1 What is language?

It is language, more obviously than anything else, that distinguishes humankind from the rest of the animal world. Humans have also been described as tool-making animals; but language itself is the most remarkable tool that they have invented, and is the one that makes most of the others possible. The most primitive tools, admittedly, may have come earlier than language: the higher apes sometimes use sticks as elementary tools, and even break them for this purpose. But tools of any greater sophistication demand the kind of human co-operation and division of labour which is hardly possible without language. Language, in fact, is the great machine-tool which makes human culture possible.

Other animals, it is true, communicate with one another, or at any rate stimulate one another to action, by means of cries. Many birds utter warning calls at the approach of danger; some animals have mating-calls; apes utter different cries to express anger, fear or pleasure. Some animals use other modes of communication: many have postures that signify submission, to prevent an attack by a rival; hive-bees indicate the direction and distance of honey from the hive by means of the famous bee-dance; dolphins seem to have a communication system which uses both sounds and bodily posture. But these various means of communication differ in important ways from human language. Animals' cries are not *articulate*. This means, basically, that they lack structure. They lack, for example, the kind of structure given by the contrast between vowels and consonants, and the kind of structure that enables us to divide a human utterance into words. We can change an utterance by replacing one word by another: a sentry can say 'Tanks

approaching from the north', or he can change one word and say 'Aircraft approaching from the north' or 'Tanks approaching from the west'; but a bird has a single indivisible alarm-cry, which means 'Danger!' This is why the number of signals that an animal can make is very limited: the Great Tit has about thirty different calls, whereas in human language the number of possible utterances is infinite. It also explains why animal cries are very *general* in meaning. These differences will become clearer if we consider some of the characteristics of human language.

What is language?

A human language is a signalling system. The written language is secondary and derivative. In the history of each individual, speech or signing is learned before writing, and there is good reason for believing that the same was true in the history of the species. There are communities that have speech without writing, but we know of no human community which has a written language without a spoken or signed one.

Vocal sounds

The vocal sounds which provide the materials for a language are produced by the various *speech organs* (see figure 1). The production of sounds requires energy, and this is usually supplied by the diaphragm and the chest muscles, which enable us to send a flow of breath up from the lungs. Some languages use additional sources of energy: it is possible to make clicking noises by muscular movements of the tongue, and popping noises by movements of the cheeks and lips, and such sounds are found in some of the African languages. It is also possible to use air flowing *into* the lungs, i.e. to utilize indrawn breath for the production of speech sounds in very short utterances. In English, however, we usually rely on the *outflow* of air from the lungs, which is modified in various ways by the position and shape of the organs that it passes through before finally emerging at the mouth or nose.

First the air from the lungs passes through the vocal cords, in the larynx. These are rather like a small pair of lips in the windpipe,

