

1.1 Language and linguistics

1.1.1 Phonetics as part of linguistics

PHONETICS as a subject of study is nowadays considered to be part of linguistics. But in departments of linguistics in universities it is still a subject with more autonomy than other areas, for various reasons: it is the only section of linguistics which deals almost entirely with the spoken language (the exception being the relationship between sounds and spellings); it is often heavily dependent on instruments and even more dependent on computers than other areas of language study; it depends on data more than other areas of linguistics; and it depends on scaled measurements more than other areas of linguistics.

Nevertheless phonetics does overlap with and inform almost all other areas of linguistics. Phonetics informs MORPHOLOGY, particularly inflexions, e.g. the morphophonemic alternations in plural formation in English as illustrated by the /s/ in in *cats*, the /z/ in *dogs*, the /ɪz/ in *losses* (see §10.4(2)). Phonetics informs SYNTAX, e.g. it highlights word class differences, e.g. between the accent placement in the noun /'tɔ:ment/ and the verb /tɔ:ment/ (see §10.5). Phonetics informs PRAGMATICS, particularly in the way intonation is used, e.g. it shows the 'reservations' in a phrase such as *Well I like 'salmon* (i.e. *but in general I am not keen on fish*) where the ˘ indicates a fall and then rise in pitch on *salmon* (see §11.6.2.6(1)). Moreover phonetics plays a leading part in analyses in SOCIO-LINGUISTICS, including variations in dialect (see, for example, §7.12) and style (see, for example, §12.5).

1.1.2 Phonetics, phonology and phonemics

We talked above about 'phonetics' but we must talk more precisely of 'phonetics and phonology', since this book is concerned with both. The PHONETICS of a language concerns the concrete characteristics (articulatory, acoustic, auditory) of the sounds used in languages while PHONOLOGY concerns how sounds function in a systemic way in a particular language. The traditional approach to phonology

called GB, and at the same time other accents have become acceptable in broadcasting. And in the last half century with the development of English into an international language the inclination in some quarters to regard GB as a monolithic standard has weakened and the need to describe variations within GB and the various types of use of English around the world has increased. So in this book, both in Chapter 7 on Standard and Regional Accents and in the subsections of Chapters 8 and 9 dealing with each consonant and vowel, while some space is devoted to the evolution of GB, more space is now also given to variation within GB and how other standards and dialects differ.

1.3 Learning

1.3.1 Functional load, phonetic cues and redundancy

While we put forward a model of English to be acquired for speaking, it must not be forgotten that a large part of language acquisition depends on listening, both listening to understand and listening to imitate. In order to understand when listening we are more dependent on some contrasts between sounds than on others; we say that some contrasts carry a higher FUNCTIONAL LOAD than others, e.g. the contrast between /i:/ and /ɪ/ carries a higher functional load than that between /u:/ and /ʊ/. Moreover while a contrast between two sounds or sets of sounds may be made in a combination of ways (this being part of the REDUNDANCY of language), some cues are always more important than others: thus the contrasts between /p,t,k/ and /b,d,g/ depend on the cues (i) voicing, (ii) aspiration (breath management), (iii) muscle tension and (iv) the effect on previous (particularly vowel) sounds. But of these cues (ii) and (iv) are far more important to listeners to GB than are (i) and (iii) and of the two (ii) is more important in syllable-initial position and (iv) in syllable-final position. So *pin* is distinguished from *bin* by the aspiration (breath) associated with the /p/ but not with the /b/ and *rope* is distinguished from *robe* by the shorter vowel in *rope*. Throughout this book we attempt to highlight the contrasts with high functional load and the cues which are more important. Besides functional load and the relative importance of specific phonetic cues, there are more general phonetic, grammatical and contextual cues which aid comprehension. If we hear an initial *th* sound [ð] we expect a vowel to follow and we know that some vowels are more likely than others. Or again, the total rhythmic shape of a word may provide an important cue to its recognition, e.g. in *I saw the sheet below* we will know if the final word is *below* or *billow* by the accenting of the first or second syllable. In a discussion about a zoo, involving a statement such as *We saw the lions and tigers* we are predisposed by the context to recognise *lions* even though the word might sound similar to *liars* or *lines*.

Thus teachers and learners of English must remember that communication does not depend on the perfect production and reception of every single element

contrastive segments, 'contrastive' here meaning 'contrasting with other segments to make a change in meaning' (see further in §5.3 below). The phonemic approach to phonology is not the only type of phonological theory but it is the most accessible to those with no training in linguistic theory. The phonemic system of a language is relatable to the writing system: the relationship between the phonemes of a language and the letters used in its writing system is called GRAPHEMICS. A phonemic description also makes it easy to describe the combinatory possibilities of the sounds (the PHONOTACTICS), e.g. that /str/ is a sequence of sounds which begins words in English but not in many other languages. For such reasons the major part of this book is set within phonemic analysis.

1.1.3 Pronunciation and spelling

The word 'pronunciation' indicates that this book is not one about alternative theories of the phonology of English, nor indeed does it seek to justify the use of the phonemic framework other than in terms of the ease of access mentioned in the previous section.

The term 'pronunciation' covers both phonetics and phonemics. Moreover it also encompasses the PROSODY of English, i.e. the 'suprasegmentals' which operate on longer stretches of utterances than sounds or phonemes. Prosody deals with how words and sentences are accented (see Chapter 10), and how pitch, loudness and length work to produce rhythm and intonation (see Chapter 11).

