are involved. It is usually necessary to be more specific: if a consonant is made with both lips, it is called bilabial (plosives and fricatives of this type are regularly encountered); if another articulator is brought into contact or near-contact with the lips, we use terms such as labiodental (lips and teeth) or linguo-labial (tongue and teeth).

Another use of the lips is to produce the effect of lip-rounding, and this is often called labialisation; the term is more often used in relation to consonants, since the term "rounded" tends to be used for vowels with rounded lips.

START ~ *INDEX*

labiodental

A consonant articulated with contact between one or both of the lips and the teeth is labiodental. By far the most common type of labiodental articulation is one where the lower lip touches the upper front teeth, as in the fricatives **f** and **v**. Labiodental plosives, nasals and approximants are also found

START ~ INDEX

laminal

This adjective is used to refer to articulations in which the tongue blade (the part of the tongue just further back than the tongue tip) is used. English alveolar consonants **t**, **d**, **n**, **s**, **z**, **l** are usually laminal.

START ~ INDEX

larynx

The larynx is a major component of our speech-producing equipment and has a number of different functions. It is located in the throat and its main biological function is to act as a valve that can stop air entering or escaping from the lungs and also (usually) prevents food and other solids from entering the lungs. It consists of a rigid framework or box made of cartilage and, inside, the vocal folds, which are two small lumps of muscular tissue like a very small pair of lips with the division between them (the glottis) running from front to back of the throat. There is a complex set of muscles inside the larynx that can open and close the vocal folds as well as changing their length and tension. See *English Phonetics and Phonology*, Section 4.1.

Loss of laryngeal function (usually through surgical laryngectomy) has a devastating effect on speech, but patients can learn to use substitute sources of voicing either from oesophageal air pressure ("belching") or from an electronic artificial voice source.

START ~ INDEX

lateral

A consonant is lateral if there is obstruction to the passage of air in the centre (mid-line) of the air-passage and the air flows to the side of the obstruction. In English the **l** phoneme is lateral both in its "clear" and its "dark" allophones: the blade of the tongue is in contact with the alveolar ridge as for a **t**, **d** or **n** but the sides of the tongue are lowered to allow the passage of air. When an alveolar plosive precedes a lateral consonant in English it is usual for it to be laterally released: this means that to go from **t** or **d** to **l** we simply lower the sides of the tongue to release the compressed air, rather than lowering and then raising the tongue blade.

Most laterals are produced with the air passage to both sides of the obstruction (they are bilateral), but sometimes we find air passing to one side only (unilateral). Other lateral consonants are found in other languages: the Welsh "ll" sound is a voiceless lateral fricative **ɬ**, and Xhosa and Zulu have a voiced lateral fricative **ɮ**; several

Southern African languages have lateral clicks (where the plosive occlusion is released laterally) and at least one language (of Papua New Guinea) has a contrast between alveolar and velar lateral. A bilabial lateral is an articulatory possibility but it seems not to be used in speech.

START ~ *INDEX*

lax

A lax sound is said to be one produced with relatively little articulatory energy. Since there is no established standard for measuring articulatory energy, this concept only has meaning if it is used in relation to some other sounds that are articulated with a comparatively greater amount of energy (the term tense is used for this). It is mainly American phonologists who use the terms lax and tense in describing English vowels: the short vowels ɪ, e, æ, ʌ, ɒ, ʊ, ə are classed as lax, while what are usually referred to as the long vowels and the diphthongs are tense. The terms can also be used of consonants as equivalent to fortis (tense) and lenis (lax), though this is not commonly done in present-day description.

START ~ *INDEX*

length

The scientific measure of the amount of time that an event takes is called duration; it is also important to study the time dimension from the point of view of what the listener hears – length is a term sometimes used in phonetics to refer to a subjective impression that is distinct from physically measurable duration. Usually, however, the term is used as if synonymous with duration. Length is important in many ways in speech: in English and most other languages, stressed syllables tend to be longer than unstressed. Some languages have phonemic differences between long and short sounds, and English is claimed by some writers to be of this type, contrasting short vowels ɪ, e, æ, ʌ, ɒ, ʊ, ə with long vowels iː, ɑː, ɔː, ɜː, uː (though other, equally valid analyses have been put forward). When languages have long/short consonant differences, as does Arabic, for example, it is usual to treat the long consonants as geminate; it is odd that this is not done equally regularly in the case of vowels.

Perhaps the most interesting example of length differences comes from Estonian, which has traditionally been said to have a three-way distinction between short, long and extra-long consonants and vowels.

START ~ *INDEX*

lenis

A lenis sound is a weakly articulated one (the word comes from Latin, where it means "smooth, gentle"). The opposite term is fortis. In general, the term lenis is used of voiced consonants (which are supposed to be less strongly articulated than voiceless ones), and is resorted to particularly for languages such as German, Russian and English where "voiced" phonemes like **b**, **d**, **g** are not always voiced.

START ~ INDEX

level (tone)

Many tone languages possess level tones; these are produced with an unchanging pitch level, and some languages have a number (some as many as four or five) of contrasting level tones. In the description of English intonation it is also necessary to recognise the existence of level tone: as a simple demonstration, consider various common one-syllable utterances such as 'well', 'yes', 'no', 'some'. Most English speakers seem to be able to recognise a level-tone pronunciation as something different from the various moving-tone possibilities such as fall, rise, fall–rise etc., and to ascribe some sort of meaning to it (usually with some feeling of boredom, hesitation or lack of surprise). It is probable that from the perceptual point of view a level tone is more closely related to a rising tone than to a falling one.

Level tone presents a problem in that the tones used in the intonation of a language like English are usually defined in terms of pitch movements, and there is no pitch movement on a level tone. It is therefore necessary to say, in identifying a syllable as carrying a level tone, that it has the prominence characteristic of the moving tones and occurs in a context where a tone would be expected to begin.

START ~ INDEX

lexicon/al

Traditionally, a lexicon is the same thing as a dictionary. In recent years, however, the word has been given a slightly different meaning for linguistic studies: it is used to refer to the total set of words that a speaker knows (i.e. has stored in her or his mind). The speaker's lexicon is, of course, much more than just a list of words: it is also a whole network of relationships between the words. There is much evidence to show that words are stored in the mind in a very complex way that enables us to recognise a word very quickly. One important but unanswered question is how alternative pronunciations are stored in the mind: do we keep a set of different ways of pronouncing a word like 'that' or 'there', or do we also have rules to specify how one form of the word may be changed into another?

START ~ INDEX

liaison

"Linking" or "joining together" of sounds is what this French word refers to. In general this is not something that speakers need to do anything active about – we produce the phonemes that belong to the words we are using in a more or less continuous stream, and the listener recognises them (or most of them) and receives the message. However, phoneticians have felt it necessary in some cases to draw attention to the way the end of one word is joined on to the beginning of the following word. In English the best-known case of liaison is the "linking r": there are many words in English (e.g. 'car', 'here', 'tyre') which in a rhotic accent such as General American or Scots would be pronounced with a final **r** but which in BBC pronunciation end in a vowel when they are pronounced before a pause or before a consonant. When they are followed by a vowel, BBC speakers pronounce **r** at the end (e.g. 'the car is' **ðə kɑːr ɪz**) – it is said that this is done to link the words without sliding the two vowels together (though it is difficult to see how such a statement could stand as an explanation of the phenomenon – lots of languages do run vowels together). Another aspect of liaison in English is the movement of a single consonant at the end of an unstressed word to the beginning of the next if that is strongly stressed: a well-known example is 'not at all', where the **t** of 'at' becomes initial (and therefore strongly aspirated) in the final syllable for many speakers.

START ~ *INDEX*

lingual

This is the adjective used of any articulation in which the tongue is involved.

START ~ *INDEX*

lips

The lips are extremely mobile and active articulators in speech. In addition to being used to make complete closure for **p**, **b**, **m** they can be brought into contact with the teeth or the tongue. The ring of muscles around the lips makes it possible for them to be rounded and protruded. They are so flexible that they can be used to produce a trill.

START ~ *INDEX*

liquid

This is an old-fashioned phonetic term that has managed to survive to the present day despite the lack of any scientific definition of it. Liquids are one type of approximant, which is a sound closely similar to vowels: some approximants are glides, in that they

involve a continuous movement from one sound quality to another (e.g. **j** in 'yet' and **w** in 'wet'). Liquids are different from glides in that they can be maintained as steady sounds – the English liquids are **r** and **l**.

START ~ INDEX

loudness

We have instrumental techniques for making scientific measurements of the amount of energy present in sounds, but we also need a word for the impression received by the human listener, and we use loudness for this. We all use greater loudness to overcome difficult communication conditions (for example, a bad telephone line) and to give strong emphasis to what we are saying, and it is clear that individuals differ from each other in the natural loudness level of their normal speaking voice. Loudness plays a relatively small role in the stressing of syllables, and it seems that in general we do not make very much linguistic use of loudness contrasts in speaking.

START ~ INDEX

low

The word low is used for two different purposes in phonetics: it is used to refer to low pitch (related to low fundamental frequency). In addition, it is used by some phoneticians as an alternative to open as a technical term for describing vowels (so that **a** and **ɑ** are low vowels).

START ~ INDEX

lungs

The biological function of the lungs is to absorb oxygen from air breathed in and to excrete carbon dioxide into the air breathed out. From the speech point of view, their major function is to provide the driving force that compresses the air we use for generating speech sounds. They are similar to large sponges, and their size and shape are determined by the rib cage that surrounds them, so that when the ribs are pressed down the lungs are compressed and when the ribs are lifted the lungs expand and fill with air. Although they hold a considerable amount of air (normally several litres, though this differs greatly between individuals) we use only a small proportion of their capacity when speaking – we would find it very tiring if we had to fill and empty the lungs as we spoke, and in fact it is impossible for us to empty our lungs completely.

START ~ INDEX

manner of articulation

One of the most important things that we need to know about a speech sound is what sort of obstruction it makes to the flow of air: a vowel makes very little obstruction, while a plosive consonant makes a total obstruction. The type of obstruction is known as the manner of articulation. Apart from vowels, we can identify a number of different manners of articulation, and the consonant chart of the International Phonetic Association classifies consonants according to their manner and their place of articulation.

START ~ *INDEX*

median

In the great majority of speech sounds the flow of air passes down the centre of the vocal tract (though in plosives there is a brief time when air does not flow at all). Some phoneticians feel we should have a technical term to characterise such sounds, and use median; however, since it is really only laterals like l that are *not* median, the term is only rarely needed.

START ~ *INDEX*

metrical phonology

This is a comparatively recent development in phonological theory, and is one of the approaches often described as "non-linear". It can be seen as a reaction against the overriding importance given to the phonemic segment in most earlier theories of phonology. In metrical phonology great importance is given to larger units and their relative strength and weakness; there is, for example, considerable interest in the structure of the syllable itself and in the patterns of strong and weak that one finds among neighbouring syllables and among the words to which the syllables belong. Another area of major interest is the rhythmical nature of speech and the structure of the foot: metrical phonology attempts to explain why shifts in word stress occur as a result of context, giving alternations like

thir'teen but 'thirteenth 'place

com'pact but 'compact 'disc

The metrical structure of an utterance is usually diagrammed in the form of a tree diagram (metrical trees), though for the purposes of explaining the different levels of stress found in an utterance more compact "metrical grids" can be constructed. This approach can be criticised for constructing very elaborate hypotheses with little empirical evidence, and for relying exclusively on a binary relationship between

elements where all polysyllabic sequences can be reduced to pairs of items of which one is strong and the other is weak.

You can read more in *English Phonetics and Phonology*, Section 14.1.

START ~ *INDEX*

minimal pair

In establishing the set of phonemes of a language, it is usual to demonstrate the independent, contrastive nature of a phoneme by citing pairs of words which differ in one sound only and have different meanings. Thus in BBC English 'fairy' **feəri** and 'fairly' **feəli** make a minimal pair and prove that **r** and **l** are separate, contrasting phonemes; the same cannot be done in, for example, Japanese since that language does not have distinct **r** and **l** phonemes.

START ~ *INDEX*

monophthong

This word, which refers to a single vowel, would be pretty meaningless on its own: it is used only in contrast with the word diphthong, which literally means a "double sound" in Ancient Greek. If we find a vowel that is not a diphthong, we can call it a monophthong.

START ~ *INDEX*

mora

This is a unit used in the study of quantity and rhythm in speech. In this study it is traditional to make use of the concept of the syllable. However, the syllable is made to play a lot of different roles in language description: in phonology we often use the syllable as the basic framework for describing how vowels and consonants can combine in a particular language, and most of the time it does not seem to matter that we use the same unit to be the thing that we count when we are looking for beats in verse or rhythmical speech. Traditionally, the syllable has also been viewed as an articulatory unit consisting (in its ideal form) of a movement from a relatively closed vocal tract to a relatively open vocal tract and back to a relatively closed one.