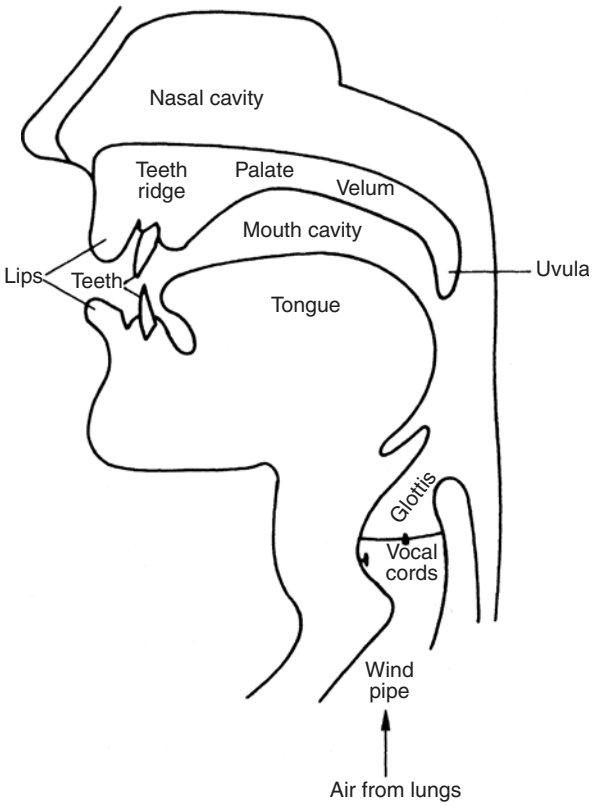

Figure 1 Main speech organs

and we are able to adjust these lips to various positions, from fully closed (when the flow of air is completely blocked) to wide open (when the flow of air is quite unobstructed). In one of the intermediate positions, the vocal cords vibrate as the air passes through, rather like the reed of a bassoon or an oboe, and produce a musical tone called **voice**. We can vary the pitch of our voice (how high or low the tone is on the musical scale), and it changes constantly as we speak, which produces the characteristic melodies of English sentences. The sounds in which voice is used are called **voiced** sounds, but some speech sounds are made with the vocal cords in

the wide open position, and are therefore **voiceless** (or **breathed**). You can detect the presence or absence of voice by covering your ears with your hands: voiced sounds then produce a loud buzzing noise in the head. For example, if you cover your ears firmly and utter a long continuous *v* sound, you will hear voice; if you change it to an *f* sound, the voice disappears. In fact the English *v* and *f* are made in exactly the same way, except that one is voiced and the other voiceless. There are many other similar pairs in English, including *z* and *s*, the *th* of *this* and the *th* of *thing* (for which we can use the symbols [ð] and [θ]), and the consonant sounds in the middle of *pleasure* and of *washer* (for which we can use the symbols [ʒ] and [ʃ]). We can play other tricks with our vocal cords: we can sing, or whisper, or speak falsetto: but the two most important positions for speech are the voiced and the voiceless.

After passing through the vocal cords, the stream of air continues upwards, and passes out through the mouth, or the nose, or both. The most backward part of the roof of the mouth, called the *velum* or the soft palate, can be moved up and down to close or open the entrance to the nasal cavity, while the mouth passage can be blocked by means of the lips or the tongue.

In a **vowel** sound, voice is switched on, and the mouth cavity is left unobstructed, so that the air passes out freely. If the nasal passage is also opened, we get a nasal vowel, like those of French *bon* 'good' or *brun* 'brown', but for the English vowels the nasal passage is normally closed (though some American speakers habitually leave the door ajar and speak with a nasal 'twang'). The quality of a vowel is determined by the position of the tongue, lower jaw and lips, because these can change the shape of the cavity that the air passes through, and different shapes give different resonances. The tongue is the most important. If we raise part of our tongue, we divide the mouth passage into two cavities of different sizes, one at the back and one at the front; the quality of the vowel is, to a great extent, determined by the relative sizes of these two cavities. To describe any vowel sound, therefore, we specify the position of the highest part of the tongue: we can do this in terms of its height (open, half-open, half-close, close) and of its retraction (front, central, or back). A little experimentation with your finger in your mouth, or with a torch and a mirror, will show

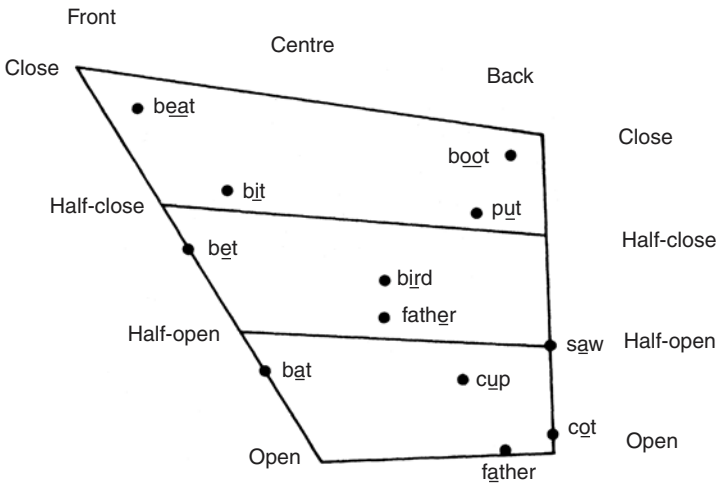

Figure 2 Vowel diagram: typical tongue positions for twelve vowels of present-day English (RP)

