I
Segments

1.1 Introduction

The fact that words, or more generally stretches of speech, can be divided up into individual segments, or speech-sounds, is familiar to speakers of languages. Thus speakers of English will generally agree that the word *bat* consists of the three sounds 'b', 'a' and 't'. They will further agree that the spelling system of English, i.e. its orthography, does not correspond in a one-to-one fashion to the 'sounds' of the language, so that a word such as *thatch*, although made up of six distinct orthographic symbols, contains only three, or perhaps four, sounds: 'th', 'a' and 'tch' (or perhaps 't' and 'ch'). This discrepancy means that phoneticians and phonologists require a system of transcription for the units of sound analogous to, but different from, that for the units of spelling. Various such systems have been proposed, and are familiar to the user of any dictionary giving the 'pronunciation' of the words of a language. In this book we will generally use the transcription system of the International Phonetic Association (IPA; see Appendix).

The transcription of the sounds of a word is not an entirely straightforward undertaking, and raises interesting theoretical questions in phonology. Thus the transcription of the English word *thatch* requires a decision (implicit or explicit) on the part of the compiler of the system as to whether the sequence *tch* represents two sounds, or **phonological segments** (specifically the two sounds found at the beginning of English *tore* /tɔː/ and *shore* /ʃɔː/),[1] or whether it is to be treated as a single sound, normally referred to as an **affricate**. In systems based on the IPA alphabet, the first option is taken, so that *chore* is represented phonemically as /tʃɔː/ and *thatch* as /θætʃ/, with *ch* or *tch* being

[1] In this book we will in general transcribe English words in the form in which they are realised in RP (Received Pronunciation), the prestige accent of British English. This is a matter of convenience; we are not thereby implying that RP has in any sense a privileged status in terms of its linguistic properties. We will, however, frequently consider other varieties where necessary; in particular we will have occasion to examine data from **rhotic** dialects, i.e. dialects in which postvocalic /r/ is pronounced. RP is non-rhotic, as evidenced by the realisations /tɔː/ and /ʃɔː/ for *tore* and *shore*; compare the pronunciations /toːr/ and /ʃoːr/ (or /toːɹ/ and /ʃoːɹ/) in a rhotic dialect such as Scots English.

represented as a sequence of /t/ and /ʃ/ (although the claim that /t/ and /ʃ/ are more closely related than a normal pair of segments can be indicated by the use of a ligature, as in /θæt͡ʃ/, or, more commonly, by combining the two symbols, as in /θætʃ/). In North American systems, however, such orthographic sequences are generally treated unambiguously as single segments, so that we find transcriptions such as /θæč/.

Notice that the concept of affricate illustrates not only that the relationship between sound and spelling is not entirely straightforward, but also, and perhaps more obviously of relevance for the phonologist, that the relationship between 'phonetic' and 'phonological' representation is also a matter of analysis. From a purely phonetic point of view, the nature of the relationship between the stop and the fricative in the final cluster of English *thatch* does not seem markedly different from that between the stop and the fricative in the final cluster of *hats*: in both cases we have a *phonetic* sequence of stop + fricative, [tʃ] and [ts], respectively (we adopt the usual convention of giving phonetic representations in square brackets, and phonological ones between slant brackets; the line under [t] in [t̪ʃ] denotes retraction of the articulation, in this case to the postalveolar place of articulation of the [ʃ]). However, while the *tch* sequence is commonly treated as an affricate in phonological analysis, phonologists do not generally make a similar claim for the *ts* sequence of *hats*. On the other hand, the phonetically more or less identical cluster in German *Satz* [zats] 'sentence' *is* so treated.