Not surprisingly, this multiple use of the syllable does not always work, and there are languages where we need to use different units for different purposes. In Japanese, for example, it is possible to construct syllables that are combinations of vowels and consonants: it is often pointed out that Japanese favours a CV (Consonant-Vowel)

syllable structure. Certainly we can divide Japanese speech into such syllables, but if Japanese speakers are asked to count the number of beats they hear in an utterance the answer is likely to be rather different from what an English speaker would expect: it appears that Japanese speakers count something other than phonological syllables. To English speakers, for example, the word 'Nippon' appears to have two beats, but for Japanese speakers it has four: the word is divided into units of time as follows:

ni | p | po | n

Since the term syllable is needed for other purposes, the term mora has been adopted for a unit of timing, so we can say that there are four morae in the word 'Nippon'.

START ~ *INDEX*

motor theory of speech perception

We still know little about how the brain recognises speech. Some researchers believe that in speech perception the brain makes use of knowledge about how speech sounds are made: for example, it is claimed that we hear very sharply defined differences between **b**, **d** and **g**, since each of these is produced by fundamentally different articulatory movements. In the case of vowels, the articulatory difference is more gradual, and the perception of vowel quality is therefore less categorical. The word *motor* is used in physiology and psychology to refer to the control of movement, so the motor theory states that the perception of speech sounds depends partly on the brain's awareness of the movements that must have been made to produce them. This theory was very influential in the 1950s and 60s but passed out of fashion; in recent years, however, we have seen something of a revival of motor theory and theories similar to it.

START ~ *INDEX*

nasal, nasalisation

A nasal consonant is one in which the air escapes only through the nose. For this to happen, two articulatory actions are necessary: firstly, the soft palate (or velum) must be lowered to allow air to escape past it, and secondly, a closure must be made in the oral cavity to prevent air from escaping through it. The closure may be at any place of articulation from bilabial at the front of the oral cavity to uvular at the back (in the latter case there is contact between the tip of the lowered soft palate and the raised back of the tongue). A closure any further back than this would prevent air from getting into the nasal cavity, so a pharyngeal or glottal nasal is a physical impossibility.

English has three commonly found nasal consonants: bilabial, alveolar and velar, for which the symbols **m**, **n** and **ŋ** are used. There is disagreement over the phonemic

status of the velar nasal: some claim that it must be a phoneme since it can be placed in contrastive contexts like ‘sum’/‘sun’/‘sung’, while others state that the velar nasal is an allophone of **n** which occurs before **k** and **g**.

In English we find nasal release of plosive consonants: when a plosive is followed by a nasal consonant the usual articulation is to release the compressed air by lowering the soft palate; this is particularly noticeable when the plosive and the nasal are homorganic (share the same place of articulation), as for example in ‘topmost’, ‘Putney’. The result is that no plosive release is heard from the speaker’s mouth before the nasal consonant.

You can read about English nasal consonants in *English Phonetics and Phonology*, Section 7.1.

When we find a vowel in which air escapes through the nose, it is usual to refer to this as a nasalised vowel, not a nasal vowel. Some languages (e.g. French) have nasalised vowel phonemes. In most other languages we find allophonic nasalisation when a vowel occurs close to a nasal consonant. In English, for example, the **ɑː** vowel in ‘can’t’ **kɑːnt** is nasalised so that the pronunciation is often (phonetically) **kɑ̃ːt**.

START ~ *INDEX*

Network English

This is a name for the American equivalent of BBC English or BBC pronunciation, the word ‘network’ referring to broadcasting networks. The Introduction to the *Cambridge English Pronouncing Dictionary* describes it as following ‘what is frequently heard from professional voices on national network news and information programmes. It is similar to what has been referred to as “General American”, which refers to a geographically (largely non-coastal) and socially-based set of pronunciation features’ (p. vi).

START ~ *INDEX*

neutralisation

In its simple form, the theory of the phoneme implies that two sounds that are in opposition to each other (e.g. **t** and **d** in English) are in this relationship in all contexts throughout the language. Closer study of phonemes has, however, shown that there are some contexts where the opposition no longer functions: for example, in a word like ‘still’ **stɪl**, the **t** is in a position (following **s** and preceding a vowel) where voiced (lenis) plosives do not occur. There is no possibility in English of the existence of a pair of words such as **stɪl** and **sdɪl**, so in this context the opposition between **t** and **d** is neutralised. One consequence of this is that one could equally well claim that the plosive in this word is a **d**, not a **t**. Common sense tells us that it is neither, but a

different phonological unit combining the characteristics of both. Some phonologists have suggested the word 'archiphoneme' for such a unit. The **i** vowel that we use to represent the vowel at the end of the word 'happy' could thus be called an archiphoneme.

START ~ *INDEX*

nucleus

Usually used in the description of intonation to refer to the most prominent syllable of the tone-unit, but also used in phonology to denote the centre or peak (i.e. vowel) of a syllable. It is one of the central principles of the "standard British" treatment of intonation that continuous speech can be broken up into units called tone-units, and that each of these will have one syllable that can be identified as the most prominent. This syllable will normally be the starting point of the major pitch movement (nuclear tone) in the tone unit. Another name for the nucleus is the tonic syllable.

START ~ *INDEX*

obstruent

Many different labels are used for types of consonant. One very general one that is sometimes useful is obstruent: consonants of this type create a substantial obstruction to the flow of air through the vocal tract. Plosives, fricatives and affricates are obstruents; nasals and approximants are not.

START ~ *INDEX*

occlusion

The term occlusion is used in some phonetics works as a technical term referring to an articulatory posture that results in the vocal tract being completely closed; the fact that the term closure is ambiguous supports the use of 'occlusion' for some purposes.

START ~ *INDEX*

oesophagus (also spelt **esophagus**)

Situated behind the trachea (or "windpipe") in the throat, the oesophagus is the tube down which food passes on its way to the stomach. It normally has little to do with speech, but it is possible for air pressure to build up (involuntarily or voluntarily) in

the trachea so as to produce a “belch”. When people have their larynx removed (usually because of cancer) they can learn to use this as an alternative airstream mechanism and speak quite effectively.

START ~ INDEX

onset

This term is used in the analysis of syllable structure (and occasionally in other areas); generally it refers to the first part of a syllable. In English this may be zero (when no consonant precedes the vowel in a syllable), one consonant, or two, or three. There are many restrictions on what clusters of consonants may occur in onsets: for example, if an English syllable has a three-consonant onset, the first consonant must be **s** and the last one must be one of **l**, **w**, **j**, **r**.

START ~ INDEX

open

One of the labels used for classifying vowels is *open*. An open vowel is one in which the tongue is low in the mouth and the jaw lowered: examples are cardinal vowel no. 4 [**a**] (similar to the **a** sound of French) and cardinal vowel no. 5 [**ɑ**] (like an exaggerated and old-fashioned English **ɑː**, as in ‘car’). The term ‘low’ is sometimes used instead of ‘open’, mainly by American phoneticians and phonologists.

START ~ INDEX

opposition

In the study of the phoneme it has been felt necessary to invent a number of terms to express the relationship between different phonemes. Sounds which are in opposition to each other are ones which can be substituted for each other in a given context (e.g. **t** and **k** in ‘patting’ and ‘packing’), producing different words. When we look at the whole set of phonemes in a language, we can often find very complex patterns of oppositions among the various groups of sounds.

START ~ INDEX

oral

Anything that is given the adjective oral is to do with the mouth. The oral cavity is the main cavity in the vocal tract. Consonants which are not nasal, and vowels which are not nasalised, may be called oral.

START ~ *INDEX*

Oxford accent

Some writers on English accents have attempted to subdivide "Received Pronunciation" into different varieties. Although the "Oxford accent" is usually taken to be the same thing as RP, it has been suggested that it may differ from that, particularly in prosody. There seems to be no scientific evidence for this, but the effect is supposed to be one of dramatic tempo variability, with alternation between extremely rapid speech on the one hand and excessive hesitation noises and drawled passages on the other. This is all rather fanciful, however, and should not be taken too seriously; if the notion has any validity, it is probably only in relation to an older generation.

START ~ *INDEX*

palatalisation

It is difficult to give a precise definition of this term, since it is used in a number of different ways. It may, for example, be used to refer to a process whereby the place of an articulation is shifted nearer to (or actually on to) the centre of the hard palate: the **s** at the end of the word 'this' may become palatalised to **ʃ** when followed by **j** at the beginning of 'year', giving **ðɪʃ jɪə**. (See coalescence.) However, in addition to this sense of the word we also find palatalisation being described as a secondary articulation in which the front of the tongue is raised close to the palate while an articulatory closure is made at another point in the vocal tract: in this sense, it is possible to find a palatalised **p** or **b**. Palatalisation is widespread in most Slavonic languages, where there are pairs of palatalised and non-palatalised consonants. The release of a palatalised consonant typically has a **j**-like quality.

START ~ *INDEX*

palate, palatal

The palate is sometimes known as the "roof of the mouth" (though the word "ceiling" would seem to be more appropriate). It can be divided into the hard palate, which runs from the alveolar ridge at the front of the mouth to the beginning of the soft palate at

the back, and the soft palate itself, which extends from the rear end of the hard palate almost to the back of the throat, terminating in the uvula, which can be seen in a mirror if you look at yourself with your mouth open. The hard palate is mainly composed of a thin layer of bone (which has a front-to-back split in it in the case of people with cleft palate), and is dome-shaped, as you can feel by exploring it with the tip of your tongue. The soft palate (for which there is an alternative name, velum) can be raised and lowered; it is lowered for normal breathing and for nasal consonants, and raised for most other speech sounds.

Consonants in which the tongue makes contact with the highest part of the hard palate are labelled palatal. These include the English **j** sound.

START ~ *INDEX*

paralinguistic(s)

It is often difficult to decide which of the features of speech that we can observe are part of the language (or linguistic system) and which are outside it. We are usually confident in classing vowel and consonant sounds as linguistically relevant, and in excluding coughs and sneezes (since these are never used contrastively). But there are various features that are "borderline", and the general term *paralinguistic* is often used for such features: these can include such things as different voice qualities, gestures, facial expressions and unusual ways of speaking such as laughing at the same time as speaking. Linguists disagree about which of these form part of the sound system of the language.

START ~ *INDEX*

passive articulator

Articulators are the parts of the body that are used in the production of speech. Some of these (e.g. the tongue, the lips) can be moved, while others (e.g. the hard palate, the teeth) are fixed. Fixed articulators are sometimes called passive articulators, and their most important function is to act as the place of an articulatory stricture.

START ~ *INDEX*

pause

The most obvious purpose of a pause is to allow the speaker to draw breath, but we pause for a number of other reasons as well. One type of pause that has been the subject of many studies by psycholinguists is the "planning pause", where the speaker is assumed to be constructing the next part of what (s)he is going to say, or is

searching for a word that is difficult to retrieve. As every actor knows, pauses can also be used for dramatic effect at significant points in a speech.

From the phonetic point of view, pauses differ from each other in two main ways: one is the length of the pause, and the other is whether the pause is silent or contains a "hesitation noise". (See also hesitation.)

START ~ INDEX

peak

In the phonological study of the syllable it is conventional to give names to its different components. The centre of the syllable is its peak; this is normally a vowel, but it is possible for a consonant to act as a peak instead. (See syllabic consonant.)

START ~ INDEX

perception

Most of the mental processes involved in understanding speech are unknown to us, but it is clear that discovering more about them can be very important in the general study of pronunciation. It is clear from what we know already that perception is strongly influenced by the listener's expectations about the speaker's voice and what the speaker is saying; many of the assumptions that a listener makes about a speaker are invalid when the speaker is not a native speaker of the language, and it is hoped that future research in speech perception will help to identify which aspects of speech are most important for successful understanding and which type of learner error has the most profound effect on intelligibility.

START ~ INDEX

pharynx

This is the tube which connects the larynx to the oral cavity. It is usually classed as an articulator; the best-known language that has consonants with pharyngeal (or pharyngal) place of articulation is Arabic, most dialects of which have voiced and voiceless pharyngeal fricatives made by constricting the muscles of the pharynx (and usually also some of the larynx muscles) to create an obstruction to the airflow from the lungs.

START ~ INDEX

phatic communion

This is a rather pompous name for an interesting phenomenon: often when people appear to be using language for social purposes it seems that the actual content of what they are saying has virtually no meaning. For example, greetings containing an apparent enquiry about the listener's health or a comment on the weather are usually not expected to be treated as a normal enquiry or comment. What is interesting from the pronunciation point of view is that such interactions only work if they are said in a prosodically appropriate way: it has been claimed that when welcoming a guest to a lively party one could announce (without anyone noticing anything wrong) that one had just finished murdering one's grandmother, as long as one used the appropriate intonation and facial expression for a greeting.

START ~ *INDEX*

phonation

This is a technical term for the vibration of the vocal folds; it is more commonly known as voicing.