Use of the word 'pronunciation' also indicates that the book makes reference to spellings, particularly in the chapters on vowels and consonants. Thus it gives guidance on how learners can pronounce what they read as well as how they can talk in conversation.

1.2 Change and variation

The central description of this book concerns English English (i.e. English as spoken in England) and more especially the description of a standard accent of English known in the last century as Received Pronunciation (RP) but nowadays better called GENERAL BRITISH (GB). The reason we call it General British is that speakers of GB can be found not only in England but also in Wales and Scotland. There are many speakers who use what we will call REGIONAL GENERAL BRITISH (RGB), i.e. GB with the inclusion of a small number of local characteristics. A standard pronunciation has been evolving at least since the invention of printing in the fifteenth century (see §6.5 and §7.1). The standard started as the accent at the courts of kings and queens, widened to be the accent of the public schools in the nineteenth century, was codified as Received Pronunciation (RP) early in the twentieth century and widened again to become the accent favoured by the BBC in the middle of the twentieth century. Since then the direction of change has been towards dilution of what was called RP. Greater

Englishes and which occasionally incorporates features which are common to particular subcontinental varieties (e.g. /t,d/ as retroflex [t,d] in Bangladesh, India and Pakistan). Secondly there is an INTERNATIONAL ENGLISH which reduces even further the consonant and (particularly) the vowel inventory to something even more easily learnable (e.g. the latter is potentially reduced from 20 to 10 vowels in long and short pairs). Thus the book has changed and evolved from exposition of RP as an almost invariant model for the foreign learner to one which describes GB only as one of a number of models for the learner of English as a second or foreign language.

or speech. To insist, for instance, on exaggeratedly clear articulation for clarity goes beyond the requirements of speech as a means of communication (although it may be necessary in certain situations, e.g. in a crowded room or in a theatrical or operatic production). The potential for redundancy becomes particularly important when considering what sort of simplified model is relevant to the many foreign learners whose need for English is limited to situations where a local English or an 'international' English is acceptable (this theme informs the discussion about choice of alternative models of English in §13.2).

1.3.2 Acquiring English as an L1

Children learning English as a first language will usually only have their family (and a wider circle of friends as they grow) to imitate as they learn the sound system of English, but a knowledge of the sorts of difficulties they face may enable adults to help all learners and in particular those with some sort of speech delay. Many children learning English as an L1 will have mastered the vowel system by the age of three but many will take at least until the age of five to master the system of consonants. Thus little special guidance is usually necessary for learning vowels but often particular guidance will help children to master the consonants, so hints are given in the various subsections in this book about difficulties which young children may have and the sort of guidance which may assist them (see, for example, §§8.7.1, 9.2.3, 9.4.2, 10.6, 11.4, 11.6.6).

1.3.3 Acquiring English as an additional language

When this book was first written learners learning English as an additional language were considered to have only two possible models: the British one, Received Pronunciation (RP), and the American one, GENERAL AMERICAN (GA). This book represented a detailed description of target RP. There was some advice in the sections on individual vowels and consonants about the particular problems which speakers from different L1 backgrounds might face. This advice has been expanded with every edition. Moreover nowadays General British (GB), the successor to RP, is less homogeneous and much more variation within GB is allowed and discussed. Other British accents have become less stigmatised and on the BBC, for instance, almost all regional British accents are heard, certainly in discussion programmes and increasingly even in news presentations; so learners of English as an additional language need some guidance on the accents they are likely to hear (see §7.12 for some of these). In many countries around the world English is used as a *lingua franca*, and in international communication and conferences the common language is almost always English even in situations where none of the participants is a native speaker of English. For these types of communication two new models of English are discussed as targets in Chapter 13. There is firstly an AMALGAM ENGLISH which does not sound like any particular native speaker variety but incorporates the more easily learnable characteristics of various

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

Notes

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

5.1 Speech sounds and linguistic units

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

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

... are able to stand by themselves and are called **WORDS**. In written forms of language, it usually happens that words are separated from each other by spaces, this being a sophisticated convention which is not reflected in speech. (Although Chapter 11 shows there may be some phonetic characteristics marking word boundaries.)

Yet there are meaningful units smaller than the word. The word *boys* may be divided into *boy* and *s* ([z]), where the presence or absence of [z] indicates the plural or singular form; *quickly* may be said to consist of *quick* and the adverbial suffix *-ly*. These smaller sound sequences are known as **MORPHEMES** AND may correspond with words, e.g. *boy*, in which case they may stand alone (= **FREE MORPHEME**), or they may not normally occur other than in association with a word, e.g. *-ly* (= **BOUND MORPHEME**). But segmentation can be made at a still lower level. The word *ran* is also a morpheme; but, if instead of saying [ran] we say [rʌn], we have, by changing an element on a lower level than the morpheme, changed the meaning and function of the word. This basic linguistic element, beyond which it is not necessary to go for practical purposes, is what we have already referred to as a **PHONEME**. A phoneme may, therefore, be thought of as the smallest contrastive linguistic unit which may bring about a change of meaning. Indeed the word 'contrast' is regularly used in linguistics to indicate a change of meaning.

5.3 Phonemes

It is possible to establish the phonemes of a language by means of a process of commutation (= substituting) or the discovery of **MINIMAL PAIRS**, i.e. pairs of words which are different in respect of only one sound segment. The series of words *pin, bin, tin, din, kin, chin, gin, fin, thin, sin, shin, win* supplies us with twelve words which are distinguished simply by a change in the first (consonantal) element of the sound sequence. These elements, or phonemes, are said to be in **CONTRAST OF OPPOSITION**; we may symbolise them as /p, b, t, d, k, ʃ, ʒ, f, θ, s, ʒ, w/. But other sound sequences will show other consonantal oppositions, e.g.