you the way your tongue changes position for different vowels. The different positions of the tongue to create different vowel sounds can be shown by means of a *vowel diagram*. This is a conventionalized cross-section of the mouth cavity seen from the left-hand side, on which a vowel is marked as a dot, representing the position of the highest point of the tongue. Figure 2 shows a vowel diagram for twelve English vowels. The accent represented is usually called 'Received Pronunciation' (RP). It was historically the pronunciation of people from families in the south of England who had been educated at public schools such as Eton or Harrow. As we shall see in chapter 9, this became the most prestigious accent in England and is still used as a reference variety and in teaching English, but it has been calculated that a very small percentage of the population actually use this accent today. RP is similar to the general educated accent of south-eastern England, though not quite identical to it.

The quality of a vowel is also affected by the position of the lips, which can be spread wide, held neutral, or rounded more or less tightly. In most forms of English, lip-rounding plays no independent part, for it is an automatic accompaniment of the four backmost

vowels, and the tightness of the rounding varies directly with the closeness of the vowel. You can easily check this with the help of a mirror and the vowel diagram (but it may not be true if you are Scottish or Irish). But this is not so in all languages: in French, the *u* of *lune* is made with a tongue-position similar to that of the *ea* of English *lean*, but is made with rounded lips, which gives it quite a different sound.

Vowels can also differ in length. In fact, the English vowels all have different lengths, but they fall into two broad groups, the long and the short. The short vowels are those heard in *pick*, *peck*, *pack*, *put*, *cut* and *cot*, together with [ə], the short central vowel which is heard in the *er of father* and the *a* of *about*.

The vowel diagram in figure 2 assumes that the vocal organs remain stationary while the vowel is uttered, but this is not always the case, for there are vowels in which the speech organs change their positions in the course of the sound. These are called **glides** or **diphthongs**. An example is the vowel heard in the word *boy*. Here the speech organs begin quite near the position they have for the vowel of *saw*, but almost immediately move towards the position they have for the vowel of *bit*, though they may not go all the way there. During most of the sound, the speech organs are moving, though they may remain in the initial position for a short time before the gliding movement begins. Other English diphthongs are heard in the words *hide*, *house*, *make*, *home*, *hare*, *here* and *poor* (though if you are from parts of the United States, Scotland or northern England you may use a pure vowel in some of these, especially in *home*). On the vowel diagram, diphthongs are represented by arrows, and examples are given in figure 3. Notice that our definition of a diphthong is concerned with *sound*, not with spelling. In popular usage, the *au* of *cause* and the *æ* of *mediæval* are often referred to as diphthongs, but these are not diphthongs in our sense of the word: they are pure vowels which happen to be represented in spelling by two letters (the digraph *au* and the ligature *æ*). Conversely, a diphthong may be represented in spelling by a single letter, like the *y* of *fly*.

We have spoken of diphthongs as single vowel sounds, not as combinations of two vowel sounds. One good reason for doing so is that a diphthong forms only one syllable, not two. A syllable is a