START ~ *INDEX*

phone

The term phoneme has become very widely used for a contrastive unit of sound in language: however, a term is also needed for a unit at the phonetic level, since there is not always a one-to-one correspondence between units at the two levels. For example, the word 'can't' is phonemically **kɑːnt** (four phonemic units), but may be pronounced **kɑ̃t** with the nasal consonant phoneme absorbed into the preceding vowel as nasalisation (three phonetic units). The term *phone* has been used for a unit at the phonetic level, but it has to be said that the term (though useful) has not become widely used; this must be at least partly due to the fact that the word is already used for a much more familiar object.

START ~ *INDEX*

phoneme

This is the fundamental unit of phonology, which has been defined and used in many different ways. Virtually all theories of phonology hold that spoken language can be broken down into a string of sound units (phonemes), and that each language has a small, relatively fixed set of these phonemes. Most phonemes can be put into groups; for example, in English we can identify a group of plosive phonemes **p**, **t**, **k**, **b**, **d**, **g**,

a group of voiceless fricatives **f**, **θ**, **s**, **ʃ**, **h,** and so on. An important question in phoneme theory is how the analyst can establish what the phonemes of a language are. The most widely accepted view is that phonemes are contrastive and one must find cases where the difference between two words is dependent on the difference between two phonemes: for example, we can prove that the difference between 'pin' and 'pan' depends on the vowel, and that **ɪ** and **æ** are different phonemes. Pairs of words that differ in just one phoneme are known as minimal pairs. We can establish the same fact about **p** and **b** by citing 'pin' and 'bin'. Of course, you can only start doing commutation tests like this when you have a provisional list of possible phonemes to test, so some basic phonetic analysis must precede this stage. Other fundamental concepts used in phonemic analysis of this sort are complementary distribution, free variation, distinctive feature and allophone.

Different analyses of a language are possible: in the case of English some phonologists claim that there are only six vowel phonemes, others that there are twenty or more (it depends on whether you count diphthongs and long vowels as single phonemes or as combinations of two phonemes).

It used to be said that learning the pronunciation of a language depended on learning the individual phonemes of the language, but this "building-block" view of pronunciation is looked on nowadays as an unhelpful oversimplification.

START ~ *INDEX*

phonemics

When the importance of the phoneme became widely accepted, in the 1930s and 40s, many attempts were made to develop scientific ways of establishing the phonemes of a language and listing each phoneme's allophones; this was known as *phonemics*. Nowadays little importance is given to this type of analysis, and it is considered a minor branch of phonology, except for the practical purpose of devising writing systems for previously unwritten languages.

START ~ *INDEX*

phonetics

Phonetics is the scientific study of speech. It has a long history, going back certainly to well over two thousand years ago. The central concerns in phonetics are the discovery of how speech sounds are produced, how they are used in spoken language, how we can record speech sounds with written symbols and how we hear and recognise different sounds. In the first of these areas, when we study the production of speech sounds we can observe what speakers do (articulatory observation) and we can try to feel what is going on inside our vocal tract (kinaesthetic observation). The second area is where phonetics overlaps with phonology: usually in phonetics we are

only interested in sounds that are used in meaningful speech, and phoneticians are interested in discovering the range and variety of sounds used in this way in all the known languages of the world. This is sometimes known as linguistic phonetics. Thirdly, there has always been a need for agreed conventions for using phonetic symbols that represent speech sounds; the International Phonetic Association has played a very important role in this. Finally, the auditory aspect of speech is very important: the ear is capable of making fine discrimination between different sounds, and sometimes it is not possible to define in articulatory terms precisely what the difference is. A good example of this is in vowel classification: while it is important to know the position and shape of the tongue and lips, it is often very important to have been trained in an agreed set of standard auditory qualities that vowels can be reliably related to. (See cardinal vowel; other important branches of phonetics are experimental, instrumental and acoustic.)

START ~ *INDEX*

phonology

The most basic activity in phonology is phonemic analysis, in which the objective is to establish what the phonemes are and arrive at the phonemic inventory of the language. Very few phonologists have ever believed that this would be an adequate analysis of the sound system of a language: it is necessary to go beyond this. One can look at suprasegmental phonology – the study of stress, rhythm and intonation, which has led in recent years to new approaches to phonology such as metrical and autosegmental theory; one can go beyond the phoneme and look into the detailed characteristics of each unit in terms of distinctive features; the way in which sounds can combine in a language is studied in phonotactics and in the analysis of syllable structure. For some phonologists the most important area is the relationships between the different phonemes – how they form groups, the nature of the oppositions between them and how those oppositions may be neutralised.

Until the second half of the twentieth century most phonology had been treated as a separate "level" that had little to do with other "higher" areas of language such as morphology and grammar. Since the 1960s the subject has been greatly influenced by generative phonology, in which phonology becomes inextricably bound up with these other areas; this has made contemporary phonology much harder to understand, but it has the advantage that it no longer appears to be an isolated and self-contained field.

START ~ *INDEX*

phonotactics

It has often been observed that languages do not allow phonemes to appear in any order; a native speaker of English can figure out fairly easily that the sequence of phonemes **strɛŋθs** makes an English word ('strengths'), that the sequence **blɛɪdʒ** would be acceptable as an English word 'blage' although that word does not happen

to exist, and that the sequence **lvɜːʒm** could not possibly be an English word. Knowledge of such facts is important in phonotactics, the study of sound sequences.

Although it is not necessary to do so, most phonotactic analyses are based on the syllable. Phonotactic studies of English come up with some strange findings: certain sequences seem to be associated with particular feelings or human characteristics, for no obvious reason. Why should 'bump', 'lump', 'hump', 'rump', 'mump(s)', 'clump' and others all be associated with large blunt shapes? Why should there be a whole family of words ending with a plosive and a syllabic **l** all having meanings to do with clumsy, awkward or difficult action ('muddle', 'fumble', 'straddle', 'cuddle', 'fiddle', 'buckle' (vb.), 'struggle', 'wriggle')? Why can't English syllables begin with **pw**, **bw**, **tl**, **dl** when **pl**, **bl**, **tw**, **dw** are acceptable?

START ~ *INDEX*

pitch

Pitch is an auditory sensation: when we hear a regularly vibrating sound such as a note played on a musical instrument, or a vowel produced by the human voice, we hear a high pitch if the rate of vibration is high and a low pitch if the rate of vibration is low. Many speech sounds are voiceless (e.g. **s**), and cannot give rise to a sensation of pitch in this way. The pitch sensation that we receive from a voiced sound corresponds quite closely to the frequency of vibration of the vocal folds; however, we usually refer to the vibration frequency as fundamental frequency in order to keep the two things distinct.

Pitch is used in many languages as an essential component of the pronunciation of a word, so that a change of pitch may cause a change in meaning: these are called tone languages. In most languages (whether or not they are tone languages) pitch plays a central role in intonation.

START ~ *INDEX*

pitch range

In studying tone and intonation, it is very important to remember that each person has her or his own pitch range, so that what is high pitch for a person with a low-pitched voice may be the same as low pitch for a person with a high-pitched voice. Consequently, whatever we say about a speaker's use of pitch must be relative to that person's personal pitch range. Each of us has a highest and a lowest pitch level for speaking, though we may occasionally go outside that range when we are very emotional.

START ~ *INDEX*

place of articulation

Consonants are made by producing an obstruction to the flow of air at some point in the vocal tract, and when we classify consonants one of the most important things to establish is the place where this obstruction is made; this is known as the place of articulation, and in conventional phonetic classification each place of articulation has an adjective that can be applied to a consonant. To give a few examples of familiar sounds, the place of articulation for **p**, **b** is bilabial, for **f**, **v** labiodental, for **θ**, **ð** dental, for **t**, **d** alveolar , for **ʃ**, **ʒ** post-alveolar, for **k**, **g** velar, and for **h** glottal. The full range of places of articulation can be seen on the IPA chart.

Sometimes it is necessary to specify more than one place of articulation for a consonant, for one of two reasons: firstly, there may be a secondary articulation – a less extreme obstruction to the airflow, but one which is thought to have a significant effect; secondly, some languages have consonants that make two simultaneous constrictions, neither of which could fairly be regarded as taking precedence over the other. A number of West African languages, such as Igbo, have consonants which involve simultaneous plosive closures at the lips and at the velum, as in, for example, the labial-velar stops **kp**, **gb** found in Igbo and Yoruba.

START ~ *INDEX*

plosion

When a plosive is released and is followed by a vowel or a pause, there is usually a small explosive noise made as the compressed air escapes. This is easier to hear in the case of English voiceless or fortis plosives, though this effect is sometimes masked by glottalisation.

START ~ *INDEX*

plosive

In many ways it is possible to regard plosives as the most basic type of consonant. They are produced by forming a complete obstruction to the flow of air out of the mouth and nose, and normally this results in a build-up of compressed air inside the chamber formed by the closure. When the closure is released, there is a small explosion (see plosion) that causes a sharp noise. Plosives are among the first sounds that are used by children when they start to speak (though nasals are likely to be the very first consonants). The basic plosive consonant type can be exploited in many different ways: plosives may have any place of articulation, may be voiced or voiceless and may have an egressive or ingressive airflow. The airflow may be from the lungs (pulmonic), from the larynx (glottalic) or generated in the mouth (velaric). We find great variation in the release of the plosive.

START ~ INDEX

polysyllabic

A linguistic unit such as a word, morpheme or phrase is polysyllabic if it contains more than one syllable.

START ~ INDEX

pragmatics

In analysing different styles of speech, and studying the use of prosody, it is very important to be able to specify what the objective of the speaker of a particular utterance was: studying speech and language data out of context has been a serious weakness of many past studies. Pragmatics is a field of study that concerns itself with the social, communicative and practical use of language, and has become recognised as a vital part of linguistics. Work in this field looks at such things as the presuppositions and background knowledge that language users need to have in order to communicate, the strategies they adopt in order to make a point convincingly and the kinds of function that language is used for.

START ~ INDEX

pre-fortis clipping

Fortis consonants have the effect of shortening a preceding vowel or sonorant consonant, so that, for example, 'bit' has a shorter vowel than 'bid'. This effect is sometimes called pre-fortis clipping.

START ~ INDEX

prominence

"Stress" or "accentuation" depends crucially on the speaker's ability to make certain syllables more noticeable than others. A syllable which "stands out" in this way is a prominent syllable. An important thing about prominence, at least in English, is the fact that there are many ways in which a syllable can be made prominent: experiments have shown that prominence is associated with greater length, greater loudness, pitch prominence (i.e. having a pitch level or movement that makes a syllable stand out from its context) and with "full" vowels and diphthongs (whereas the vowels **ə** "schwa", **i**, **u** and syllabic consonants are only found in unstressed syllables). Despite

the complexity of this set of interrelated factors, it seems that the listener simply hears syllables as more prominent or less prominent.

START ~ INDEX

pronouncing dictionary (pronunciation dictionary)

It is probably only the English language, with its complex and unpredictable spelling system, that needs a special kind of dictionary to tell you how to pronounce words which you know how to write. With a pronouncing dictionary, the user looks up the required word in its spelling form and reads the pronunciation in the form of phonetic or phonemic transcription. (Actually, one of the earliest pronunciation dictionaries, published in 1913, worked the other way round, giving the spelling for a word which the user already knew and looked up in phonemic form. It is not reported to have been a big success.) Normally, several alternative pronunciations will be offered, with an indication of which is the most usual and possibly some information on other accents (e.g. a dictionary based on the BBC accent, or "Received Pronunciation", might also give one or more American pronunciation for a word). The importance of pronouncing dictionaries has declined to some extent in recent years as most modern English-language dictionaries now include pronunciation information in phonemic transcription for each entry, but they are still widely used.

START ~ INDEX

pronunciation

It is not very helpful to be told that pronunciation is the act of producing the sounds of a language. The aspects of this subject that concern most people are (1) standards of pronunciation and (2) the learning of pronunciation. In the case of (1) standards of pronunciation, the principal factor is the choice of model accent: once this decision is made, any deviation from the model tends to attract criticism from people who are concerned with standards; the best-known example of this is the way people complain about "bad" pronunciation in an "official" speaker of the BBC, but similar complaints are made about the way children pronounce their native language in school, or the way immigrant children fail to achieve native-speaker competence in the pronunciation of the "host" language. These are areas that are as much political as phonetic, and it is difficult to see how people will ever agree on them. In the area of (2) pronunciation teaching and learning, a great deal of research and development has been carried out since the early 20th century by phoneticians. It should be remembered that, useful though practical phonetics is in the teaching and learning of pronunciation, it is not essential, and many people learn to pronounce a language that they are learning simply through imitation and correction by a teacher or a native speaker.

START ~ INDEX

prosody/ic

It is traditional in the study of language to regard speech as being basically composed of a sequence of sounds (vowels and consonants); the term prosody and its adjective prosodic is then used to refer to those features of speech (such as pitch) that can be added to those sounds, usually to a sequence of more than one sound. This approach can sometimes give the misleading impression that prosody is something optional, added like a coat of paint, when in reality at least some aspects of prosody are inextricably bound up with the rest of speech. The word suprasegmental has practically the same meaning.

