PHONETIC INTRODUCTION

(i) Syllable, vowel and consonant

In any continuous piece of utterance we may perceive certain variations of prominence, characterizing its constituent sounds in such a way that the more prominent alternate with the less prominent in a more or less regular succession. A diagrammatic representation of the opening of the Aeneid, for example, would appear somewhat as follows in terms of relative prominence:

```
| a | r | m | a | u | i | r | u | m | q | e | c | a | n | o |
```

It will be seen that the heights of the ‘peaks’ and the depths of the ‘valleys’ are various; but it is their relative and not their absolute measurement that is important from the standpoint of linguistic structure (omitted for present purposes is the heightening of certain peaks as a consequence of stress or intonation). In the above example there appear seven peaks, with six valleys between them, thus:

peaks \( a - a - i - u - e - a - o \)

valleys \( m - u - r - qu - c - n \)

The number of syllables in a piece generally corresponds to the number of peaks of prominence. The sounds which habitually occur at these peaks we term vowels, and those which occur in the valleys consonants.

The classification is not, however, entirely straightforward. Thus the \( r \) of uirum is in the valley, but that of arma is on the slope; the point here is that in uirum the \( r \) is less prominent than both the preceding \( i \) and the following \( u \), whereas the \( r \) in arma, whilst less prominent than the preceding \( a \), is more prominent than the following \( m \). Similar considerations apply to the \( m \)’s
PHONETIC INTRODUCTION

of arma and of urumque. At this point it will be advisable to consider the nature of the so far undefined ‘prominence’. In Daniel Jones’ words, ‘The prominence of sounds may be due to inherent sonority (carrying power), to length or to stress or to special intonation, or to combinations of these’; so far as concerns the vowel/consonant distinction, inherent sonority is the most generally relevant factor—but there are exceptions. Thus the initial $u$ of urum lies in a valley, whereas the $u$ of the second syllable forms a peak; yet the articulation of both by the tongue and lips is more or less identical. Here the point is that in its position before $i$ the initial $u$ is reduced to a very short duration,\(^1\) with consequent loss of prominence (although its inherent sonority is comparable with that of the $i$); the $u$ of the second syllable, on the other hand, is of high sonority and prominence in contrast with the surrounding $r$ and $m$. Similar principles would apply to the two $i$’s in a word such as rustit.

Sounds which may function either as peaks or as valleys of prominence, whilst classified as vowels in their peak (or ‘nuclear’) function, are generally termed semi-vowels, and classed with the consonants, in their valley (or ‘marginal’) function. Thus Latin $i$ and $u$ may represent both vowels and consonants, and Latin does not distinguish the two functions in writing—unlike e.g. English, which distinguishes $i$ and $u$ from $y$ and $w$.

Finally, it may be noted that two vowels can follow one another as independent peaks, by means of some diminution of energy between them: thus e.g. Latin a-it, faci-at, abi-it, mortu-us, medi-us, tenu-is.

(ii) Consonants

A primary classification of consonants is into the categories of voiced and voiceless. Voiced sounds involve an approximation of the two edges of the vocal cords, so that when air passes through them it sets up a characteristic vibration, known technically as ‘glottal tone’ or voice; voiceless sounds involve a

\(^1\) Probably also with some relaxation of lip-rounding.
CONSONANTS

clear separation of the cords, so that no such vibration occurs. The difference may be exemplified by the English (voiced) z and (voiceless) s. If the cords are closed, the vibration of the former can be clearly heard by the speaker; the vibration can also be felt by placing a finger on the protuberance of the thyroid cartilage (‘Adam’s apple’).

Sounds may be further classified according to the position or organ involved in their articulation. Thus labial (or bilabial) involves the articulation of the two lips (e.g. English p), labio-dental the articulation of the upper teeth and lower lip (e.g. English f), dental the articulation of the tongue-tip and upper teeth (e.g. English th), alveolar the articulation of the tongue-tip and upper gums (e.g. English t), velar the articulation of the back of the tongue and the back of the palate (e.g. English k).