- (1) *tame, dame, game, lame, maim, name*, adding /g, l, m, n/ to our inventory;
- (2) *pot, tot, cot, lot, yacht, hot, rot*, adding /j, h, r/ (the sound of the letter <y> is phonemically transcribed /j/);
- (3) *pie, tie, buy, thigh, thy, vie*, adding /ð, v/ (here the spelling <th> is transcribed /ð/);
- (4) *two, do, who, woo, zoo*, adding /z/.

Such comparative procedures reveal twenty-two consonantal phonemes capable of contrastive function initially in a word.

It is not sufficient, however, to consider merely one position in the word. Possibilities of phonemic opposition have to be investigated in medial and final positions as well as in initial position. If this is done in English, we discover in

... the same examples instances of different sounds may occur, in respect of their function in a language, as the same linguistic unit. In such a pronunciation of *tot* as is noted above, the initial /t/ is described as consisting of:

- (1) a voiceless closure made by the tongue tip and rims against the alveolar ridge and side teeth with air from the lungs building up compression behind the closure—[t];
- (2) a slow release of the closure and the compressed air, so that friction is heard—[s];
- (3) an interval before the beginning of the next sound, during which there is friction in the glottis (and voiceless resonance in the supraglottal cavities)—[h].

The second manifestation of /t/, on the other hand, might have an articulation which could be described phonetically as follows:

- (1) an alveolar stop made as before, but with a simultaneous stop made in the glottis;
- (2) release of the glottal closure but retention of the alveolar closure while the soft palate is lowered and compressed air escapes through the nasal cavity.

The first [t] might be briefly described as a voiceless alveolar plosive, released with affrication and aspiration; the second as an unexploded voiceless alveolar plosive made with a simultaneous glottal stop. These two different articulations function as the same linguistic unit, the first sound occurring in syllable-initial position when accented and the second in syllable-final position (particularly where a pause follows). Such an abstract linguistic unit, which will include sounds of different types, is called a **PHONEME**; the different phonetic realisations of a phoneme are known as its **ALLOPHONES**.

5.2 The linguistic hierarchy

Thus speech and language require different types of unit in their analysis. An utterance, on the concrete speech level, will consist of the continuous physiological activity which results in a continuum of sound; the largest unit will, therefore, be the span of sound occurring between two silences. Within this unit of varying extent it may be possible to find smaller segments. It is from the abstract linguistic level of analysis that we receive guidance as to how the utterance may be usefully segmented in the case of any particular language. We might find, for instance, that an utterance such as 'The boys ran quickly away and were soon out of sight' is spoken without a pause or interruption for breath; on the articulatory level, it consists of one breath-group. But, on the linguistic level, we know that this utterance is capable of being analysed as a sentence consisting of two **CLAUSES**. Moreover, certain extensive sequences occurring within the utterance might be meaningfully replaced by other sound sequences, e.g. *boys* might be replaced by

letter, *leather*, *leisure* or *seater*, *seeker*, *Caesar*, *seizure*. This phoneme /ʒ/ is rare in initial and final positions (e.g. in *genre* and *rouge*). Moreover, in final positions, we do not find /h/ or /r/ in most British speech (the letter <r> being silent in words like *car*, *serve*, *hear*) and it is also questionable whether we should consider /w,j/ as separate, final, contrastive units (see §8.2). We do, however, find one more phoneme that is common in medial and final positions but unknown initially, viz. /ŋ/ cf. *simmer*, *sinner*, *singer* or *some*, *son*, *sung*.

Such an analysis of the consonantal phonemes of English gives us a total of twenty-four phonemes, of which four (/h,r,ʒ,ŋ/) are of *restricted occurrence*—or six, if /w,j/ are not admitted finally. Similar procedures can be used to establish the vowel phonemes of English (see Chapter 8).

5.3.1 Diversity of phonemic solutions

It is important to emphasise the fact that it is frequently possible to make several different statements of the phonemic structure of a language, all of which may be equally valid from a logical standpoint. The solution chosen will be the one which is most convenient as regards the use to which the phonemic analysis is to be put. Thus, one solution might be appropriate when it is a question of teaching a language to a particular group of foreign learners, when similarities and differences between two languages may need to be underlined; another solution might be appropriate if it is a question of using the phonemic analysis as a basis for an orthography, when sociolinguistic considerations (for example, relations with other countries having particular orthographic conventions) have to be taken into account. Even without such considerations, discrepancies in analysis frequently arise in the case of such sound combinations as affricates (e.g. [tʃ,dʒ,ʧ,ʤ,ʧr,ʤr]) and diphthongs (e.g. [eɪ,əʊ,aɪ,əʊ]), which may be treated as single phonemes or combinations of two. Such problems concerning particular English sounds will be dealt with when vowels and consonants are considered in detail.