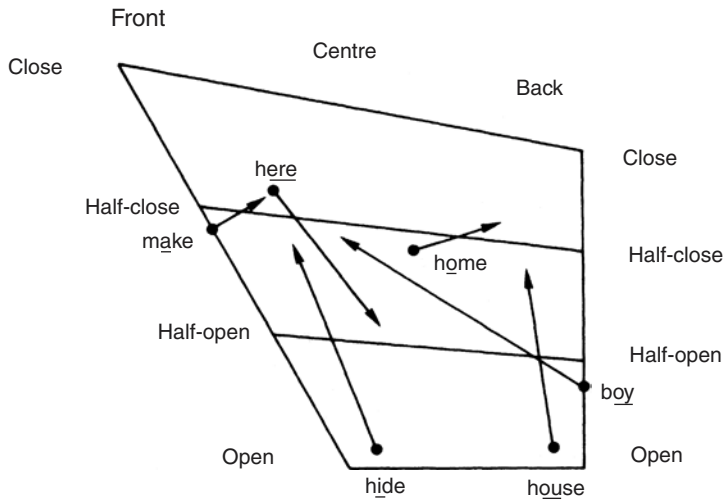

Figure 3 Vowel diagram: six diphthongs of present-day English (RP)

peak of prominence in the chain of utterance. If you could measure the acoustic power output of a speaker as it varies with time, you would find that it goes continually up and down, forming little peaks and valleys: the peaks are syllables. The words *lair* and *here* form only one peak each, and so only one syllable, whereas the words *player* and *newer* are usually pronounced with two peaks and so contain two syllables. It is thus desirable to distinguish between a diphthong (which is one syllable: for instance *f**a**ce*) and a sequence of two vowels (which is two syllables: for instance *hel**iu**m*). Alternatively, a diphthong can be analysed as the combination of a vowel with a semivowel (a non-syllabic glide, like the *y* in *yes*), and this analysis is adopted by many linguists, especially Americans.

In all vowels, the mouth passage is unobstructed. If it is obstructed at any time during the production of a speech sound, the resulting sound will be a **consonant**. In English, there are three main types of consonant: **fricatives, stops** and **sonorants**.

Fricatives are made by narrowing the air passage so much that the stream of air produces audible friction. In *f* and *v*, the constriction is made by pressing the lower lip against the top teeth, while in *th* ([θ] and [ð]) the tip of the tongue is pressed against the

upper teeth. In *s* and *z*, the front of the tongue is pressed against the teeth-ridge (that is, the convex part of the roof of the mouth immediately behind the upper teeth), and the air allowed to flow down a narrow channel in the middle of the tongue, while for [ʃ] and [ʒ] the passage is made wider and flatter. The English *h* consonant can perhaps also be classed as a fricative, but in this case the friction occurs in the glottis, and the mouth passage is completely unobstructed.

In stop consonants, the flow of air from the lungs is completely blocked at some point, and pressure built up behind the blockage; then the blockage is suddenly removed, and there is an outrush of air. The exact sound produced will depend on where and how the blockage is made, and on the speed of the release. In *p* and *b*, the blockage is made by pressing the two lips together. In *t* and *d*, the tip of the tongue is pressed against the teeth-ridge (not against the teeth themselves, as in many other languages). In *k* and *g*, the back part of the tongue is lifted and pressed against the soft palate. In these six sounds, the release is very sudden. In *ch* (as in *church*) and *j* (as in *judge*), which are made in much the same position as *t* and *d*, the release of the blockage is slower, and this gives a different effect, so that *ch* sounds something like a *t* followed very rapidly by a *sh*. Stops with rapid release are called **plosives**, and those with slow release **affricates**. There is also a plosive called the **glottal stop**, in which the blockage is made by complete closure of the vocal cords. This was previously thought to be a feature of Cockney, but, as we shall see in chapter 11, its use is now widespread in many varieties of British English.

In the sonorant consonants, use is made of resonant cavities, as in the vowels, but there is some kind of obstruction in the mouth passage. The English sonorants are the nasals, *m*, *n* and *ng* (as in *sing*), the lateral consonant *l*, and the approximant *r*. In the nasals, the nasal passage is open but the mouth passage is blocked, the blockages being similar to those made for the plosives *b*, *d* and *g* respectively. In the lateral, the centre of the mouth is blocked by the tongue, while the air is allowed to escape down one side, or down both. In English these are all normally voiced, though they may become voiceless or partially voiceless under certain conditions, for example when they follow an *s*. In Welsh, you will hear an *l* sound

(spelt *ll*, as in *Llanelli*) which is regularly voiceless, but this is a fricative consonant rather than a sonorant.