The reasons for these differences (which we will not explore in any detail here) are thus phonological, rather than phonetic, although it is usually claimed that for something to be considered phonologically an affricate it must in any case have the phonetic property of **homorganicity**: i.e. the stop and the fricative must have the same place of articulation, so that [ts] (where both elements are alveolar) and [tʃ] (where both elements are postalveolar) are both conceivable phonological affricates, while a sequence such as [ps] in English *cups* would not be. This claim is associated with the fact that it is just these homorganic sequences which may display a different distribution from 'normal' sequences of consonants. Affricates can generally occur *both* in syllable-initial position *and* in syllable-final position in a language, and thus violate the 'mirror-image' constraint on syllable structure.[2] This constraint states that a consonant cluster which can be syllable-initial in a language cannot be syllable-final, while the same cluster with its consonants in reverse order shows the opposite properties. English is typical in having initial /kl-/ and final /-lk/ (*class*, *sulk*), but not initial */lk-/ or final */-kl/ within a single

[2] We consider syllable structure in Chapter 3.

syllable. Contrast this with the distribution of affricates: /tʃ/ can be both initial and final in English (*chip* /tʃɪp/ and *pitch* /pɪtʃ/), as can /ts/ in German (*Ziel* /tsiːl/ 'goal' and *Satz*). On the other hand, the English sequence /ts/, like other stop + fricative sequences (e.g. /ps/, /ks/), occurs only in syllable-final position (and then almost exclusively as the result of morphological suffixation: e.g. *hats* = HAT + PLURAL).[3]

A full discussion of the status of affricates would take us much further. We return in §1.4 to the status of segments (or sequences) such as these, which exemplify the problem of dealing with what have been referred to as 'complex segments', and we will see that these phenomena have been the trigger for a great deal of interesting work in theories dealing with representation in phonology. Let us first, however, consider a rather more fundamental question regarding phonological representation: does the phonological segment have any internal structure? That is, is there anything which we can say about the way in which sounds behave by assuming some sort of internal structure which we could not say by having segments as the smallest phonological units?

1.2 Evidence for internal structure

It is not difficult to demonstrate that phonological segments in languages can be grouped together, in the sense that particular sets of segments may undergo what seems to be the same kind of **phonological process**. We are assuming here, fairly non-controversially, that it is reasonable to talk about phonological processes, in which a particular segment, or, more importantly here, a group of segments, is affected in some way. These may be either 'events' in the history of a language or relationships holding between the most abstract phonological representation of a segment or group of segments and its surface phonetic realisation.[4]

One such phonological process is that of **nasal place assimilation**, whereby a nasal consonant has the same place of articulation as a following obstruent (i.e. a stop, fricative or affricate). In English, for example, the effects of this process can be identified in various contexts, as in (1):[5]

[3] We indicate morphemes, i.e. minimal syntactic units, by the use of small capitals, as here.

[4] In the context of this book, however, we will beg the question of exactly what is meant by a surface 'phonetic' representation. For practical purposes, the 'surface' representations we consider will be fairly 'shallow' or 'concrete' *phonological* representations. Nevertheless, we will continue to refer to such representations as phonetic. More generally, as we noted in the Preface, we are assuming a model of phonology which is essentially derivational, in the tradition of Chomsky and Halle (1968). We do not adopt here the constraint-based model of Optimality Theory (see, e.g., McCarthy and Prince 1993; Prince and Smolensky 1993; Kager 1999). This is a matter of convenience, however, as we claim that much of what we have to say about the phonological representation of words is independent of whether we adopt a derivational or a constraint-based approach.

[5] The asterisks in (1c) denote that a sequence is ill formed.

(1)	a.	Edinburgh	[ɛmbrə]		
		handbook	[hæmbʊk]		
	b.	unpopular	[ʌmpɒpjələ]		
		unfair	[ʌɱfɛə]		
	c.	camber	[kæmbə]	*[kænbə]	*[kæŋbə]
		canter	[kæntə]	*[kæmtə]	*[kæŋtə]
		canker	[kæŋkə]	*[kæmkə]	*[kænkə]