5.3.2 Distinctive features

Up to now we have obtained an inventory of phonemes for English which is no more than a set of relationships or oppositions. The essence of the phoneme /p/, for instance, is that it is not /t/ or /k/ or /s/, etc. This is a negative definition, which it is desirable to amplify by means of positive information of a phonetic type. Thus, we may say that /p/ is, from a phonetic point of view, characteristically voiceless (compared with voiced /b/); labial (compared with the places of articulation of such sounds as /t/ or /k/); plosive (compared with /f/). The /p/ phoneme may, therefore, be defined positively by stating the combination of DISTINCTIVE FEATURES which identify it within the English phonemic system: voiceless, labial, plosive.

is originally conceived, the distinctive features of a language were stated in articulatory terms using as a basis the phonetic classification of consonants described in the previous chapter. So the distinctive features of English /p/ were voiceless, labial and plosive. Here there are three dimensions of variation: voicing, place and manner. But it was conceded that the distinctive features of a language might involve more or less than three dimensions. For example, in some languages (e.g. in Tamil, a language of South India) voicing is not a distinctive feature (so changing from [p] to [b] does not bring about a change of meaning) and so only place and manner are distinctive. In other languages we may need to state four dimensions of variation. In Hindi not only is voicing (and place and manner) distinctive but aspiration is also separately distinctive from voice; compare, for example, /kaan/ 'ear', /k^haan/ 'mine', /gaan/ 'anthem', /g^haan/ 'quantity'. Such articulatory distinctive features sometimes involve two terms (voiceless vs voiced, aspirated vs unaspirated), sometimes three (e.g. labial /p,b/ vs alveolar /t,d/ vs velar /k,g/ in English) and sometimes more.

Later developments in the theory of distinctive features have involved explaining all the contrasts of a language in terms of BINARY distinctive features and suggesting that there is a set of binary features (involving around 12 or 13 distinctions) which will account for all languages. An apparent three-term distinction like labial vs alveolar vs velar is turned into two features with plus or minus values; using 'coronal' to mean 'made with the blade of the tongue raised above the neutral position' and 'anterior' to mean 'made in front of the hard plate', the English plosives /p,b,t,d,k,g/ are then defined as follows:

	p,b	t,d	k,g
coronal	-	+	-
anterior	+	+	-

In the most well-known set of binary distinctive features,² many features are still articulatory although some are auditory or acoustic (e.g. 'strident').

In this book we use distinctive feature analysis (of the more traditional kind which allows non-binary dimensions) where such analysis is not in doubt and where it is obviously explanatory. This means that we frequently refer to feature analysis when describing the consonants of English, but use it very little when describing the vowels, since almost all distinctive feature analysis in this area is disputed and not always helpful.

5.3.3 Allophones

No two realisations of a phoneme are the same. This is true even when the same word is repeated; thus, when the word *cat* is said twice, there are likely to be slight phonetic variations in the two realisations of the phoneme sequence /k-t-a-t/. Nevertheless, the phonetic similarities between the utterances will probably be more striking than the differences. But variants of the same phoneme will

allophonic variations of their phonemes and will, for instance, say that the various allophones of /l/ which we have discussed are the 'same' sound; [h] and [ɲ], however, they will always consider to be 'different' sounds. When they make a statement of this kind, they are usually referring to the function of the sounds in the language system and can thereby offer helpful, intuitive, information regarding the phonemic organisation of their language. In the case of a language such as English, prejudices induced by the existence of written forms have naturally to be taken into account in evaluating the native speaker's reaction.

5.3.4 Neutralisation

It sometimes happens that a sound may appear to belong to either of two phonemes. In English, examples of this kind are to be found in the plosive series. The contrast between English /p,t,k/ and /b,d,g/ is shown in word-initial position by pairs like *pin/bin*, *team/deem*, *come/gum*. However, following /s/ there is no such contrast. Words beginning /sp-, st-, sk-/ are not contrasted with words beginning /sb-, sd-, sg-/ although a distinction sometimes occurs word-medially, as in *disperse/disperse* and *discussed/discusst* (which suggests a syllable division between the /s/ and the following plosive). In such circumstances we say that the contrast between /p,t,k/ and /b,d,g/, the contrast between voiceless and voiced plosives, is NEUTRALISED following /s/ in word-initial position. Words like *spin*, *steam* and *scar* could equally well be transcribed with /b,d,g/ as with /p,t,k/. Indeed, even though the writing system itself suggests /p,t,k/ (/k/ may be written with <k> or <c>), the sounds which actually occur following /s/ can in some respects be considered closer to /b,d,g/ since the aspiration which generally accompanies /p,t,k/ in initial position is not present after /s/ (although vowels following /p,t,k/ generally start from a higher pitch and vowels following /sp,st,sk/ have this higher pitch, which argues for /p,t,k/).³

Another case of neutralisation concerns the allophones of /m/ and /n/ before /f/ or /v/, in words like *symphony* and *infant*. The nasal consonant in each case is likely to be [m] in fluent speech, i.e. a labiodental sound anticipating the labiodental [f]. Here again, /m/ and /n/ are not opposed, so that the sound could be allocated to either the /m/ or the /n/ phoneme. In practice, since in a slow pronunciation an [m] would tend to be used in *symphony* and an [n] in *infant*, the [m] is usually regarded as an allophone of /m/ in the one case and of /n/ in the other.