The *r* consonant has various realizations in different varieties of English, but in Received Pronunciation, and in much American English, it is an **approximant**. This is a consonant in which the articulators approach one another, but not closely enough to produce a fricative or a stop. In *r*, the tip of the tongue approaches the teeth-ridge, as if for *d*, but does not make contact, and the tongue is usually curled slightly backward, with the tip raised. In some varieties of English *r* is a trill, in which the tip of the tongue vibrates rapidly, or a flap, in which the tip of the tongue makes a single tap against the teeth-ridge. In some languages, the consonant written as *r* is a different sort of sound: in the best-known varieties of French and German, it is not made with the tip of the tongue, but with the uvula (the small fleshy appendage to the soft palate, which can be seen hanging at the back of the mouth), and in many Indian languages there is a retroflex *r* made by curling the tongue right back and articulating against the roof of the mouth.

In English, sonorant consonants can form syllables. It is sometimes asserted that every syllable must contain a vowel, but this is not so, as can be seen from words like *table* and *button*: in normal pronunciation, each of these has two syllables, the second of which contains no vowel. Syllabic *r* is very common in American speech, in positions where RP instead has the vowel [ə] (called 'schwa'), in words like *perceive*.

With the sonorant consonants we can also group the English semivowels, heard in the *y* of *yes* and the *w* of *wet*. A semivowel is a glide, like a diphthong; but, unlike a diphthong, it does not constitute a syllable. To make the *y* of *yes*, we put our tongue in the position for a short *i* (as in *pin*), and then glide to the position for the following *e*. Similarly, to make *w*, we put our tongue in the position for short *u* (as in *put*), and again glide to the following vowel.

Phonetic symbols

Even in this short account of the English speech sounds, it has already become apparent that it is difficult to discuss the subject without making use of special symbols. We have in English no

single unambiguous spelling to represent the consonant sound in the middle of the word *pleasure* or the first vowel of the word *about*, or to distinguish between the voiced and voiceless *th* of *this* and *thing*, and for this reason we have already introduced the phonetic symbols [ʒ], [ə], [ð] and [θ] to represent these sounds. In the course of this book, we shall use phonetic symbols when they make things simpler and clearer, but shall often use ordinary letter symbols in cases where no ambiguity can arise. When we introduce a new phonetic symbol, we shall of course indicate what it stands for, but for convenience of reference we give below two tables in which all the symbols used are gathered together. In table 1.1, we give a list of symbols which can be used for the transcription of present-day English (Received Pronunciation), together with illustrative examples.

The examples assume Received Pronunciation. Speakers of General American (the most widespread accent in the United States) use the same vowel in *hot* as in *father*, pronounce the /r/ in *air* and *bird*, and lack the centring diphthongs /ɪə/, /ɛə/ and /ʊə/ (the word *here*, for example, being /hɪr/). The symbol [ː] is used to denote vowel-length, so that [ə] is short and [ɜː] long. In General American, however, vowel-length is less significant than in RP, and it is usual to transcribe it without using length-marks, so that for example *tree* is transcribed /tri/.

Similarly, the examples will not fit all speakers in Britain. If you are a northerner, you may well use the same vowel in *put* as in *cut*, where RP makes a distinction. If you are a Scot, you may use the same vowel in *put* as in *goose*. If you come from the West Midlands, or south Lancashire, or the Sheffield area, you may pronounce *sing* as /sɪŋg/, with a [g] after the [ŋ].

Diphthongs are represented by two symbols, the first showing the vowel position in which the diphthong starts, and the second showing the position towards which it glides. So the diphthong in the word *here* begins in about the same position as the [ɪ] of *pin*, and glides towards the central vowel [ə]. We therefore represent it by the notation [ɪə]. The symbol ['] is used to mark stress, and is placed before the syllable that is stressed, so that *admit* is transcribed [əd'mɪt].