(1) shows examples of agreement in place of articulation between the nasal and the following obstruent. (1a, b) involve optional assimilations, particularly associated with fast-speech situations: realisations such as /ɛdɪnbʌrə/ and /ʌnpɒpjələ/, which do not show assimilation, also occur, of course. Those in (1b) can be analysed morphologically as involving a prefix ending underlyingly in the alveolar nasal /n/; e.g. UN + FAIR /ʌn + fɛə/. This analysis is supported by the fact that in such cases there are only two possible phonetic realisations of the nasal in the prefix: either as [n] or as the nasal which is homorganic with the following consonant. In addition, if there is no question of a possible assimilation, as in (2), where the following morpheme begins with a vowel or /h/, the only possible realisation is [n]:

(2)	unequal	[ʌniːkwəl]
	unhappy	[ʌnhæpɪ]

The forms in (1c) demonstrate a general constraint on English intervocalic clusters (at least those immediately following a stressed vowel within a single morpheme), which states that a sequence of nasal + stop must be homorganic. These differ from (1a, b), however, in that we are no longer dealing with cases in which, say, the labial nasal can be said to be *derived* from an alveolar nasal, as in [ɛmbrə] or [ʌmpɒpjələ] – there is no possibility of *camber* or *canker* occurring with /n/, as in *[kænbə] or *[kænkə], and there is no internal morphological structure which would lead us to suspect that these words have some kind of prefix CAN-.

Thus the process of nasal place assimilation is instantiated in various ways in English, and indeed in many other languages. However, our concerns here are not primarily with the status of the various different types of examples in the phonology of English; rather they focus on the characterisation of this type of process. In other words, how can we formalise the constraint represented in various ways by the data in (1)? Let us consider first (1a, b), in which we see that a cluster of /n/ followed by a stop may become homorganic in English. If the smallest available phonological units are complete segments, then we might represent the processes as in (3) (for the sake of simplicity, we ignore the case of nasals preceding /f/):

(3) a. /n/ → [m] / __ {/p/, /b/}
b. /n/ → [ŋ] / __ {/k/, /g/}

We use here a traditional **linear** type of notation for phonological rules:[6] the arrow denotes 'is realised as'; the underlying segment is given in slant brackets and its surface phonetic realisation in square brackets; the horizontal line denotes the environment in which the segment affected by the rule occurs, in this case preceding {/p/, /b/}; and the braces denote a set of segments. (3a), then, can be read as: 'Underlying /n/ is realised as phonetic [m] when it precedes either /p/ or /b/.'

There are various objections which can be raised with respect to the formulations of nasal place assimilation in (3). The common core of these objections is that the two parts do not look any more likely to be recurrent phonological rules than, say, any of the processes in (4), which are not likely to occur in any language:

(4) a. /n/ → [m] / __ {/k/, /g/}
b. /n/ → [ŋ] / __ {/p/, /b/}
c. /n/ → [m] / __ {/k/, /d/}
d. /n/ → [l] / __ {/t/, /d/}

Formally, the various rules in (4) are no more or less complex than those in (3), which express recurrent processes – surely an undesirable state of affairs. More particularly, the type of formulation in (3) and (4) is inadequate in two ways. In the first place, the formalism fails to relate the change characterised by a particular rule to the environment in which it occurs. Thus (4a), in which an alveolar nasal becomes labial in the environment of velar stops, is no more difficult to formulate than (3a), in which the same change takes place in the environment of labial stops. Yet (3a) is a natural process of assimilation, while (4a) is not. Secondly, the formalism does not show that the sets of consonants in the environments in (3a, b) are ones that we would expect to find triggering the same kind of change, whereas that in (4c), a set consisting of a voiceless velar stop and a voiced alveolar stop, would be most unlikely to be responsible for the change in (4c) (or, indeed, any other assimilation process). Again, though, (4c) is no more difficult to formulate than any of the other rules in (3) and (4).