A number of aspects of speech can be identified as significant and regularly used prosodic features; the most thoroughly investigated is intonation, but others include stress, rhythm, voice quality, loudness and tempo (speed).

START ~ *INDEX*

public school accent

Foreigners are often surprised to find that in Britain, so-called public schools are *private* schools, and are used almost exclusively to educate the children of the wealthy. They are one of the strongest forces for conservatism and the preservation of privilege in British society, and one of the ways in which they preserve traditional conventions is to encourage in their pupils the use of "Received Pronunciation" (RP), also known as BBC pronunciation. This accent is therefore sometimes referred to as the "public-school accent".

START ~ *INDEX*

pulmonic

Almost all the sounds we make in speaking are created with the help of air compressed by the lungs. The adjective used for this lung-created airstream is 'pulmonic': the pulmonic airstream may be ingressive (as in breathing in) but for speaking is practically always egressive.

START ~ *INDEX*

pure vowel

This term is used to refer to a vowel in which there is no detectable change in quality from beginning to end; an alternative name is monophthong. These are contrasted with vowels containing a movement, such as the glide in a diphthong.

START ~ INDEX

rate

The word *rate* is used in talking about the speed at which we speak; in laboratory studies of speech it is usual to express this in terms of syllables per second, or sometimes (less usefully) in words per minute. An alternative term is tempo.

START ~ INDEX

realisation

As a technical term, this word is used to refer to the act of pronouncing a phoneme. Since phonemes are said to be abstract units, they are not physically real. However, when we speak we produce sounds, and these are the physical realisations of the phonemes. Each realisation is different from every other (since you can never do exactly the same thing twice), but also some realisations are noticeably different in quality from others (e.g. the English phoneme **l** is sometimes realised as a "clear l" and sometimes as a "dark l"). In this case it is more appropriate to call the sounds allophones.

START ~ INDEX

Received Pronunciation (RP)

RP has been for centuries the accent of British English usually chosen for the purposes of description and teaching, in spite of the fact that it is only spoken by a small minority of the population; it is also known as the "public school" accent, and as "BBC pronunciation". There are clear historical reasons for the adoption of RP as the model accent: in the first half of the twentieth century virtually any English person qualified to teach in a university and write textbooks would have been educated at private schools: RP was (and to a considerable extent still is) mainly the accent of the privately educated. It would therefore have been a bizarre decision at that time to choose to teach any other accent to foreign learners. It survived as the model accent for various reasons: one was its widespread use in "prestige" broadcasting, such as news-reading; secondly, it was claimed to belong to no particular region, being found in all parts of Britain (though in reality it was very much more widespread in London and the south-east of England than anywhere else); and thirdly, it became accepted as a common currency – an accent that (it was claimed) everyone in Britain knows and understands.

Some detailed descriptions of RP have suggested that it is possible to identify different varieties within RP, such as "advanced", or "conservative". Another suggestion is that there is an exaggerated version that can be called "hyper-RP". But

these sub-species do not appear to be easy to identify reliably. My own opinion is that RP was a convenient fiction, but one which had regrettable associations with high social class and privilege. I prefer to treat the BBC accent as the best model for the description of English, and to consign "Received Pronunciation" to history.

START ~ *INDEX*

reduction

When a syllable in English is unstressed, it frequently happens that it is pronounced differently from the "same" syllable when stressed; the process is one of weakening, where vowels tend to become more schwa-like (i.e. they are centralised), and plosives tend to become fricatives. The reduced forms of vowels can be clearly seen in the set of words 'photograph' **ˈfəʊtəgrɑːf**, 'photography' **fəˈtɒgrəfi**, 'photographic' **ˌfəʊtəˈgræfɪk** – when one of the three syllables does not receive stress its vowel is reduced to **ə**. This is felt to be an important characteristic of English phonetics, and something that is not found in all languages. It is possible that the difference between languages which exhibit vowel reduction and those which do not is closely parallel to the proposed difference between "stress-timed" and "syllable-timed" languages.

START ~ *INDEX*

register

Several uses are made of this word: in singing, it is used to refer to different styles of voice production that the singer may select, particularly head register and chest register. The term is also used by some phoneticians to refer to similar options in speaking (see voice quality). A further use of the term is in the typology of tone languages: it has been proposed that all tone languages could be categorised either as contour languages or as register languages. In the latter, the most important characteristic of a tone is its pitch level relative to the speaker's pitch range, rather than the shape of any pitch movement.

START ~ *INDEX*

release

Only consonants which involve a complete, air-tight closure are properly described as having a release component, which means that only plosive and affricate consonants are to be considered. When air is compressed behind a complete closure in the vocal tract, the release may be one of several different sorts. Firstly, the release may happen when the air pressure is near its maximum, resulting in a loud explosive sound, or it may happen (particularly in final position) that the speaker allows the air pressure to reduce before the release, so that the resulting noise is much less. Since an airstream

is involved, the release may be egressive (the usual situation) or ingressive (as in clicks and implosives). In addition, the release may be simple or complex. If it is simple, the released air escapes in a rush directly from the oral cavity into the atmosphere (assuming an egressive airstream); if a vowel follows and the start of voicing is delayed we say that the plosive is aspirated. The release is complex if the passage of the released air is modified by some other articulation that follows immediately. If the release is followed by fricative noise produced in the same place of articulation as the plosive closure, we describe the resulting plosive-plus-fricative sound as an affricate. Alternatively, there may be nasal release or lateral release.

START ~ *INDEX*

resonance

This term is widely used in non-scientific ways, and also with technical senses in phonetics and speech acoustics. In its non-technical sense it is often found in music, especially singing (e.g. "his bass voice had a rich resonance"); in auditory phonetics it is sometimes used to refer to particular sound qualities (e.g. "her **l** sound has a dark resonance"). But in acoustic terminology the word is used in a different way. Many people first discover resonance while singing in the bath: singing a particular note creates a powerful "booming" effect, while other notes do not have the same effect. Like bathrooms, vocal tracts have natural resonant frequencies. In speech acoustics, the vocal tract is thought of as a continuous tube with different dimensions at different places along its length. As with all tubes and chambers, it is possible to identify particular frequencies at which there are resonances – these are observable as peaks of energy, or formants. In the case of voiced speech sounds, the acoustic energy generated in the larynx passes through the vocal tract and at most frequencies much of the energy is lost; however, at the few frequencies where the sound wave resonates most of the energy passes through, creating peaks of energy at those frequencies. In the case of voiceless sounds, resonance is more difficult to explain.

START ~ *INDEX*

retroflex

A retroflex articulation is one in which the tip of the tongue is curled upward and backward. The **r** sound of BBC English and General American is sometimes described as being retroflex, though in normal speech the degree of retroflexion is relatively small. Other languages have retroflex consonants with a more noticeable auditory quality, the best known examples being the great majority the languages of the Indian sub-continent. The sound of retroflex consonants is fairly familiar to English listeners, since first-generation immigrants from India and Pakistan tend to carry the retroflex quality into their pronunciation of English and this is often mimicked.

In American English and some accents of south-west England it is common for vowels preceding **r** (e.g. **ɑː** in 'car', or **ɜː** in 'bird') to be affected by the consonant so that they have a retroflex quality for most of their duration. This "r-colouring" is most common in back or central vowels where the forward part of the tongue is relatively free to change shape.

START ~ *INDEX*

rhotic/ity

This term is used to describe varieties of English pronunciation in which the **r** phoneme is found in all phonological contexts. In BBC Pronunciation, **r** is only found before vowels (as in 'red' **red**, 'around' **əraʊnd**), but never before consonants or before a pause. In rhotic accents, on the other hand, **r** may occur before consonants (as in 'cart' **kɑːrt**) and before a pause (as in 'car' **kɑːr**). While BBC pronunciation is non-rhotic, many accents of the British Isles are rhotic, including most of the south and west of England, much of Wales, and all of Scotland and Ireland. Most speakers of American English speak with a rhotic accent, but there are non-rhotic areas including the Boston area, lower-class New York and the Deep South.

Foreign learners encounter a lot of difficulty in learning not to pronounce **r** in the wrong places, and life would be easier for most learners of English if the model chosen were rhotic.

START ~ *INDEX*

rhyme

Rhyming verse has pairs of lines that end with the same sequence of sounds. If we examine the sound sequences that must match each other, we find that these consist of the vowel and any final consonants of the last syllable: thus 'moon' and 'June' rhyme, and the initial consonants of these two words are not important (of course, we do find longer-running rhymes than this in verse, particularly the comic variety, e.g. 'ability' rhyming with 'senility', 'Harvard' with 'discovered').

The concept of rhyme has become useful in the phonological analysis of the syllable as a way of referring to the vowel peak of the syllable plus any sounds following the peak within the syllable (the coda). Thus in the word 'spoon' the rhyme is **uːn**, in 'tea' it is **iː** and in 'strengths' it is **eŋθs** or **eŋkθs**.

START ~ *INDEX*

rhythm

Speech is perceived as a sequence of events in time, and the word rhythm is used to refer to the way events are distributed in time. Obvious examples of vocal rhythms are chanting as part of games (for example, children calling words while skipping, or football crowds calling their team's name) or in connection with work (e.g. sailors' chants used to synchronise the pulling on an anchor rope). In conversational speech the rhythms are vastly more complicated, but it is clear that the timing of speech is not random. An extreme view (though a quite common one) is that English speech has a rhythm that allows us to divide it up into more or less equal intervals of time called feet, each of which begins with a stressed syllable: this is called the stress-timed rhythm hypothesis. Languages where the length of each syllable remains more or less the same as that of its neighbours whether or not it is stressed are called syllable-timed. Most evidence from the study of real speech suggests that such rhythms only exist in very careful, controlled speaking, but it appears from psychological research that listeners' brains tend to hear timing regularities even where there is little or no physical regularity.

START ~ *INDEX*

root (of tongue)

The base of the tongue, where it is attached to the rear end of the lower jaw, is known as the root. This has usually been assumed to have no linguistic function. However, it has been discovered that some non-European languages have vowels that differ from each other in terms of quality, and the only articulatory difference between them appears to be that some are pronounced with the tongue root moved forward and some have the tongue root further back.

START ~ *INDEX*

rounding

Practically any vowel or consonant may be produced with different amounts of lip-rounding. The lips are rounded by muscles that act rather like a drawstring round the neck of a bag, bringing the edges of the lips towards each other. Except in unusual cases, this results not only in the mouth opening adopting a round shape, but also in a protrusion or "pushing forward" of the lips; Swedish is described as having a rounded vowel without lip protrusion, however. In theory any vowel position (defined in terms of height and frontness/backness) may be produced rounded or unrounded, though we do not necessarily find all possible vowels with and without rounding in natural languages. Consonants, too, may have rounded lips (in w, the basic consonantal articulation itself consists of lip-rounding): this lip-rounding in consonants is regarded as a secondary articulation, and it is usual to refer to it as labialisation. In BBC pronunciation, it is common to find ʃ, ʒ, tʃ, dʒ and r with slight lip-rounding.

START ~ INDEX

sandhi

The ways in which speech sounds influence each other when they are neighbours is of great interest to contemporary phoneticians and phonologists (see assimilation and coalescence), but the subject is also one which interested the Sanskrit grammarians of India (who introduced the term) over two thousand years ago. The notion of sandhi is used mainly in the area between morphology and phonology, and is not much used in the study of pronunciation. It is most commonly found in discussion of tone languages and the contextual influences on tones.

START ~ INDEX

schwa

One of the most noticeable features of English pronunciation is the phonetic difference between stressed and unstressed syllables. In most languages, any of the vowels of the language can occur in any syllable whether that syllable is stressed or not; in English, however, a syllable which bears no stress is more likely to have one of a small number of weak vowels, and the most common weak vowel is one which never occurs in a stressed syllable. That vowel is the schwa vowel (symbolised **ə**), which is generally described as being unrounded, central (i.e. between front and back) and mid (i.e. between close and open). Statistically, this is reported to be the most frequently occurring vowel of English (over 10% of all vowels). It is ironic that the most frequent English vowel has no regular letter for its spelling. The name *schwa* comes from Hebrew, which does have a symbol for this sound.

Many foreign learners of English have difficulty in learning to pronounce schwa.