5.3.5 Phonemic systems

Statements concerning phonemic categories and allophonic variants can be made in respect of only one variety of one language. It does not follow that, because [j] and [ʃ] are not contrastive in English and belong to the same phoneme, that this is so in other languages—in Russian [j] and [ʃ] constitute separate phonemes.

frequently show consistent phonetic differences, such consistent variants are referred to as ALLOPHONES. We have seen (§5.1) how different the initial and final allophones of /t/ in the word *tot* may be. Or again, the [k] sounds which occur initially in the words *key* and *car* are phonetically clearly different: the first can be felt to be a forward articulation, near the hard palate, whereas the second is made further back on the soft palate. This difference of articulation is brought about by the nature of the following vowel, [i:] having a more advanced articulation than [a:]; the allophonic variation is in this case conditioned by the context. In some varieties of English the two [l] sounds of *lull* [lʌl] show a variation of a different kind. The first [l], the so-called 'clear' [l] with a front vowel resonance, has a quality very different from that of the final 'dark' [ɫ] with a back vowel resonance. Here the difference of quality is related to the position of the phoneme in the word or syllable and depends on whether a vowel or a consonant or a pause follows. It is possible, therefore, to predict in a given language which allophones of a phoneme will occur in any particular context or situation: they are said to be in conditioned variation or COMPLEMENTARY DISTRIBUTION. Statements of complementary distribution can refer to preceding or following sounds (e.g. fronted [k] before front vowels like /i:/ in *key* but retracted [k] before back vowels like /a:/ in *car*); to positions in syllables (plosives are strongly aspirated when initial in accented syllables); or to positions in any grammatical unit, e.g. words (vowels may optionally be preceded by a glottal stop when word-initial) or morphemes (Cockney has a different allophone of /ɔ:/ in morpheme-medial and morpheme-final positions (cf. *board* [bɔəd] vs *bored* [bɔəd]).

Complementary distribution does not take into account those variant realisations of the same phoneme in the same situation which may constitute the difference between two utterances of the same word. When the same speaker produces noticeably different pronunciations of the word *cat* (e.g. by exploding or not exploding the final /t/), the different realisations of the phonemes are said to be in FREE VARIATION. Again, the word *very* may be pronounced [veri] (where the middle consonant is an approximant) or [veri] (where the middle consonant is a tap). The approximant and the tap are here in free variation. Variants in free variation are also allophones (since, like those in complementary distribution, they are not involved in changes of meaning).

It is usually the case that there is some phonetic similarity between the allophones of a phoneme: for example, both the [l] sounds discussed above, as well as the voiceless fricative variety which follows /p/ or /k/ in words such as *please* and *clean*, are lateral articulations. It sometimes happens that two sounds occur in complementary distribution, but are not treated as allophones of the same phoneme because of their total phonetic dissimilarity. This is the case of [h] and [ɲ] in English; they are never significantly opposed, since [h] occurs typically in initial positions in the syllable or word and [ɲ] in final positions. A purely logical arrangement might include these two sounds within the same phoneme, so that *hung* might be transcribed phonemically as either /hʌh/ or /hʌɲ/; but such a solution would ignore the total lack of phonetic similarity and also the feeling

to remember this in the kind of complication in both the system and its realisation, not only for present-day English but also when it is a question of investigating past states of the language. (For a more detailed analysis of variation between dialects, see §7.12.)

5.4 Transcription

The transcription of an utterance (analysed in terms of a linear sequence of sounds) will naturally differ according to whether the aim is to indicate detailed sound values—an ALLOPHONIC (or NARROW) transcription—or the sequence of significant functional elements—a PHONEMIC (or BROAD) transcription.

In the former, an allophonic type of transcription, an attempt is made to include a considerable amount of information concerning our knowledge of articulatory activity or our auditory perception of allophonic features. The *International Phonetic Alphabet* (IPA), shown in Table 1, provides numerous diacritics for a purpose such as this; e.g. the word *titles* might be transcribed as [ˈtɪzːəˈtɪt̪l̪z̪]. Such a notation would show the affrication and aspiration of the initial [t], the fact that the first element of the diphthong is centralised from Cardinal 4 and is long compared with the second element, which is a centralised Cardinal 2, that the [ɪ] has a back vowel resonance and is partly devoiced in its first stage, and that the final [z] is completely devoiced; additionally it is shown that the first syllable is accented. Such a notation is relatively explicit and detailed, but gives no more than an impression of the complexity of the utterance as revealed by various methods of physiological and acoustic investigation. This type of transcription (though usually not as detailed as this) is useful when the focus is on particular details of pronunciation.

In phonemic transcription a different principle operates—namely, that of one symbol per phoneme. Thus a phonemic transcription of the type of English described in this book uses forty-four different symbols (twenty-four consonants and twenty vowels). The basis on which an actual symbol is chosen depends on two further principles: (a) using the phonetic symbols of the most frequent allophones and (b) replacing non-Roman symbols arising from (a) by Roman symbols where these are not already in use. Thus the phonetic symbol for the most common allophone of the phoneme at the beginning of *red* is [r] but the phonemic transcription replaces /r/ by /r/ on the basis of (b). But in the transcription of vowels Romanisation (i.e. the principle under (b)) is not completely carried through in this book, e.g. the transcription uses /n/ and /ɔ:/ for the vowels in *not* and *caught* where it would be possible to use /o/ and /o:/. Transcription of these vowels as used here is called COMPARATIVE PHONEMIC because it allows comparison with vowels in other languages to be made, even though a phonemic transcription is being used. It follows from the principles mentioned above that, even using the IPA, it is possible to construct different sets of symbols for the forty-four symbols of English, although the one used in this book is the most common one in use for the type of English described.