If the tongue or lips form a complete closure, during which air is prevented from passing through the mouth until the closure is released, the resulting sound is termed a stop. Stops are further subdivided into plosives and nasals according to whether the nasal passages are closed or open during the articulation of the stop; thus English has the plosives p, b (bilabial, voiceless and voiced), t, d (alveolar), k, g (velar), and the nasals m (bilabial), n (alveolar), and as ng in sing (velar). In most languages the nasals are all inherently voiced.

The ‘plosion’ of the plosives refers to the effect which is produced when the oral closure is released. If the vocal cords are left open for a brief period after the plosion, producing an audible type of ‘h-sound’, the consonant is termed aspirated —there is clear aspiration, for example, of voiceless plosives at the beginning of words in English and German. In French, on the other hand, the vocal cords are closed almost simultaneously with the oral plosion, and the result is a relatively unaspirated sound.

If the articulating organs are not completely closed, but if the channel between them is so narrow as to cause an audible effect as the air passes through it, the resulting sound is termed a fricative. English examples are f and v (labio-dental,
PHONETIC INTRODUCTION

voiceless and voiced), and s and z (alveolar). The aspirate, 
\( h \), is also sometimes called a ‘glottal fricative’.

If one side of the tongue forms a closure, but the other side 
permits air to flow freely,\(^1\) the result is a lateral consonant, 
such as the English \( l \). Such sounds are sometimes classed with 
the \( r \)-sounds as ‘liquids’ (see p. 32).

(iii) Vowels

Variations of vowel quality are affected primarily by the raising 
of different portions of the tongue’s surface towards the palate, 
and by different degrees of such raising resulting in different 
degrees of aperture between tongue and palate. Vowels may 
thus be classified according to (a) how far forward or open 
they are articulated (i.e. involving more forward or more backward 
areas of the tongue and palate), and (b) how close or open 
they are (i.e. involving greater or lesser raising of the tongue).

The relations of the vowels to one another may then be con- 
veniently represented in terms of a two-dimensional diagram. 
When so represented they tend to fall into a triangular or 
quadilateral pattern,\(^2\) such as:

![Vowel Diagram]

Vowels intermediate between front and back are referred to as 
central, and vowels intermediate between close and open as 
mid (the so-called ‘neutral’ vowel of standard southern British 
English, as at the end of sofa or finger, is a mid-central vowel).

\(^1\) Alternatively there may be a central closure, with air-flow on both sides.

\(^2\) It should be mentioned that such a pattern applies more exactly to the 
acoustic effects of the vowels than to their actual physiological articulation.
VOWELS

Associated with the features already mentioned are various degrees of lip-rounding; generally speaking back vowels are associated with rounding and front vowels with its absence (lip-spread). Thus the English u and i, in e.g. put, pit, are respectively close back rounded and close front unrounded. Sometimes, however, rounding is associated with a front vowel and spreading with a back vowel—thus the French u, German ü, and classical Greek υ are front rounded vowels, and back unrounded vowels occur in some languages.

Vowels are normally articulated with the nasal passages closed (by raising the soft palate or ‘velum’), but if the nasal passages are left open the result is a nasaled vowel (as e.g. in French or, phonetically transcribed [ɔ]).

Diphthongs are formed by articulating a vowel and then, within the same syllable, making a gradual change of articulation (or ‘glide’) in the direction of another vowel. Most commonly, but not inevitably, the first element of a diphthong is more open than the second. Thus the diphthong of English high involves a glide from a towards i, of how from a towards u, and of hay from e towards i.

In many languages vowels fall into two degrees of length, long and short. By and large the difference corresponds to a greater as opposed to a lesser duration—but not invariably so. Other features, such as muscular tension, difference of quality, and tendency to diphthongization, may be at least as important (they are, for example, in distinguishing the so-called ‘short’ vowel of English bit from the so-called ‘long’ vowel of beat).†

(iv) Accent‡

Accent is a general term covering two distinct linguistic functions, and two different modes of implementing these functions. The two functions of accent are termed ‘delimitative’ and ‘culminative’. The first of these, as its name suggests, concerns the fact that in certain languages there are restrictions on the position of the accent within the word such that, given this position, it is possible to infer from it the boundaries of words.
PHONETIC INTRODUCTION