START ~ INDEX

secondary articulation

In classifying consonants it is usual to identify the place of articulation of the major constriction; however, in the case of most consonants it is possible to add an additional stricture at some other point in the vocal tract. A simple example is lip-rounding: English **ʃ**, for example, is often pronounced with rounded lips, and in this case the rounding is a secondary articulation (where the primary articulation is the post-alveolar fricative constriction). Velarisation is another secondary articulation: in this case the back of the tongue is raised while a more extreme constriction is made elsewhere. This mechanism is used extensively in Arabic for the production of the "emphatic" consonants, and in English is the means for giving a "dark l" its distinctive quality.

segment

Phoneticians and phonologists disagree about segments: when we analyse an utterance, we can identify a number of phonological and grammatical elements, partly as a result of our knowledge of the language. Consequently, we are able to write down something we hear in words separated by spaces, and (with proper training) transcribe with phonemic symbols the sounds that we hear. However, when we examine speech sounds in connected speech closely, we find many cases where it is difficult to identify separate sound units (segments) that correspond to phonemes, since many of the articulatory movements that create the sounds tend to be continuous rather than sharply switched. For example, pre-consonantal **n** sounds in English (e.g. 'kind' **kaɪnd**) are often almost undetectable except in the form of nasalisation of the vowel preceding them; sequences of fricatives often overlap, so that it is difficult or impossible to split the sequence **ʃs** in 'fish soup', or **fθs** in 'fifths'. As a result, some people believe that dividing speech up into segments (segmentation) is fundamentally misguided; the opposite view is that since segmentation appears to be possible in most cases, and speakers seem to be aware of segments in their speech, we should not reject segmentation because there are problematical cases.

semivowel

It has long been recognised that most languages contain a class of sound that functions in a way similar to consonants but is phonetically similar to vowels: in English, for example, the sounds **w** and **j** (as found in 'wet' and 'yet') are of this type: they are used in the first part of syllables, preceding vowels, but if **w** and **j** are pronounced slowly, it can be clearly heard that in quality they resemble the vowels [**u**] and [**i**] respectively. (See also contoid and vocoid.) The term semivowel has been in use for a long time for such sounds, though it is not a very helpful or meaningful name; the term approximant is more often used today. Americans usually use the symbol **y** for the sound in 'yes', but European phoneticians reserve this symbol for a close front rounded vowel.

English has words which are pronounced differently according to whether they are followed by a vowel or a consonant: these are 'the' **ðɪ** or **ðə** and the indefinite article 'a/an', and it is the pre-consonantal form that we find before **j** and **w**. In addition, "linking r", which is found in BBC and other non-rhotic accents, does not appear before semivowels. It is by looking at evidence such as this that we can conclude that as far as English is concerned, **j** and **w** are in the same phonological class as the other consonants despite their vowel-like phonetic nature.

In French there are three sounds traditionally classed as semivowels: in addition to **j** and **w** there is a sound based on the front rounded vowel **y** (as in 'tu', 'lu'); this semivowel is symbolised **ɥ** and is found in initial position in the word 'huit' **ɥit** (eight) and in consonant clusters such as **fr** in **frɥi** ('fruit'). The IPA chart also lists a semivowel **ʍ** corresponding to the back close unrounded vowel **ɯ**. Like the others, this is classed as an approximant.

START ~ *INDEX*

sentence stress

The main question that is asked in studying so-called sentence stress is which syllable (or word) of a particular sentence is most strongly stressed (or accented). We should be clear that in any given sentence of more than one syllable there is no logical necessity for there to be just one syllable that stands out from all the others. Much writing on this subject has been done on the basis of short, invented sentences designed to have just one obvious sentence stress, but in real life we often find exceptions to this. In a sentence of more than five or six words we tend to break the string of words into separate tone-units, each of which will be likely to have a strong stress. For example:

> If she hadnt been rich | she couldnt have bought it

In addition we find cases where syllables in two neighbouring words seem to be equally strongly stressed. For example:

> Ive \burnt /most of them. (with pitch fall on 'burnt' and pitch rise on 'most')

Given that (in English, at least), sentence stress is a rather badly-defined notion, is it at least possible to make generalisations about stress placement in simple sentences? It is widely believed that the most likely place for sentence stress to fall is on the appropriate syllable of the last lexical word of the sentence: in this case, "appropriate syllable" refers to the syllable indicated by the rules for word stress, while lexical word refers to words such as nouns, verbs, adjectives and adverbs. This rule accounts for the stress pattern of many sentences, but there is considerable controversy over how to account for the many exceptions: some linguists say that the sentence stress tends to be placed on the word which is most important to the meaning of the sentence, while others say that the placement of the stress is determined by the underlying syntactic structure.

Many other languages seem to exhibit very similar use of stress, but it is not possible in the present state of our knowledge to say whether there are universal tendencies in all languages to position sentence stress in predictable ways.

START ~ *INDEX*

sibilant

It is sometimes necessary to make subdivisions within the very large set of possible fricative sounds. As explained under fricative, one possible division is between those fricatives which make a sharp or strong hissing noise (e.g. **s**, **ʃ**) and those which produce only a soft noise (e.g. **f**, **θ**). In English we use the sibilant sound **ʃ** to command silence (e.g. in a classroom). Some other cultures use **s**, but it is hard to imagine anyone using **f** or **θ** for this purpose.

START ~ *INDEX*

slip of the tongue (speech error)

Much has been discovered about the control of speech production in the brain as a result of studying the errors we make in speaking. These are traditionally known as "slips of the tongue", though as has often been pointed out, it is not usually the tongue that slips, but the brain which is attempting to control it. Some errors involve unintentionally saying the wrong word (a type of slip that the great psychoanalyst Freud was particularly interested in), or being unable to think of a word that one knows. Many slips involve phonemes occurring in the wrong place, either through perseveration (i.e. repeating a segment that has occurred before, as in 'cup of key' for 'cup of tea') or transposition (the slip known as a Spoonerism), as in 'tasted a worm' instead of 'wasted a term'. My favourite example of a Spoonerism is one I heard myself on the radio recently, where the speaker said 'hypodeemic nerdle' **haɪpədiːmɪk nɜːdl̩** instead of 'hypodermic needle' **haɪpədɜːmɪk niːdl̩** – stressed syllables of the two words were interchanged. Such slips apparently never result in an unacceptable sequence of phonemes: for example, 'brake fluid' could be mispronounced through a Spoonerism as 'frake bluid', but 'brake switch' could never be mispronounced in this way since it would result in 'srake bwitch', and English syllables do not normally begin with **sr** or **bw**.

Some researchers have made large collections of recorded speech errors, and there are many discoveries still to be made in this field.

START ~ *INDEX*

slit

In a fricative made by forming a constriction between the tongue and the palate, the hole through which the air escapes may be narrow and deep (groove) or wide and shallow (slit). (See groove.)

START ~ *INDEX*

soft palate

Most of the roof of the mouth consists of hard palate, which has bone beneath the skin. Towards the back of the mouth, the layer of bone comes to an end but the layer of soft tissue continues for some distance, ending eventually in a loose appendage that can easily be seen by looking in a mirror: this dangling object is the uvula, but the layer of soft tissue to which it is attached is called the soft palate (it is also sometimes named the velum). In normal breathing it is allowed to hang down so that air may pass above it and escape through the nose, but for most speech sounds it is lifted up and pressed against the upper back wall of the throat so that no air can escape through the nose. This is necessary for a plosive, for example, so that air may be compressed within the vocal tract. However, for nasal consonants (e.g. **m**, **n**) the soft palate must be lowered since air can escape only through the nose in these sounds. In nasalised vowels (such vowels are found in considerable numbers in French, for example) the soft palate is lowered and air escapes through the mouth and the nose together.

START ~ *INDEX*

sonorant

Many technical terms have been invented in phonology to refer to particular groups or families of sounds. A sonorant is a sound which is voiced and does not cause enough obstruction to the airflow to prevent normal voicing from continuing. Thus vowels, nasals, laterals and other approximants such as English **j**, **w**, **r** are sonorants, while plosives, fricatives and affricates are non-sonorants.

START ~ *INDEX*

sonority

It is possible to describe sounds in terms of how powerful they sound to the listener; a vowel sound such as **a** is said to be more sonorant than the fricative **f**, for example. It is said that if we hear a word such as 'banana' as consisting of three syllables, it is because we can hear three peaks of sonority corresponding to the vowels. Some phonologists claim that there is a sonority hierarchy among classes of sound that governs the way they combine with other sounds: in descending order of sonority, we would find firstly open vowels like **a**, then closer vowels (e.g. **i**, **u**); "liquids" such as **l**, **r**, followed by nasals, fricatives and finally plosives (the least sonorant).

START ~ *INDEX*

spectrogram, spectrography

In the development of the laboratory study of speech, the technique that has been the most fundamental tool in acoustic analysis is spectrography. In its earliest days, this was carried out on special machines that analysed a few seconds of speech and burned patterns on heat-sensitive paper, but all spectrography is now done by computers. A spectrography program on a computer produces a sort of picture, in shades of grey or in a variety of colours, of the recorded sounds, and this spectrogram is shown on the computer screen and can be printed. With practice, an analyst can identify many fine details of speech sounds. The cover of *English Phonetics and Phonology* has a spectrogram on the cover, of a male voice (mine) saying 'English Phonetics and Phonology', and you can see an explanation of this in the section called 'About the Book' on this website.

It is important to get the terms right, though they are confusing. The picture is a *spectrogram*, while the analysing device used to make it is a *spectrograph.*

START ~ *INDEX*

spreading (lip)

The quality of many sounds can be modified by changing the shape of the lips; the best known example is lip-rounding (labialisation), but another is lip-spreading, produced by pulling the corners of the mouth away from each other as in a smile. Phonetics books tend to be rather inconsistent about this, sometimes implying that any sound that is not rounded has spread lips, but elsewhere treating lip-spreading as being something different from neutral lip shape (in which there is no special configuration of the lips).

START ~ *INDEX*

stop

This term is often used as if synonymous with plosive. However, some writers on phonetics use it to refer to the class of sounds in which there is complete closure specifically in the oral cavity. In this case, sounds such as **m**, **n** are also stops; more precisely, they are nasal stops.

START ~ *INDEX*

stress

Stress is a large topic and despite the fact that it has been extensively studied for a very long time there remain many areas of disagreement or lack of understanding. To

begin with a basic point, it is almost certainly true that in all languages some syllables are in some sense stronger than other syllables; these are syllables that have the potential to be described as stressed. It is also probably true that the difference between strong and weak syllables is of some linguistic importance in every language – strong and weak syllables do not occur at random. However, languages differ in the linguistic function of such differences: in English, for example, the position of stress can change the meaning of a word, as in the case of 'import' (noun) and 'import' (verb), and so forms part of the phonological composition of the word. It is usually claimed that in the case of French there is no possibility of moving the stress to different syllables except in cases of special emphasis or contrast, since stress (if there is any that can be detected) always falls on the last syllable of a word. In tone languages it is often difficult or impossible for someone who is not a native speaker of the language to identify stress functioning separately from tone: syllables may sound stronger or weaker according to the tone they bear.

It is necessary to consider what factors make a syllable count as stressed. It seems likely that stressed syllables are produced with greater effort than unstressed, and that this effort is manifested in the air pressure generated in the lungs for producing the syllable and also in the articulatory movements in the vocal tract. These effects of stress produce in turn various audible results: one is pitch prominence, in which the stressed syllable stands out from its context (for example, being higher if its unstressed neighbours are low in pitch, or lower if those neighbours are high; often a pitch glide such as a fall or rise is used to give greater pitch prominence); another effect of stress is that stressed syllables tend to be longer – this is very noticeable in English, less so in some other languages; also, stressed syllables tend to be louder than unstressed, though experiments have shown that differences in loudness alone are not very noticeable to most listeners. It has been suggested by many writers that the term accent should be used to refer to some of the manifestations of stress (particularly pitch prominence), but the word, though widely used, never seems to have acquired a distinct meaning of its own.

One of the areas in which there is little agreement is that of levels of stress: some descriptions of languages manage with just two levels (stressed and unstressed), while others use more. In English, one can argue that if one takes the word 'indicator' as an example, the first syllable is the most strongly stressed, the third syllable is the next most strongly stressed and the second and fourth syllables are weakly stressed, or unstressed. This gives us three levels: it is possible to argue for more, though this rarely seems to give any practical benefit.

In terms of its linguistic function, stress is often treated under two different headings: word stress and sentence stress. These two areas are discussed under their separate headings.

START ~ *INDEX*

stress-shift

It quite often happens in English that the stress pattern of a word is different when the word occurs in particular contexts compared with its stress pattern when said in isolation: for example, the word ‘fifteenth’ in isolation is stressed on the second syllable, but in ‘fifteenth place’ the stress is on the first syllable. This also happens in place names: the name ‘Wolverhampton’ is stressed on the third syllable, but in the name of the football team ‘Wolverhampton Wanderers’ the stress is usually found on the first syllable. This is known as stress-shift. Explanations by proponents of metrical phonology have suggested that the shift is made in order to avoid two strong stresses coming close together and to preserve the rhythmical regularity of their speech, but such explanations, though attractive, do not have any experimental or scientific justification. English speakers are quite capable of producing strong stresses next to each other when appropriate.

START ~ *INDEX*

stress-timing

It is sometimes claimed that different languages and dialects have different types of rhythm. Stress-timed rhythm is one of these rhythmical types, and is said to be characterised by a tendency for stressed syllables to occur at equal intervals of time. (See rhythm, isochrony, foot, syllable-timing.)