Of again, although /j/ is a phoneme in most varieties of English, in Italian the vowel nasal [ŋ] is an allophone of /n/ which occurs only before /k/ and /g/. Indeed, in English, too, /ŋ/ has not always had phonemic status. Nowadays, [ŋ] might be considered an allophone of /n/ before /k/ and /g/, as in *sink* and *finger*, if it were not for the fact that the /g/ in words such as *sing* was lost about four hundred years ago; once this situation had arisen, a phonemic opposition existed between *sin* and *sing*. In some parts of north-west England, the situation is still the same as it was four hundred years ago, e.g. not only is *sink* pronounced [sɪŋk] but *sing* is pronounced [sɪŋg] and in such dialects [ŋ] can be considered an allophone of /n/. The number of phonemes may differ as between different varieties of the same language. In present-day English spoken in the south of England, the words *cat*, *half*, *cart* contain the phonemes /a/, /ɑ:/ and /ɑ:/ respectively. But one type of Scottish English has only one vowel phoneme for all three words, the words being phonemically /kat, haf, kart/ (the pre-consonantal /r/ being pronounced in Scottish English). Such a dialect of English has one phoneme less than speech in the south of England, since the opposition *Sam/psalm* is lost. On the other hand, this smaller number of phonemes is sometimes counterbalanced by the regular opposition of the first elements of such a pair as *witch/which*, which establishes a phonemic contrast between /w/ and /m/.

It should not be assumed that the phonemic systems of two dialects differ only in having a lesser or greater number of phonemes. The sound sequence [set], i.e. with a vowel in the region of Cardinal 3, may be a realisation of *sat* in one dialect and of *set* in another; the phonemic categories commonly represented as /e, a/, may nevertheless be present in both dialects, all the short front vowels /i, e, a/ being closer in the first dialect than that in the second. Or again, the diphthong [əʊ] is a realisation of the phoneme of *boat* in most of the south of England, but is frequently a realisation of the vowel in *boat* in Cockney; however, the same number of vowel phonemes occurs in both kinds of English.

Moreover, speakers of different dialects may distribute their phonemes differently in words as when a speaker from the north of England pronounces *after*, *bath* and *pass* with /a/ where a speaker from the south of England pronounces them with /ɑ:/. Even speakers of the same dialect (as well as those of different dialects) may distribute the same number of phonemes differently among the words they use. In southern England, some will say *elastic* with /a/ in the second syllable, others /ɑ:/ and some will say /ˈjuːnɪzən/ for *unison*, others /ˈjuːnɪsn/.

Lastly, even individuals are inconsistent; in certain situations, they may change the number of their phonemes, e.g. the occasional use of /m/ in southern England in words like *which*; and they may not always use the same phoneme in a particular word or group of words, e.g. the varying use, in the same person's speech, of /n/ or /ɔ:/ in words like *off*.

To sum up, we may conclude that a phonemic analysis of a number of varieties of one language is likely to reveal: different phonemic systems; different realisations of phonemes; different distribution of phonemes in words (and this last even within the speech of one individual according to the situation). It is important

it must be remembered that a phonemic transcription does not by itself indicate how a sequence is to be pronounced. Only if we know the conventions which tell us how a phoneme is to be realised in different positions do we know its correct pronunciation. Nevertheless a phonemic transcription is particularly useful as a corrective instrument in a language like English where the orthography does not consistently mirror present-day pronunciation.

By now it will have become clear that slant brackets are used for a phonemic transcription, e.g. /tʰɪtʌlz/ while square brackets indicate an allophonic transcription, e.g. [tʰɪˈtʰɪtʌlz]. Sometimes we may wish to show just the phonetic detail of one segment in an otherwise phonemic transcription. In such cases square brackets must still be used, e.g. [tʰɪtʰɪlz]. Slant brackets are only used if the whole sequence is represented phonemically.

5.5 Syllables

The concept of a SYLLABLE as something at a higher level than that of the phoneme or sound segment, yet distinct from that of the word or morpheme, has existed since ancient times. It is significant that most alphabets which now have as their basis the representation of phonemes by letters (however irregular) have reached this state by way of a form of writing which symbolised a group of sounds—a syllabary. Indeed, the basis of the writing of many languages, e.g. that of the Semitic group, remains syllabic. But definitions of the syllable have always presented some difficulty. The best-known approach is that which used to be called a theory of prominence⁴ but is nowadays better known as the sonority hierarchy.

5.5.1 The sonority hierarchy

In any utterance some sounds stand out as more prominent or sonorous than others, i.e. they are felt by listeners to be more sonorous than their neighbours. Another way of judging the sonority of a sound is to imagine its 'carrying power'. A vowel like [a] clearly has more carrying power than a consonant like [z] which in turn has more carrying power than a [b]. Indeed the last sound, a plosive, has virtually no sonority at all unless followed by a vowel. A SONORITY scale or hierarchy can be set up which represents the relative sonority of various classes of sound; while there is some argument over some of the details of such a hierarchy, the main elements are not disputed. One version of the hierarchy is as follows (the most sonorous classes are at the top of the scale):

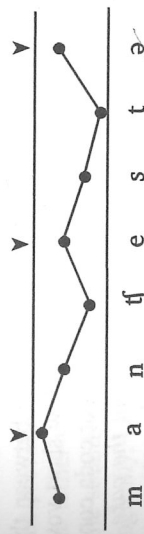
- open vowels
- close vowels
- glides /j, w/
- liquids /l, r/
- nasals

affricates
plosives

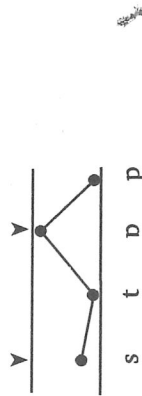
Intermediate vowels are appropriately placed between open and close. Within the last three categories voiced sounds are more sonorous than voiceless sounds.