This state of affairs clearly arises because we have neither isolated the phonetic properties which are shared by the set of segments involved in the process – nasality in the case of the input and the output (why should the output of (3a) be [m] rather than, say, [l]?); place of articulation in the

[6] See the Preface for a discussion of the difference between linear and non-linear approaches to phonological representation.

case of the output and the environment – nor incorporated them in our rule. In other words, we have failed to take account of the fact that it is the phonetic properties of segments which are responsible for their phonological behaviour, i.e. that phonological segments are not indivisible wholes, but are made up of properties, or, as they are usually referred to, **features**, which to a large extent correspond to the properties familiar from traditional phonetic description.

Furthermore, the fact that a change such as (4c) is an unlikely candidate for an assimilation rule shows that the class of segments triggering the process must share a particular property – in the case of (3a), for example, the property of labiality. A further examination of the phonologies of languages of the world would quickly show that a class of segments like this forms what is referred to as a **natural class**, i.e. a set of segments which *recurrently* participates as a class in phonological processes, such as the ones sketched above. Thus a set of segments which shares some phonetic property or combination of properties, to the exclusion of other sets of segments, forms a natural class.

Let us now identify a number of (ad hoc) phonological features which are relevant here, specifically [nasal], [labial], [alveolar] and [velar]. (Features are by convention enclosed in square brackets.)

We can use these features to write a general rule to characterise the assimilation processes illustrated by (3):

(5) a. $\begin{bmatrix}\text{nasal}\\ \text{alveolar}\end{bmatrix} \rightarrow$ [labial] / __ [labial]

b. $\begin{bmatrix}\text{nasal}\\ \text{alveolar}\end{bmatrix} \rightarrow$ [velar] / __ [velar]

However, we can formulate a rather more general statement about nasal place assimilation in English, which will also incorporate the data in (1c), in which there appears to be no reason to derive [m] and [ŋ] from an underlying /n/. This general statement about the class of nasals is given in (6):

(6) a. [nasal] → [labial] / __ [labial]
b. [nasal] → [alveolar] / __ [alveolar]
c. [nasal] → [velar] / __ [velar]

(6) successfully shows that the rule is a statement about a particular class of segments, nasals, characterised by a single feature which serves to distinguish the class from any other segments in the language. In other words, only nasals undergo the processes characterised by the rule, and no other segments in the language. Furthermore, it shows that the outputs and environments share a feature, namely the feature characterising place of articulation, which

makes just these processes more likely to occur than those in (4), for example. (6) is a non-arbitrary process, then.

Examples like these, which are typical of the way in which phonological processes operate in language, provide evidence for incorporating features in phonological description. It is with the nature of these features, and more particularly the question of whether they are organised in any way in the representation of segments, that we will be largely concerned in the remainder of this chapter.

However, at this point, let us note that the particular formulation in (6) will turn out to be far from adequate on a number of grounds, which do not, however, affect the validity of the points just made. Let us consider here just two of the problems.

(6) appears to consist of three sub-processes, whereas, as we have seen, nasal place assimilation is a single process in English. In traditional linear phonology, it is usual to 'collapse' rules like those in (6), all of which share the same input, to give (7):

(7) $$[\text{nasal}] \rightarrow \left\{ \begin{array}{l} [\text{labial}] \,/\, \_\_\, [\text{labial}] \\ [\text{alveolar}] \,/\, \_\_\, [\text{alveolar}] \\ [\text{velar}] \,/\, \_\_\, [\text{velar}] \end{array} \right\}$$

The three expressions contained in braces are to be seen as alternatives; i.e. nasals are labial before labials, alveolar before alveolars and velar before velars. Thus the 'shared' part of the rule – the input – is mentioned only once.[7]

However, conventions such as that used in (7) still permit the collapse of unrelated rules, as well as rules which apparently belong together. Thus some languages have a rule whereby a nasal consonant becomes voiceless preceding a voiceless (aspirated) consonant. In some dialects of Icelandic, for example, *hempa* /hɛmpʰa/ 'cassock' is realised as [hɛm̥pa], with devoicing of the /m/. There seems to be no formal reason why the rule characterising this process cannot be collapsed with (7), especially as Icelandic also has nasal place assimilation processes:

(8) $$[\text{nasal}] \rightarrow \left\{ \begin{array}{l} [\text{labial}] \,/\, \_\_\, [\text{labial}] \\ [\text{alveolar}] \,/\, \_\_\, [\text{alveolar}] \\ [\text{velar}] \,/\, \_\_\, [\text{velar}] \\ [\text{voiceless}] \,/\, \_\_\, [\text{voiceless}] \end{array} \right\}$$

In other words, we have still failed to show that the features involved in the nasal assimilation process, i.e. [labial], [alveolar] and [velar], are related to

[7] A fuller formulation of the rule in question would also involve reference to other features; we ignore this here, as before.

each other in some way, i.e. that they characterise place of articulation, whereas [voiceless] is not related to any of the other three in this way.

A second problem is that, merely by incorporating features in our rules, rather than the segments of (3) and (4), we have not removed the possibility of formulating what are sometimes referred to as 'crazy rules'. Thus (9) is as easy to formulate as (7):

(9) $$[\text{nasal}] \rightarrow \left\{ \begin{array}{l} [\text{labial}] \,/\, \_\_\, [\text{alveolar}] \\ [\text{alveolar}] \,/\, \_\_\, [\text{velar}] \\ [\text{velar}] \,/\, \_\_\, [\text{labial}] \end{array} \right\}$$

Underlying these criticisms of the formal conventions of linear phonology is the belief that a phonological theory should be as restrictive as possible, in the sense that an ideal system should be able to represent only phonologically natural events and states, and should not be able to characterise unnatural events such as (4) or (9). This belief underpins many **non-linear** alternatives to the formulations above, alternatives which we will begin to consider in §1.4. For the moment, however, we turn in greater detail to the nature of the features which will be required in phonology.

1.3 Phonological features

The idea that segments are made up of phonological features has a long tradition, and received its first comprehensive formalisation in Jakobson *et al.* (1951). The most widely known system is that proposed by Chomsky and Halle (1968; henceforth *SPE*), which differs from the Jakobsonian model in a number of respects, most notably in that the later features are based entirely on articulatory parameters, whereas those of Jakobson *et al.* were defined primarily in terms of acoustic properties. A second important difference involves the fact that many of the Jakobsonian features were relevant to the description and characterisation of both vowels and consonants, while the *SPE* system used largely separate sets of features. Feature theory is not unique to linear approaches to phonology; indeed, much work within non-linear phonology adopts the set of features proposed in the linear framework of *SPE*. However, non-linear phonology typically differs from linear accounts of the segment in incorporating a greater degree of internal structure than a simple list of features, as we shall demonstrate later in this chapter.

As there is a great deal of discussion of individual features available in the literature (e.g. Kenstowicz and Kisseberth 1979; Lass 1984a: chs. 5–6; Keating 1988a; Clements and Hume 1995), we shall not attempt to provide a comprehensive account of the features which would be required to characterise the segments making up the phonological system of English, for example. Rather,

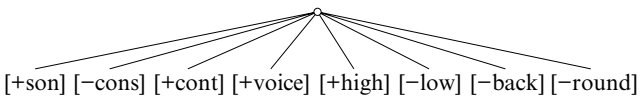
we shall introduce individual features as and when they become relevant, and only provide extensive discussion when necessary. Here the focus will be on how features interact in the representation of the segment, and in particular on the degree of structure required.

In the linear model of *SPE*, segments were viewed as consisting simply of an unordered list of binary features, which were established on grounds similar to those discussed above, i.e. the potential of a feature to define a natural class of segments. The features characterising a segment were organised into a **feature-matrix** in which the features were simply listed along with their value (either + or –) for the segment in question; thus the feature-matrix for the English vowel /iː/, for example, contains the following features, among others:

(10)
$$\begin{bmatrix} +\text{sonorant} \\ -\text{consonantal} \\ +\text{continuant} \\ +\text{voice} \\ +\text{high} \\ -\text{low} \\ -\text{back} \\ -\text{round} \end{bmatrix}$$

Within recent non-linear phonology, in which a more elaborate internal structure has been assigned to the segment, it has become customary to use a different type of formalism to represent the segment. We return in §§1.3.1 and 1.3.5 to the kind of motivation that can be adduced for suggesting a greater degree of structure than is embodied in (10); however, to facilitate comparison, we take the opportunity at this point of providing a 'non-linear' equivalent of (10), in which all of the features making up the segment are ASSOCIATED to a single segmental NODE, represented in (11) by '∘':

(11)



This node is generally referred to as the ROOT NODE – see §1.4.