START ~ *INDEX*

stricture

In classifying speech sounds it is necessary to have a clear idea of the degree to which the flow of air is obstructed in the production of the sound. In the case of most vowels there is very little obstruction, but most consonants have a noticeable one; it is usual to refer to this obstruction as a stricture, and the classification of consonants is usually based on the specification of the place of the stricture (e.g. the lips for a bilabial consonant) and the manner of the stricture (e.g. plosive, nasal, fricative).

START ~ *INDEX*

strong form

English has a number of short words which have both strong and weak forms: for example, the word ‘that’ is sometimes pronounced **ðæt** (strong) and sometimes **ðət** (weak). The linguistic context generally determines which one is to be used. The difference between strong and weak forms is explained under weak form.

START ~ INDEX

style

Something which every speaker is able to do is speak in different styles: there are variations in formality ranging from ceremonial and religious styles to intimate communication within a family or a couple; most people are able to adjust their speech to overcome difficult communicating conditions (such as a bad telephone line), and most people know how to tell jokes effectively. But at present we have very little idea what form this knowledge might have in the speaker's mind.

START ~ INDEX

subglottal pressure

Almost all speech sounds depend on having air pushed out of the lungs in order to generate the sound. For voicing to be possible, the pressure of air below the glottis must be higher than the pressure above the glottis (i.e. in the mouth) – otherwise, voicing will not happen. Variation in subglottal pressure is closely related to variations in pitch and stress.

START ~ INDEX

supraglottal

This adjective is used of places in the vocal tract above the glottis (which is inside the larynx). Thus any articulation which involves the pharynx or any other part of the vocal tract above this is supraglottal.

START ~ INDEX

suprasegmental

The term suprasegmental was invented to refer to aspects of sound such as intonation that did not seem to be properties of individual segments (i.e. the vowels and consonants of which speech is composed). The term has tended to be used predominantly by American writers, and much British work has preferred to use the term prosodic instead. There has never been full agreement about how many suprasegmental features are to be found in speech, but pitch, loudness, tempo, rhythm and stress are the most commonly mentioned ones.

START ~ INDEX

Sweet, Henry

Henry Sweet (1845–1912) was a great pioneer of phonetics based in Oxford University. He made extremely important contributions not only to the theory of phonetics (which he described as "the indispensable foundation to the study of language") but also to spelling reform, shorthand, philology, linguistics and language teaching. His best known works include the *Primer of Phonetics*, *The Sounds of English* and *The Practical Study of Languages*.

See Higgins, Henry.

START ~ *INDEX*

syllabic consonant

The great majority of syllables in all languages have a vowel at their centre, and may have one or more consonants preceding and following the vowel (though languages differ greatly in the possible occurrences of consonants in syllables). However, in a few cases we find syllables which contain nothing that could conventionally be classed as a vowel. Sometimes this is a normal state of affairs in a particular language (consider the first syllables of the Czech names 'Brno' and 'Vltava'); in some other languages syllabic consonants appear to arise as a consequence of a weak vowel becoming lost. In German, for example, the word 'abend' may be pronounced in slow, careful speech as **abənt** but in more rapid speech as **abn̩t** or **abm̩t**. In English some syllabic consonants appear to have become practically obligatory in present-day speech: words such as 'bottle' and 'button' would not sound acceptable in BBC pronunciation if pronounced **bɒtəl**, **bʌtən** (though these are normal in some other English accents), and are instead pronounced **bɒtl̩**, **bʌtn̩**. In many other cases in English it appears to be possible either to pronounce **m, n, ŋ, l, r** as syllabic consonants or to pronounce them with a preceding vowel, as in 'open' **əʊpn̩** or **əʊpən**, 'orderly' **ɔːdl̩i** or **ɔːdəli**, 'history' **hɪstr̩i** or **hɪstəri**. The matter is more confusing because of the fact that speakers do not agree in their intuitions about whether a consonant (particularly **l**) is syllabic or not: while most would agree that, for example, 'cuddle' and 'cycle' are disyllabic (i.e. contain two syllables), 'cuddly' and 'cycling' are disyllabic for some people (and therefore do not contain a syllabic consonant) while for others they are trisyllabic. More research is needed in this area for English.

In Japanese we find that some consonants appear to be able to stand as syllables by themselves, according to the intuitions of native speakers who are asked to divide speech up into rhythmical beats. See mora.

START ~ *INDEX*

syllable

The syllable is a fundamentally important unit both in phonetics and in phonology. It is a good idea to keep phonetic notions of the syllable separate from phonological ones. Phonetically we can observe that the flow of speech typically consists of an alternation between vowel-like states (where the vocal tract is comparatively open and unobstructed) and consonant-like states where some obstruction to the airflow is made. Silence and pause are to be regarded as being of consonantal type in this case. So from the speech production point of view a syllable consists of a movement from a constricted or silent state to a vowel-like state and then back to constricted or silent. From the acoustic point of view, this means that the speech signal shows a series of peaks of energy corresponding to vowel-like states separated by troughs of lower energy (see sonority). However, this view of the syllable appears often not to fit the facts when we look at the phonemic structure of syllables and at speakers' views about them. One of the most difficult areas is that of syllabic consonants.

Phonologists are interested in the structure of the syllable, since there appear to be interesting observations to be made about which phonemes may occur at the beginning, in the middle and at the end of syllables. The study of sequences of phonemes is called phonotactics, and it seems that the phonotactic possibilities of a language are determined by syllabic structure; this means that any sequence of sounds that a native speaker produces can be broken down into syllables without any segments being left over. For example, in 'Their strengths triumphed frequently', we find the rather daunting sequences of consonant phonemes **ŋθstr** and **mftfr**, but using what we know of English phonotactics we can split these clusters into one part that belongs to the end of one syllable and another part that belongs to the beginning of another. Thus the first one can only be divided **ŋθ | str** or **ŋθs | tr** and the second can only be **mft | fr**. Phonological treatments of syllable structure usually call the first part of a syllable the onset, the middle part the peak and the end part the coda; the combination of peak and coda is called the rhyme.

Syllables are claimed to be the most basic unit in speech: every language has syllables, and babies learn to produce syllables before they can manage to say a word of their native language. When a person has a speech disorder, their speech will still display syllabic organisation, and slips of the tongue also show that syllabic regularity tends to be preserved even in "faulty" speech.

START ~ *INDEX*

syllable-timing

Languages in which all syllables tend to have an equal time value in the rhythm of the language are said to be syllable-timed; this tendency is contrasted with stress-timing, where the time between stressed syllables is said to tend to be equal irrespective of the number of unstressed syllables in between. Spanish and French are often claimed to be syllable-timed; many phoneticians, however, doubt whether any language is truly syllable-timed.

START ~ INDEX

symbol

One of the most basic activities in phonetics is the use of written symbols to represent speech sounds or particular properties of speech sounds. The use of such symbols for studying and describing English is particularly important, since the spelling system is very far from representing the pronunciation of most words. Many different types of symbol have been tried, but they are almost all based on the idea of having one symbol per phoneme. For many languages it would be perfectly feasible to use a set of syllable symbols instead (though this would not do for English, which would need around 10,000 such symbols). There is an obvious parallel with alphabetic writing, and although phoneticians have in the past experimented with specially-devised symbols which represent phonetic properties in a systematic way, it is the letters of the Roman alphabet that form the basis of the majority of widely-used phonetic symbols, with letters from other writing systems (e.g. Old English **ð**, Greek **θ**) being used to supplement these. Most of the principles for the design of the symbols we use today have been developed by the International Phonetic Association.

START ~ INDEX

synthetic speech

The speech synthesiser is a widely-used tool in speech research: it produces artificial speech, and when the speech synthesis is carefully done the result is indistinguishable from a recording of a human being speaking. Its main use is to produce very finely controlled changes in speech sounds so that listeners' judgements can be experimentally tested. For example, to test if it is true that the most important difference between a pair of words like 'cart' **kɑːt** and 'card' **kɑːd** is that the vowel is shorter before the voiceless final consonant, we can create a large number of syllables resembling **kɑːt** or **kɑːd** in which everything is kept constant except the length of the vowel, and then ask listeners to say whether they hear 'cart' or 'card'. In this way we can map the perceptual boundaries between phonemes. There are many other types of experiment that can be done with synthetic speech.

Synthetic speech is produced by means of computer software. Many phonetics experts have worked on a special application of speech synthesis known as *speech synthesis by rule*, in which a computer is given a written text and must convert it into intelligible speech with appropriate contextual allophones, correct timing and stress and, if possible, appropriate intonation. Synthesis-by-rule systems are useful for such applications as reading machines for blind people, and computerised telephone information systems like "talking timetables". This technology is also used for less serious applications such as talking toys and computer games.

START ~ INDEX

tail

In the analysis of intonation, all syllables that follow the tonic syllable (also called nuclear syllable) up to the tone-unit boundary constitute the tail. Thus in the utterance 'I want two of them', the tail is 'of them'. See *English Phonetics and Phonology* 16.2, page 131.

START ~ INDEX

tap

Many languages have a sound which resembles t or d, being made by a complete closure between the tongue and the alveolar region, but which is very brief and is produced by a sharp upward throw of the tongue blade. As soon as contact is made, the effects of gravity and air pressure cause the tongue to fall again. This tap sound (for which the phonetic symbol is ɾ) is noticeable in Scottish accents as the realisation of the r phoneme, and in American English it is often heard as a (voiced) realisation of t when it occurs after a stressed vowel and before an unstressed one (e.g. the phrase 'getting better' is pronounced gɛɾɪŋ bɛɾə˞). A widely-used alternative way of symbolising this sound is t̬.

In BBC English it used to be quite common to hear a tap for r at the end of a stressed syllable in careful or emphatic speech (e.g. 'very' vɛɾi), though this is less often heard in modern speech. It is now increasingly common to hear the American-style tapped t̬ in England as an allophone of t following a stressed vowel and preceding an unstressed one.

Several varieties of tap are possible: they may be voiced or voiceless – Scottish pre-pausal r is often realised as a voiceless tap, as in 'here' hiɾ̥. They may also be produced with the soft palate lowered, resulting in a nasalised tap which is sometimes heard in the American pronunciation of words like 'mental' meɾ̃əl. A closely related sound is the flap, and the trill also has some similar characteristics.

START ~ INDEX

teeth

The teeth play some important roles in speech. In dental consonants the tip of the tongue is in contact with some of the front teeth. Sometimes this contact is with the inner surface of the upper front teeth, but some speakers place the tongue tip against the lower front teeth and have a secondary contact between the tongue blade and the upper teeth or the alveolar ridge: this happens for some English pronunciations of θ, ð and some French pronunciations of t, d, s, z.

In dental, alveolar and palatal articulations it is necessary to keep a contact between the sides of the tongue and the inside of the upper molar teeth in order to prevent the escape of air.

START ~ INDEX

tempo

Every speaker knows how to speak at different rates, and much research has been done in recent years to study what differences in pronunciation are found between words said in slow speech and the same words produced in fast speech. While some aspects of speaking rate are not linguistically important (e.g. one individual speaker's speaking rate when compared with some other individual's), there is evidence to suggest that we do use such variation contrastively to help to convey something about our attitudes and emotions. This linguistic use of speaking rate is frequently called tempo. In research in this area it is felt necessary to use two different measures: the rate including pauses and hesitations (speaking rate) and the rate with these excluded (articulation rate). Although typing speed is often measured in words per minute, in the study of speech rate it is usual to measure either syllables per second or phonemes per second. Most speakers seem to produce speech at a rate of five or six syllables per second, or ten to twelve phonemes per second.

START ~ INDEX

tense

See lax.

START ~ INDEX

tessitura

This is not a commonly used term in phonetics, but it has been put forward as a technical term (borrowed from singing terminology) to refer to what is sometimes called pitch range. Speakers have their own natural tessitura (the range between the lowest and highest pitch they normally use), but also may extend or shift this for special purposes. The speech of sports commentators provides a lot of suitable research material for this.

START ~ INDEX

timbre (tamber)

It is sometimes useful to have a general word to refer to the quality of a sound, and timbre is sometimes used in that role. It is one of the many words that phonetics has adopted from musical terminology. The word is sometimes spelt tamber.

START ~ *INDEX*

tip

It is useful to divide the tongue up into sections or zones for the purposes of describing its use in articulation. The end of the tongue nearest to the front teeth is called the tip. Sounds made with the tip of the tongue are called apical.

START ~ *INDEX*

ToBI

This is an alternative way of analysing and transcribing intonation which was developed by American researchers in the 1990s. Its basic principle is that intonation can be represented by sequences of high tone (H) and low tone (L). Since most tones in intonation are in fact moving, ToBI links the H and L elements together, so that, for example, a rise is a sequence of L followed by H. The ToBI system was developed and tested to ensure that users could be trained to use it and to be consistent with other users, and in research use it has always been a computer-based system in which the user transcribes the intonation on the computer screen, adding the symbols to the acoustic signal.