Thus in Czech or Hungarian, words are normally accented by stress on their first syllable; the occurrence of stress in these languages thus indicates the beginning of a word. In Armenian, words are normally stressed on their final syllable, so that the occurrence of stress here indicates the end of a word. In Polish, words are normally stressed on their penultimate syllable, so that the occurrence of stress indicates a word-boundary after the next syllable. The accent of classical Latin is delimitative in a rather complex manner (see p. 83), but that of the majority of Greek dialects only trivially so. In English or Russian, for example, where the accent is free (cf. English *import, impört*; Russian *múka* ‘torment’, *múká* ‘flour’), the accent is of course not delimitative, since it is impossible to predict word-boundaries from it. In such cases the accent has only its ‘culminative’ function of indicating the number of full words in the utterance\(^1\) (a function that is also included in the delimitative accent); the culminative function may in fact be considered as a phonetic expression of the individuality of the word, focused upon a particular portion of it.

Whether delimitative or merely culminative in function, two specific modes of accentuation must be recognized, (a) PITCH, or TONAL accent, and (b) STRESS, or DYNAMIC accent. The tonal accent involves a raising of the voice-pitch at a particular point, and the dynamic accent involves an increase in the muscular effort (primarily by the abdominal muscles).

It is important to distinguish tone from INTONATION. The former refers to the pitch-patterns operative within individual words, whereas ‘intonation’ refers to the pitch-pattern operative over the whole clause or sentence. There may of course be, and there usually is, considerable interaction between these two patterns; thus the pitch-pattern of a given word may vary greatly in accordance with the pitch-pattern of the sentence; such an effect is sometimes referred to as a ‘perturbation’ of the word-tones. Rather similar considerations apply in the case of stress, though one might expect the ‘perturbation’ to be less in

\(^1\) Since it is free, however, it is capable, unlike the delimitative accent, of carrying differences of meaning (as in the English and Russian examples cited).
ACCENT

those languages where the word-stress is strong, as, for example, in English; even here, however, some variation is possible, as for instance in the word fundamental in the two sentences it’s quite fundamental and it’s a fundamental principle. In French, stress is a feature of the word only as an isolate (in which case it falls on the final syllable); in connected speech, however, it is rather a feature of the sense-group.

Naturally the syllables in a word have varying degrees of stress, but by the stressed syllable we mean the syllable which carries the main stress.

Ceteris paribus, stressed sounds produce greater intensity of air-pressure and are perceived as louder than others; but, as already mentioned, the overall prominence of a sound depends upon other features also, such as inherent sonority, duration and intonation; and it is not always easy to disentangle the various causes contributing to its perception.

The distribution of accentual types amongst the languages of the world is a matter for observation rather than prediction. Some more or less universal rules do, however, seem to be emerging. For example, it has been claimed that if a language has significant distinctions of vowel-length, as Latin or Greek, it will not generally have a free dynamic accent; and if a language requires an analysis of its syllabic peaks into ‘mora’ (as classical Greek), its accent is likely to be tonal, but if (as in Latin) no such analysis is demanded, the accent is likely to be dynamic.

(v) Speech and writing

In the study of a ‘dead’ language there is inevitably a main emphasis on the written word. But it is well to remember that writing is secondary to speech, and however much it may deviate from it, has speech as its ultimate basis. The written symbols correspond, in a more or less complete manner, to phonological or grammatical elements of speech; and, as Martinet points out, ‘vocal quality is directly responsible for the linearity of speech and the consequent linearity of script’. It is therefore in a sense misleading to speak of written symbols as
PHONETIC INTRODUCTION

being pronounced—rather is it the other way round, the symbols represent spoken elements. But when, as in the case of Latin today, most utterance consists of reading from a written text, the traditional terminology of ‘pronouncing letters’ may reasonably be tolerated, and is in fact maintained in this book.

In Latin, as in modern European languages, the correspondence is between symbols (letters) and phonological elements, and is much more regular than in some languages, such as English or French or Modern Greek, which notoriously use different symbols or combinations of symbols to indicate the same sound.

It is sometimes stated that an ideal writing-system would have a symbol for every sound—that it would in fact be a kind of ‘visible speech’. Since, however, the number of sounds in a language is infinite, and the ‘same’ sound probably never precisely recurs, this requirement is quite impracticable. It is also unnecessary, as alphabets from earliest times have recognized. The number of symbols can be reduced to manageable proportions without any resultant ambiguity by a process which has long been unconsciously followed, and the theoretical basis of which has been worked out in recent years.