The terms 'glide' and 'liquid' represent a division of the class 'approximant' (see §4.3.4(5)): glides are short movements away from a vowel-like position (e.g. English /j, w/), while liquid covers sounds like English /l, r/, which have narrowing without friction but are not relatable to vowel sounds. Trills and flaps are usually included with liquids, although this is not agreed by all (it fits well enough for English since trilled [r] and flapped [r] are variants of the usual approximant [ɹ] of GB).

Using the sonority hierarchy we can then draw a contour representing the varying prominences of an utterance, e.g.



The number of syllables in an utterance equates with the number of peaks of sonority, in this case three (marked with arrow heads). This accords with native speakers' intuition. However, there are some cases where contours plotted with the sonority hierarchy do not produce results which accord with our intuition. Many such cases in English involve /s/ in clusters, as, for example, in *stop*:

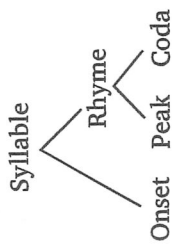


The contour of *stop* implies two syllables, while native speaker intuition is certain that there is only one syllable. This suggests that sounds below a certain level on the hierarchy cannot constitute peaks, i.e. that classes from fricatives downwards cannot constitute peaks in English (though the cut-off point may be drawn at different levels in different languages). Formal statements about the clustering possibilities of English consonants sometimes treat /s/ as an 'appendix'⁵ to syllables which may consequently violate restrictions on their sonority (see §10.10.1).

5.5.2 Syllable constituency

In the previous section reference was made to syllable peaks. In a word like *print* /prɪnt/ the vowel /ɪ/ constitutes the peak and the consonants around it are

sometimes said to constitute the syllabic margins, with those before the peak being called the onset and those after the peak being called the coda. The ONSET, PEAK and CODA of a syllable form a hierarchy of constituents, in which the coda is more closely associated with the peak than with the onset. This can be represented diagrammatically as



Evidence for the greater coherence of the peak and the coda compared with that between the onset and the peak comes from the use of rhyming in verse in which *pat*, *cat*, *sat* rhyme but *pat*, *pad*, *pack* do not rhyme (and hence the use of the term 'rhyme' itself), added to which there are very often restrictions between the peak and the coda in ways in which there are not between the onset and the peak, e.g. the consonant /ŋ/ as in *sing* can only follow short vowels. Moreover, onset consonants are involved in slips of the tongue but coda consonants are not, e.g. *pat the cake* may be produced with a slip to *cat the cake* or even *cat the pake* but slips do not produce *pak the cake* or *pak the cate*.

As implied by the sonority scale discussed above and illustrated in a word like *prints*, onsets generally involve increasing sonority up to the peak (e.g. /r/ is more sonorous than /p/ while the peak /r/ is more sonorous than /r/) while codas generally involve decreasing sonority (e.g. /t/ is less sonorous than /n/ which is less sonorous than the peak /r/). As mentioned above in this section the final /s/ is an exception which does not constitute a syllable despite being of higher sonority than the /t/ which precedes it.

5.5.3 Syllable boundaries

While the onsets and codas of syllables are obviously clearly identifiable at the beginnings and ends of words, dividing word-medial sequences of consonants between coda and onset can be problematical. In many languages such dividing of words into syllables is a relatively straightforward process (e.g. in Bantu languages, in Japanese and in French). In other languages, like English, it is not. The sonority hierarchy tells us how many syllables there are in an utterance by showing us a number of peaks of sonority. Such peaks represent the centres of syllables (usually vowels). Conversely it would seem reasonable for the troughs of sonority to represent the boundaries between syllables. Sounds following the trough would then be in ascending sonority up to the peak and sounds following the peak would be in descending sonority down to the trough. But problems arise because the hierarchy does not tell us whether to place a trough consonant itself with the preceding or the following syllable; an additional

problem is caused by the appended /s/ mentioned in the previous sections. So, for example, syllable division is problematical in words like *funny*, *bitter*, *mattress*, *extra* /'ekstrə/. Various principles can be applied to decide between alternatives: align syllable boundaries with morpheme boundaries where present (the morphemic principle); align syllable boundaries to parallel syllable codas and onsets at the ends and beginnings of words (the phonotactic principle); align syllable boundaries to best predict allophonic variation, e.g. the devoicing of /r/ following /t/. Unfortunately such principles often conflict with one another. A further principle is often invoked in such cases, the maximal onset principle, which assigns consonants to onsets wherever possible and is said to be a universal in languages; but this itself often conflicts with one or more of the principles above. The syllabification of word-medial sequences in English is dealt with in detail in §10.10.3.

5.6 Vowel and consonant

It was seen in the previous chapter that attempts to arrive at a universal phonetic definition of the terms VOWEL and CONSONANT encounter difficulties as regards certain borderline sounds such as [j,w,r] in English. If, however, the syllable is defined phonologically, i.e. from the point of view of distribution of phonemes, a solution can be given to most of these problems. It will be found that the phonemes of a language usually fall into two classes, those which are typically central in the syllable (occurring at the peak) and those which typically occur at the margins (or onsets and codas) of syllables. The term 'vowel' can then be applied to those phonemes having the former function and 'consonant' to those having the latter. The frictionless English sounds /j,w,r/, which, according to most phonetic descriptions, are vowel-like, nevertheless function in the language as consonants, i.e. are marginal in the syllable. A further illustration of the consonantal function of /j,w,r/ is provided by the behaviour of the English articles when they combine with words beginning with these phonemes. *The* is pronounced /ðɪ/ before a vowel and /ðə/ before a consonant; we also have the forms *a* /ə/ or *an* /ən/ according to whether a consonant or vowel follows. Since it is normal to pronounce *the yacht*, *the watch*, *the rabbit* with /ðə/ and to prefix /ə/ to *yacht*, *watch* and *rabbit* rather than /ən/, /j,w,r/ can be treated as belonging to the consonant class of phonemes, despite their vowel-like quality.