Unfortunately, as so often happens with approaches to intonation, a system with a simple basic design gets loaded with more and more detail (often as a result of people publishing papers that point out weaknesses of the system as it stands). Versions of ToBI have been developed for other languages, for other dialects of English and for multi-dialectal comparative studies, and it has to be said that it is now forbiddingly complex for the new user.

A highly simplified account of ToBI can be read in *English Phonetics and Phonology*, Section 17.4, but to get a comprehensive introduction it is best to read tutorial material on the ToBI website at http://www.ling.ohio-state.edu/~tobi/

START ~ *INDEX*

tone

Although this word has a very wide range of meanings and uses in ordinary language, its meaning in phonetics and phonology is quite restricted: it refers to an identifiable movement or level of pitch that is used in a linguistically contrastive way. In some languages (known as tone languages) the linguistic function of tone is to change the meaning of a word: in Mandarin Chinese, for example, ˉ**ma** said with high pitch means 'mother' while ˏ**ma** said on a low rising tone means 'hemp'. In other languages, tone forms the central part of intonation, and the difference between, for example, a rising and a falling tone on a particular word may cause a different interpretation of the sentence in which it occurs. In the case of tone languages it is usual to identify tones as being a property of individual syllables, whereas an intonational tone may be spread over many syllables.

In the analysis of English intonation, tone refers to one of the pitch possibilities for the tonic (or nuclear) syllable, a set usually including fall, rise, fall–rise and rise–fall, though others are suggested by various writers.

START ~ *INDEX*

tone language

As explained in the section on tone, some languages make use of tone for distinguishing word meanings, or, in some cases, for indicating different aspects of grammar. It is probably the case that the majority of the people in the world speak a tone language as their native language, and the peripheral role assigned to the subject of tone by European-language-speaking phoneticians and phonologists shows a regrettable bias that has only recently begun to be corrected. It is conventional (though not strictly accurate) to divide tone languages into contour languages (where the most important distinguishing characteristic of tones is the shape of their pitch contour) and register languages where the height of the pitch is the most important thing. Chinese, and other languages of south-east Asia, are said to be contour languages while most African tone languages (mainly in the South and West of Africa) are classed as register languages. The Amerindian tone languages of Central and South America seem to be difficult to fit into this classification.

Pitch is not the only determining factor in tone: some languages use voice quality differences in a similar way. North Vietnamese, for example, has "creaky" or "glottalized" tones.

START ~ *INDEX*

tone-unit

In the study of intonation it is usual to divide speech into larger units than syllables. If one studies only short sentences said in isolation it may be sufficient to make no

subdivision of the utterance, unless perhaps to mark out rhythmical units such as the foot, but in longer utterances there must be some points at which the analyst marks a break between the end of one pattern and the beginning of the next. These breaks divide speech into tone-units, and are called tone-unit boundaries. If the study of intonation is part of phonology, these boundaries should be identifiable with reference to their effect on pronunciation rather than to grammatical information about word and clause boundaries; statistically, however, we find that in most cases tone-unit boundaries do fall at obvious syntactic boundaries, and it would be rather odd to divide two tone-units in the middle of a phrase. The most obvious factor to look for in trying to establish boundaries is the presence of a pause, and in slow careful speech (e.g. in lectures, sermons and political speeches) this may be done quite regularly. However, it seems that we detect tone-unit boundaries even when the speaker does not make a pause, if there is an identifiable break or discontinuity in the rhythm or in the intonation pattern.

There is evidence that we use a larger number of shorter tone-units in informal conversational speech, and fewer, longer tone units in formal styles.

START ~ *INDEX*

tongue

The tongue is such an important organ for the production of speech that many languages base their word for 'language' on it. It is composed almost entirely of muscle tissue, and the muscles can achieve extraordinary control over the shape and movement of the tongue. The mechanism for protruding the tongue forward out of the mouth between the front teeth, for example, is one which would be very difficult for any engineer to design with no rigid components and no fixed external point to use for pulling.

The tongue is usually subdivided for the purposes of description: the furthest forward section is the tip, and behind this is the blade. The widest part of the tongue is called the front, behind which is the back, which extends past the back teeth and down the forward part of the pharynx. Finally, where the tongue ends and is joined to the rear end of the lower jaw is the root, which has little linguistic function, though it is suggested that this can moved forward and backward to change vowel quality, and that this adjustment is used in some African languages.

The manner of articulation of many consonants depends on the versatility of the tongue. Plosives involving the tongue require an air-tight closure: in the case of those made with the tongue tip or blade, a closure between the forward part of the tongue and the palate or the front teeth is made, as well as one between the sides of the tongue and inner surfaces of the upper molar teeth. Velar and uvular plosives require an air-tight closure between the back of the tongue and the underside of the soft palate. Other articulations include laterals (where the tongue makes central contact but allows air to escape over its sides), and tongue-tip trill, tap and flap. Retroflex consonants are made by curling the tip of the tongue backwards. Finally, the tongue is also used to create an airstream for "click" consonants.

It is sometimes necessary for the tongue to be removed surgically (usually as a result of cancer) in an operation called glossectomy; surprisingly, patients are able to speak intelligibly after this operation when they have had time to practise new ways of articulating.

START ~ *INDEX*

tonic

This adjective is used in the description of intonation. A tonic syllable is one which carries a tone, i.e. has a noticeable degree of prominence. In theories of intonation where only one tone may occur in a tone-unit, the tonic syllable therefore is the point of strongest stress.

START ~ *INDEX*

trachea

This is more popularly known as the "windpipe": it is the tube carrying air which descends from the larynx to the lungs. It runs close to the oesophagus, which carries food and drink down to the stomach. When something that should be going down the oesophagus starts going down the trachea instead, we get rid of it by coughing.

START ~ *INDEX*

transcription

In present-day usage, transcription is the writing down of a spoken utterance using a suitable set of symbols. In its original meaning the word implied converting from one representation (e.g. written text) into another (e.g. phonetic symbols). Transcription exercises are a long-established exercise for teaching phonetics. There are many different types of transcription: the most fundamental division that can be made is between phonemic and phonetic transcription. In the case of the former, the only symbols that may be used are those which represent one of the phonemes of the language, and extra symbols are excluded. In a phonetic transcription the transcriber may use the full range of phonetic symbols if these are required; a narrow phonetic transcription is one which carries a lot of fine detail about the precise phonetic quality of sounds, while a broad phonetic transcription gives a more limited amount of phonetic information.

Many different types of phonemic transcription have been discussed: many of the issues are too complex to go into here, but the fundamental question is whether a phonemic transcription should only represent what can be heard, or whether it should also include sounds that the native speaker feels belong to the words heard, even if

those sounds are not physically present. Take the word 'football', which every native speaker of English can see is made from 'foot' and 'ball': in ordinary speech it is likely that no **t** will be pronounced, though there will probably be a brief **p** sound in its place. Those who favour a more abstract phonemic transcription will say that the word is still phonemically **fʊtbɔːl**, and the bilabial stop is just a bit of allophonic variation that is not worth recording at this level.

START ~ *INDEX*

trill

The parts of the body that are used in speaking (the vocal apparatus) include some "wobbly bits" that can be made to vibrate. When this type of vibration is made as a speech sound, it is called a trill. The possibilities include a bilabial trill, where the lips vibrate (used as a mild insult, this is sometimes called "blowing a raspberry", or, in the USA, a "Bronx Cheer"); a tongue-tip trill (often called a "rolled r") which is produced in many languages for a sound represented alphabetically as 'r' or 'rr', and a uvular trill (which is a rather dramatic way of pronouncing a "uvular r" as found in French, German and many other European languages, most commonly used in acting and singing – Edith Piaf's singing pronunciation is a good example). The vibration of the vocal folds that we normally call voicing is, strictly speaking, another trill, but it is not normally classed with the other trills. Nor is the sound produced by snoring, which is a trill of the soft palate caused by ingressive airflow during breathing in.

When trills occur in languages, they are almost always voiced: it is difficult to explain why this is so.

START ~ *INDEX*

triphthong

A triphthong is a vowel glide with three distinguishable vowel qualities – in other words, it is similar to a diphthong but comprising three rather than two vowel qualities. In English there are said to be five triphthongs, formed by adding **ə** to the diphthongs **eɪ**, **aɪ**, **ɔɪ**, **əʊ**, **aʊ**, these triphthongs are found in the words 'layer' **leɪə**, 'liar' **laɪə**, 'loyal' **lɔɪəl**, 'power' **paʊə**, 'mower' **məʊə**. Things are not this simple, however. There are many other examples of sequences of three vowel qualities, e.g. 'play- off' **pleɪɒf**, 'reopen' **riəʊpən**, so the five listed above must have some special characteristic. One possibility is that speakers hear them as one syllable; this may be the case, but there does not seem to be any clear way of proving this. This is a matter which depends to some extent on the accent: many BBC speakers pronounce these sequences almost as pure vowels (prolongations of the first element of the triphthong), so that the word 'Ireland', for example, sounds like **ɑːlənd**; in Lancashire and Yorkshire accents, on the other hand, the middle vowel (**ɪ** or **ʊ**) is pronounced with such a close vowel quality that it would seem more appropriate to

transcribe the triphthongs with **j** or **w** in the middle (e.g. ‘fire’ **fajə**), emphasising the disyllabic aspect of their pronunciation.

START ~ INDEX

turn-taking

The analysis of conversation has become an important part of linguistic and phonetic research, and one of the major areas to be studied is how participants in a conversation manage to take turns to speak without interrupting each other too much. There are many subtle ways of giving the necessary signals, many of which make use of prosodic features in speech such as a change of rhythm.

START ~ INDEX

upspeak

This is a joking name for a popular style of intonation used mainly by young people, in which a rising tone is used where a fall would be expected. This has the effect of making statements sound like questions. It is often indicated by writers such as novelists and journalists by the use of question marks. For example: “I saw John last night? He was, like, completely out of his mind?”

START ~ INDEX

utterance

The sentence is a unit of grammar, not of phonology, and is often treated as an abstract entity. There is a need for a parallel term that refers to a piece of continuous speech without making implications about its grammatical status, and the term utterance is widely used for this purpose.

START ~ INDEX

uvula

The uvula (a little lump of soft tissue that you can observe in the back of your mouth dangling from the end of your soft palate, if you look in a mirror with your mouth open) is something that the human race could probably manage perfectly well without, but one of the few useful things it does is to act as a place of articulation for a range of consonants articulated in the back of the mouth. There are uvular plosives: the voiceless one **q** is found as a phoneme in many dialects of Arabic, while the

voiced one ɢ is rather more elusive. Uvular fricatives are found quite commonly: German, Hebrew, Dutch and Spanish, for example, have voiceless ones, and French, Arabic and Danish have voiced ones. The uvular nasal ɴ is found in some Inuit languages. The uvula itself moves only when it vibrates in a uvular trill.

START ~ *INDEX*

velaric airstream

Speech sounds are made by moving air (see airstream), and the human speech-production system has a number of ways of making air move. One of the most basic is the sucking mechanism that is used first by babies for feeding, and by humans in later stages of life for such things as sucking liquid through a straw or drawing smoke from a cigarette. The basic mechanism for this is the air-tight closure between the back of the tongue and the soft palate: if the tongue is then retracted, pressure in the oral cavity is lowered and suction results. Consonants produced with this mechanism are called clicks.

START ~ *INDEX*

velarisation

Velarisation is one of the processes known as secondary articulations in which a constriction in the vocal tract is added to the primary constriction which gives a consonant its place of articulation. In the case of English "dark l", the **l** phoneme is articulated with its usual primary constriction in the alveolar region, while the back of the tongue is raised as for an **u** vowel creating a secondary constriction. Arabic has a number of consonant phonemes that are velarised, and are known as "emphatic" consonants.

START ~ *INDEX*

velum, velar

Velum is another name for the soft palate, and *velar* is the adjective corresponding to it. The two terms *velum* and *soft palate* can be used interchangeably in most contexts, but only the word velum lends itself to adjective formation, giving words such as *velar* which is used for the place of articulation of, for example, **k** and **g**, *velic*, used (rarely) for a closure between the upper surface of the velum and the top of the pharynx, and velaric, for the airstream produced in the mouth with a closure between the tongue and the soft palate.

START ~ *INDEX*

vocal cord, vocal fold

The terms ‘vocal cord’ and ‘vocal fold’ are effectively identical, but the latter term is more often used in present-day phonetics. The vocal folds form an essential part of the larynx, and their various states have a number of important linguistic functions. They may be firmly closed to produce what is sometimes called a glottal stop, and while they are closed the larynx may be moved up or down to produce an egressive or ingressive glottalic airstream as used in ejective and implosive consonants. When brought into light contact with each other the vocal folds tend to vibrate if air is forced through them, producing phonation or voicing. This vibration can be made to vary in many ways, resulting in differences in such things as pitch, loudness and voice quality. If a narrow opening is made between the vocal folds, friction noise can result and this is found in whispering and in the glottal fricative **h**. A more widely open glottis is found in most voiceless consonants.