In (11), as in (10), the features are unordered with respect to each other; any change in this ordering (vertical in the case of the feature-matrix in (10), horizontal in the case of the feature 'tree' in (11)) does not in this case yield anything different from the segment /iː/. We return in due course to the different claims made by the formalisms; in the meantime we devote a little space to the features themselves.

1.3.1 Major class features

The first two features in the matrix in (10) give the 'major class' to which the segment belongs, i.e. vowel; vowels are non-consonantal and, like liquids (i.e. *l* and *r* sounds) and nasals, they are sonorant. In the *SPE* model, sonorancy was defined in articulatory terms, as involving 'a vocal tract configuration in which spontaneous voicing is possible' (*SPE*: 302), but an acoustic definition is equally plausible: sonorant segments have relatively more periodic acoustic energy than non-sonorants (cf. Lass 1984a: 83). By characterising vowels, liquids and nasals as sharing the feature-value [+sonorant], of course, we are making the claim that they form a natural class (cf. §1.2), i.e. that there are phonological processes affecting just this group of segments, and no others. Equally, by assigning the value [–sonorant] to a particular group of segments (the class normally referred to as obstruents, made up of stops, fricatives and affricates), we are claiming that this group too should function as a class. It is not difficult to find processes to demonstrate this; thus the class of obstruents is typically the only class to display 'final devoicing' in many languages, as in various Scottish dialects of English, and Dutch, from which the examples in (12) are taken:

(12)

	singular				*plural*		
a.	rib	'rib'	/rɪb/	[rɪp]	ribben	/rɪbən/	[rɪbə]
	bed	'bed'	/bɛd/	[bɛt]	bedden	/bɛdən/	[bɛdə]
b.	lip	'lip'	/lɪp/	[lɪp]	lippen	/lɪpən/	[lɪpə]
	kat	'cat'	/kɑt/	[kɑt]	katten	/kɑtən/	[kɑtə]
	nek	'neck'	/nɛk/	[nɛk]	nekken	/nɛkən/	[nɛkə]
c.	kam	'comb'	/kɑm/	[kɑm]	kammen	/kɑmən/	[kɑmə]
	man	'man'	/mɑn/	[mɑn]	mannen	/mɑnən/	[mɑnə]
	ring	'ring'	/rɪŋ/	[rɪŋ]	ringen	/rɪŋən/	[rɪŋə]
	nar	'fool'	/nɑr/	[nɑr]	narren	/nɑrən/	[nɑrə]
	bel	'bell'	/bɛl/	[bɛl]	bellen	/bɛlən/	[bɛlə]

The obstruents in the singular forms of (12a, b), which are syllable-final, must be voiceless, irrespective of whether they are voiced (12a) or voiceless (12b) in other contexts, such as in the plural forms, where they occur inter-vocalically. Because the obstruents in (12a) are voiced in other contexts, we assume that they are phonologically, i.e. underlyingly, voiced. In other words, we ascribe their voicelessness in (12a) to the environment in which they occur, i.e. syllable-final position.[8]

[8] Notice that if we had assumed that the obstruents in (12a) were underlyingly voiceless, rather than voiced, we would not have been able to predict whether they would surface intervocalically as voiced (as in *bedden*) or voiceless (as in *katten*). However, it should not be thought that a state of affairs in which an underlying voiceless obstruent becomes voiced intervocalically in a language is impossible; indeed, intervocalic voicing is a very common process.