The English lateral and nasal sounds are commonly classed phonetically as of the consonantal type because of the complete or partial mouth closure with which they are articulated. From a functional viewpoint, too, they generally behave as consonants, since they are usually marginal in the syllable. Sometimes, however, they operate as a separate peak of sonority, e.g. in *middle* [mɪdl̩] and *button* [bʌtn̩], and thus function in the peaks of syllables. In such occurrences they are referred to as syllabic laterals and nasals.

It is clear that if the elements of the utterance are divided into two categories, some units which are assigned to one class according to phonetic criteria may

can into the other class when it is a question of phonological analysis of English. But the latter prevails when we are doing a phonological analysis of English.

5.7 Prosodic features⁶

As we have seen in Chapter 3, a sound not only has a quality, whose phonetic nature can be described and whose function in the language can be determined, but also has features of length, pitch and loudness. There may be phonemic oppositions in a language based solely or in part on length differences; alternatively differences in the length of a phoneme may relate to different contexts, as when English vowels are generally shorter before voiceless consonants than before voiced consonants.

The features of pitch, length and loudness may contribute to patterns which extend over larger chunks of utterance than the single segment and when used thus are called suprasegmental, or PROSODIC. Pitch is used to make differences of TONE in tone languages, where a syllable or word consisting of the same segmental sequence has different lexical meanings according to the pitch used with it (e.g. in Chinese). Outside tone languages (and even within tone languages, although to a lesser extent) pitch also makes differences of INTONATION whereby different pitch contours produce differences of attitudinal or discursive meaning (discourse here refers to the way successive chunks of utterances are linked together). While tone is a feature of syllables or words, intonation is a feature of phrases or clauses. Some combination of the features of pitch, length and loudness will also produce ACCENT, whereby particular syllables are made to stand out from those around them. There are a number of other prosodic features whose linguistic use is far less understood. These include RHYTHM, the extent to which there is a regular 'beat' in speech (see §11.2); TEMPO (the average conversational tempo of speakers of General British is around four syllables per second);⁷ and VOICE QUALITY, which includes both supralaryngeal settings of the mouth and tongue and laryngeal settings (or phonation types) involving either the vocal cords or the larynx as a whole. Sometimes a voice quality conveys meaning as when a creaky voice indicates boredom; sometimes a quality is appropriate to a situation, e.g. breathy voice is known as 'bedroom voice' and whispery voice as 'library voice'.

5.8 Paralinguistic and extralinguistic features

In addition to prosodic features which spread over more than one segment, there are also PARALINGUISTIC features, which are essentially interruptive rather than co-occurrent. The most common interruptive effect is pause, which functions often as part of the intonation system where it is one of the indicators of an intonational phrase boundary, but at other times functions as a hesitation marker. In the latter case a filled pause is often involved, by some combination of [ʔ] [m] and [ə] in General British, but by other sounds in other dialects and languages (e.g. by an [n] in Russian). Many other paralinguistic effects are more commonly

called VOCALISATIONS: these include single sounds or sequences of sounds like [ʔ] for 'be quiet', [pst] as an attention-getter and [] (a reduplicated dental click), for 'irritation' or 'naughty' (often written *tut-tut*) and various conventionalised types of cough and whistle. Since non-native speakers are likely to pause to think of the right word or grammar far more often than the native speaker, hesitation markers are of particular importance for them. With the acquisition of correct hesitations non-native learners can, if they so wish, dramatically increase their ability to sound like an Englishman.

While prosodic and paralinguistic features are used to convey meaning (although this meaning is in various ways outside the central phonemic system), the term EXTRALINGUISTIC is used for those features over which the speaker has no immediate control. Some of these features may be physical, e.g. sex, age and larynx size. Others may simply be speaker habits, e.g. a particular speaker may always speak with a creaky voice. Yet others may be specific to languages, e.g. speakers of one language may make much more use of an ingressive pulmonic airstream than other languages—this is reported to be so in Finnish—or to particular accents, e.g. Scouse, the dialect of Liverpool, is said to have an adenoïdal quality, produced by retracting and raising the tongue, tightening the pharynx, raising the larynx and keeping the jaws close together, even for open vowels.⁸ Many extralinguistic features are of course ones which may also function prosodically or paralinguistically, e.g. breathy voice may be understood as 'bedroom voice', although particular speakers may have this as a constant characteristic of their speech; and voice qualities involving a raised or lowered larynx, while being habitual, may also be interpreted as 'strained' or 'gloomy' respectively.⁹ For further description of voice quality see under §11.8 below.

Notes

- 1 It is customary to distinguish sound segments from linguistic sound units (phonemes) by using [] to enclose the former and // to enclose the latter.
- 2 Chomsky & Halle (1968).
- 3 Wingate (1982).
- 4 Jones (1918) [1960: 55].
- 5 Giegerich (1992: 147–50).
- 6 For a detailed classification of prosodic and paralinguistic features, see Crystal & Quirk (1964).
- 7 Byrd (1992a) found men speaking 6.2 per cent faster than women.
- 8 Knowles (1987).
- 9 Laver (1974, 1980).