You can read more on this in *English Phonetics and Phonology*, Section 4.1.

START ~ *INDEX*

vocalic

This word is the adjective meaning “vowel-like”, and is the opposite of “consonantal”.

START ~ *INDEX*

vocal tract

It is convenient to think of the passage from the lungs to the lips as a tube (or a pair of tubes if we think of the nasal passages as a separate passage); below the larynx is the trachea, the air passage leading to the lungs. The part above the larynx is called the vocal tract.

START ~ *INDEX*

vocoid

As is explained under contoid, phoneticians have felt the need to invent terms for sounds which have the phonetic characteristics usually attributed to vowels and consonants. Since sounds which are phonetically like consonants may function like phonological vowels, and sounds which are phonetically like vowels may function phonologically as consonants, the terms vocoid and contoid were invented to be used with purely phonetic reference, leaving the terms ‘vowel’ and ‘consonant’ to be used with phonological reference.

START ~ INDEX

voice

This word, with its very widespread use in everyday language, does not really have an agreed technical sense in phonetics. When we wish to refer simply to the vibration of the vocal folds we most frequently use the term voicing, but when we are interested in the quality of the resulting sound we often speak of voice (for example in "voice quality"). In the training of singers, it is always "the voice" that is said to be trained, though of course many of the sounds that we produce when speaking (or singing) are actually voiceless.

START ~ INDEX

voice onset time (VOT)

All languages distinguish between voiced and voiceless consonants, and plosives are the most common consonants to be distinguished in this way. However, this is not a simple matter of a plosive being either completely voiced or completely voiceless: the timing of the voicing in relation to the consonant articulation is very important. In one particular case this is so noticeable that it has for a long time been given its own name: aspiration, in which the beginning of full voicing does not happen until some time after the release of the plosive (usually voiceless). This delay, or lag, has been the subject of much experimental investigation which has led to the development of a scientific measure of voice timing called voice onset time or VOT: the onset of voicing in a plosive may lag behind the plosive release, or it may precede ("lead") it, resulting in a fully or partially voiced plosive. Both can be represented on the VOT scale, one case having positive values and the other negative values; these are usually measured in thousandths of a second (milliseconds, or msec): for example, a Spanish **b** (in which voicing begins early) might have a VOT value of -138 msec, while an English **b** with only a little voicing just before plosive release might have -10; Spanish **p,** which is unaspirated, might have +4 msec while English **p** (aspirated) might have +60 msec.

START ~ INDEX

voice quality

Speakers differ from each other in terms of voice quality (which is the main reason for our being able to recognise individuals' voices even over the telephone), but they also introduce quite a lot of variation into their voices for particular purposes, some of which could be classed as linguistically relevant. A considerable amount of research in this field has been carried out in recent years, and we have a better understanding

of the meaning of such terms as creak, breathy voice and harshness, as well as longer-established terms such as falsetto.

Many descriptions of voice quality have assumed that all the relevant variables are located in the larynx, while above the larynx is the area that is responsible for the quality of individual speech sounds; however, it is now clear that this is an oversimplification, and that the supralaryngeal area is responsible for a number of overall voice quality characteristics, particularly those which can be categorised as articulatory settings.

Good examples of the kinds of use to which voice quality variation may be put in speaking can be heard in television advertising, where "soft" or "breathy" quality tends to be used for advertising cosmetics, toilet paper and detergents; "creaky voice" tends to be associated with products that the advertisers wish to portray as associated with high social class and even snobbery (e.g. expensive sherry and luxury cars), accompanied by an exaggeratedly "posh" accent, while products aimed exclusively at men (e.g. beer, men's deodorants) seem to aim for an exaggeratedly "manly" voice with some harshness.

START ~ *INDEX*

voicing

This term refers to the vibration of the vocal folds, and is also known as phonation. Vowels, nasals and approximants (i.e. sonorants) are usually voiced, though in particular contexts the voicing may be weak or absent. Sounds such as voiceless fricatives and voiceless plosives are the most frequently found sounds that do not have voicing.

START ~ *INDEX*

vowel

Vowels are the class of sound which makes the least obstruction to the flow of air. They are almost always found at the centre of a syllable, and it is rare to find any sound other than a vowel which is able to stand alone as a whole syllable. In phonetic terms, each vowel has a number of properties that distinguish it from other vowels. These include the shape of the lips, which may be rounded (as for an **uː** vowel), neutral (as for **ə**) or spread (as in a smile, or an **iː** vowel – photographers traditionally ask their subjects to say "cheese" **tʃiːz** so that they will seem to be smiling). Secondly, the front, the middle or the back of the tongue may be raised, giving different vowel qualities: the BBC **æ** vowel ('cat') is a front vowel, while the **ɑː** of 'cart' is a back vowel. The tongue (and the lower jaw) may be raised close to the roof of the mouth, or the tongue may be left low in the mouth with the jaw comparatively open. In British phonetics we talk about 'close' and 'open' vowels, whereas American

phoneticians more often talk about ‘high’ and ‘low’ vowels. The meaning is clear in either case.

Vowels also differ in other ways: they may be nasalised by being pronounced with the soft palate lowered as for **n** or **m** – this effect is phonemically contrastive in French, where we find minimal pairs such as très **tʀɛ** (‘very’) and ‘train’ **tʀɛ̃** (‘train’), where the ˜ diacritic indicates nasality. Nasalised vowels are found frequently in English, usually close to nasal consonants: a word like ‘morning’ **mɔːnɪŋ** is likely to have at least partially nasalised vowels throughout the whole word, since the soft palate must be lowered for each of the consonants. Vowels may be voiced, as the great majority are, or voiceless, as happens in some languages: in Portuguese, for example, unstressed vowels in the last syllable of a word are often voiceless and in English the first vowel in ‘perhaps’ or ‘potato’ is often voiceless. Less usual is the case of stressed voiceless vowels, but these are found in French: close vowels, particularly **i** but also the close front rounded **y** and the back rounded **u**, become voiceless for some speakers when they are word-final before a pause (for example ‘oui’ **wi̥** , ‘midi’ **midi̥**, and also ‘entendu’ **ɑ̃tɑ̃dy̥**, ‘tout’ **tu̥**).

It is claimed that in some languages (probably including English) there is a distinction to be made between tense and lax vowels, the former being made with greater force than the latter.

START ~ *INDEX*

weak form

A very important aspect of the dynamics of English pronunciation is that many very common words have not only a strong or full pronunciation (which is used when the word is said in isolation), but also one or more weak forms which are used when the word occurs in certain contexts. Words which have weak forms are, for the most part, function words such as conjunctions (e.g. ‘and’, ‘but’, ‘or’), articles (e.g. ‘a’, ‘an’, ‘the’), pronouns (e.g. ‘she’, ‘he’, ‘her’, ‘him’), prepositions (e.g. ‘for’, ‘to’, ‘at’) and some auxiliary and modal verbs (e.g. ‘do’, ‘must’, ‘should’). Generally the strong form of such words is used when the word is being quoted (e.g. the word ‘and’ is given its strong form in the sentence “We use the word ‘and’ to join clauses”), when it is being contrasted (e.g. ‘for’ in “There are arguments for and against”) and when it is at the end of a sentence (e.g. ‘from’ in “Where did you get it from”). Often the pronunciation of a weak-form word is so different from its strong form that if it were heard in isolation it would be impossible to recognise it: for example, ‘and’ can become **n̩** in ‘us and them’, ‘fish and chips’, and ‘of’ can become **f̩** or **v̩** in ‘of course’. The reason for this is that to someone who knows the language well these words are usually highly predictable in their normal context.

See *English Phonetics and Phonology*, Chapter 12.

START ~ *INDEX*

weak syllable

In English phonology it is possible to identify a type of syllable that is called weak. Such syllables are never stressed, and in rapid speech are sometimes reduced so much that they no longer count as syllables. The majority of weak syllables contain the schwa (**ə**) vowel, but the vowels **i**, **u**, **ɪ** also appear in such syllables. Instead of a vowel, weak syllables may contain syllabic consonants such as **l̩** (as in 'bottle') or **n̩** (as in 'button').

You can read about weak syllables in *English Phonetics and Phonology*, Chapter 9.

START ~ *INDEX*

whisper

Whispering seems to be used all over the world as a way of speaking in conditions where it is necessary to be quiet. Actually, it is not very good for this: for example, whispering does not make voiceless sounds like **s** and **t** any quieter. It seems to wake sleeping babies and adults much more often than does soft voiced speech, and it seems to carry further in places like churches and concert halls. Physiologically, what happens in whispering is that the vocal folds are brought fairly close together until there is a small space between them, and air from the lungs is then forced through the hole to create friction noise which acts as a substitute for the voicing that would normally be produced. A surprising discovery is that when a speaker whispers it is still possible to recognise their intonation, or the tones of tone languages: theoretically, intonation can only result from the vibration of the vocal folds, but it seems that speakers can modify their vocal tracts to produce the effect of intonation by other means.

START ~ *INDEX*

word stress

Not all languages make use of the possibility of using stress on different syllables of a polysyllabic word: in English, however, the stress pattern is an essential component of the phonological form of a word, and learners of English either have to learn the stress pattern of each word, or to learn rules to guide them in how to assign stress correctly (or, quite probably, both). Sentence stress is a different problem, and learners also need to be aware of the phenomenon of stress-shift in which stress moves from one syllable to another in particular contexts.

It is usual to treat each word, when said on its own, as having just one primary (i.e. strongest) stress; if it is a monosyllabic word, then of course there is no more to say. If

the word contains more than one syllable, then other syllables will have other levels of stress, and secondary stress is often found in words like ˌover'whelming (with primary word stress on the ‘whelm’ syllable and secondary stress on the first syllable).

START ~ INDEX

X-ray

In the development of experimental phonetics, radiography has played a very important role and much of what we know about the dimensions and movements of the vocal tract has resulted from the examination of X-ray photos and film. In the last twenty years there has been a sharp decline in the amount of radiographic research in speech since the risk from the radiation is now known to be higher than was suspected before. The technique known as the X-ray Microbeam, developed in Japan and the USA revived this research for some time: a computer controls the direction of a very narrow beam of low-intensity radiation and builds up a picture of articulatory movements through rapid scanning. The equipment was extremely expensive, but produced valuable results. In present-day research, other techniques such as measuring the movements of articulators by means of electromagnetic tracking or magnetic resonance imaging (MRI) are more widely used.

START ~ INDEX

INDEX

front
function word
fundamental frequency

geminate
General American
generative phonology
glide
glottalic
glottal stop, glottalisation
glottis
groove
guttural

head
hesitation
Higgins, Henry
hoarse
homophone
homorganic

implosive
ingressive
instrumental phonetics
intensity
interdental
International Phonetic Association/Alphabet
intonation
isochrony

Jones, Daniel
juncture

key
kinaesthetic/esia

labial(ised)
labiodental
laminal
larynx
lateral
lax
length
lenis
level
lexicon/al
liaison
lingual
lips
liquid
loudness
low
lungs

manner of articulation
median
metrical phonology
minimal pair
monophthong
mora
motor theory of speech perception

nasal, nasalisation
Network English
neutralisation
nucleus

obstruent
occlusion
oesophagus
onset
open
opposition
oral
Oxford accent

palatalisation
palate, palatal
paralinguistic(s)
passive articulator
pause
peak
perception
pharynx
phatic communion
phonation
phone
phoneme
phonemics
phonetics
phonology
phonotactics
pitch
pitch range
place of articulation
plosion
plosive
polysyllabic
pragmatics
pre-fortis clipping

prominence
pronouncing dictionary
pronunciation
prosody/ic
public school accent
pulmonic
pure vowel

rate
realisation
Received Pronunciation (RP)
reduction
register
release
resonance
retroflex
rhotic/ity
rhyme
rhythm
root (of tongue)
rounding

sandhi
schwa
secondary articulation
segment
semivowel
sentence stress
sibilant
slip of the tongue (speech error)
slit
soft palate
sonorant
sonority
spectrogram, spectrography
spreading (lip)
stop
stress
stress-shift
stress-timing
stricture
strong form
style
subglottal pressure
supraglottal
suprasegmental
Sweet, Henry
syllabic consonant
syllable
syllable-timing
symbol
synthetic speech

tail
tap
teeth
tempo
tense
tessitura
timbre (tamber)
tip
ToBI
tone
tone language
tone-unit
tongue
tonic
trachea
transcription
trill
triphthong
turn-taking

upspeak
utterance
uvula

velaric airstream
velarisation
velum, velar
vocal cord, vocal fold
vocalic
vocal tract
vocoid
voice
voice onset time (VOT)
voice quality
voicing
vowel

weak form
weak syllable
whisper
word stress

X-ray