What is required is not one symbol per sound, but one symbol per phoneme. A ‘phoneme’ is a class of similar sounds that are significantly different from other sounds, e.g. the class of t-sounds in English tin, hat, etc., or the class of d-sounds in din, had, etc. The (voiceless) t-phoneme and the (voiced) d-phoneme are different phonemes in English, and so require distinct symbols, because tin has a different meaning from din, hat has a different meaning from had, etc.; in technical terminology, the members of the t and d phonemes are in ‘parallel distribution’, i.e. they can contrast significantly with one another, and with members of other phonemes, in otherwise identical immediate environments ((-)in, ha(-), etc.).

On the other hand, the fact that an initial t in English (as in tin) is more strongly aspirated than a final t (as in hat) is not responsible for any difference of meaning, since the two varieties occur only in different environments, and so cannot contrast
SPEECH AND WRITING

with one another—they are in ‘complementary’ and not parallel distribution. They are thus both members (or ‘allophones’) of the same t-phoneme; only one symbol is required to write them, since the difference in sound is predictable from their environment, i.e. initial or final position as the case may be. It should be noted, however, that the phonemic distribution of sounds varies from language to language; in a language such as classical Greek or modern Hindi, for example, aspirated and unaspirated t-sounds belong to separate phonemes, since the occurrence of one or the other is not predictable from environment and they may contrast significantly (e.g. Greek τελέω ‘stretch’, βέλω ‘strike’; Hindi sāṭ ‘seven’, sāth ‘with’).

The number of phonemes in a language varies; the number of consonants, for example, varies from 8 in Hawaiian, through 24 in English and 32 in Sanskrit, to 80 in the Caucasian Ubykh. Latin, according to the analysis adopted,¹ has from 15 to 18 consonant phonemes in native words.

This ‘phonemic’ principle, then, is an economic principle, ensuring that the minimal number of symbols are used consistent with unambiguous representation of speech. And Latin spelling comes very near to being completely phonemic. The principal shortcoming in this respect concerns the vowels, since no distinction is made in standard orthography between short and long—thus, for example, malus ‘bad’ and uictum (from uino) are not distinguished from malus ‘apple-tree’ and uictum (from uiuo); also no distinction is made between consonantal and vocalic i and u, as in adiect, adiens and insitus, minuit, etc. (uiuiui provides a case of actual ambiguity).

When indicating particular sounds in a phonetic notation it is customary to enclose the symbols in square brackets, e.g. [tʰ] to represent the initial sound of English tin; phonemic symbols, on the other hand, are conventionally set between obliques, e.g. /t/ for the phoneme which includes the initial sound of tin and the final sound of hat. In a book intended primarily for the

¹ Depending upon whether the [ŋ] sounds of magnus, incipio, etc. (see pp. 23, 27) are classed together as a separate phoneme, and whether gu, gu are treated as single phonemes (though represented by digraphs) or as sequences of i, g and consonantal u (see pp. 16, 25).
PHONETIC INTRODUCTION

classical and general reader rather than the technical linguist and phonetician it has seemed desirable to keep phonetic symbols to a minimum. This inevitably involves some theoretical mixing of phonetic, phonemic, and graphic levels of statement, but no practical confusion should thereby be caused. A rigorous separation of levels (as is necessary on e.g. pp. 15, 28) would lead to greater complexity of statement, which would tend to obscure the primary purpose of this study. For the same reason the conventions of the International Phonetic Alphabet have in some cases been modified in the direction of more familiar forms—e.g. by the use of [y] instead of [j] for the palatal semivowel (where the latter could be misleading to the general English reader), and by the use of the macron instead of the colon for vowel length.

Note: Where English equivalents are given for Latin sounds, the reference, unless otherwise stated, is to the standard or ‘Received Pronunciation’ (R.P.) of southern British English. The choice of this form of English as a basis of comparison is made on purely practical grounds. It is impossible to cite examples that will be equally applicable to all nationalities and dialects of English, and one must perforce select a standard; and ‘R.P.’ is by far the best documented and familiar of such standards. Nevertheless, care has been taken to select examples which, so far as possible, will not be positively misleading to speakers of other forms of English.