THE PHONOLOGICAL STRUCTURE OF WORDS

AN INTRODUCTION

COLIN J. EWEN
UNIVERSITY OF LEIDEN

HARRY VAN DER HULST
UNIVERSITY OF CONNECTICUT

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK www.cup.cam.ac.uk
40 West 20th Street, New York, NY 10011–4211, USA www.cup.org
10 Stamford Road, Oakleigh, Melbourne 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain

First published 2001

Printed in the United Kingdom at the University Press, Cambridge

Typeface 9.5/13 pt Times NR [GT]

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication data
Ewen, Colin J.
The phonological structure of words: an introduction / Colin J. Ewen, Harry van der Hulst.
p. cm. – (Cambridge textbooks in linguistics)
Includes bibliographical references and index.
ISBN 0 521 35019 0 (hardback) – ISBN 0 521 35914 7 (paperback)
1. Grammar, Comparative and general – Phonology. I. Hulst, Harry van der. II. Title. III. Series.
P217.E94 2000
414 – dc21 00-025958

ISBN 0 521 350190 hardback
ISBN 0 521 359147 paperback

CONTENTS

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 /θæʧ/). 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, [t̠ʃ] 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 [t̠ʃ] 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 [t̠ʃ] (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. $$\begin{bmatrix}\text{nasal}\\ \text{alveolar}\end{bmatrix} \rightarrow [\text{labial}] \;/\; \_\_ \; [\text{labial}]$$

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

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* /hɛmp^{h}a/ '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)
$$\text{[nasal]} \rightarrow \left\{ \begin{array}{l} \text{[labial] / \_\_ [alveolar]} \\ \text{[alveolar] / \_\_ [velar]} \\ \text{[velar] / \_\_ [labial]} \end{array} \right\}$$

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.

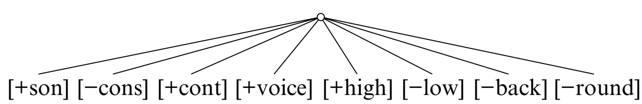
As there is a great deal of discussion of individual features available in the literature (e.g. Kenstowicz and Kisseberth 1979; Lass 1984a: chs. 5–6; Keating 1988a; Clements and Hume 1995), we shall not attempt to provide a comprehensive account of the features which would be required to characterise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

The liquids and nasals in (12c), however, remain voiced in all contexts. Thus the rule of final devoicing in Dutch must make reference to the natural class of non-sonorant consonants, and can be formulated as (13):

(13) *Dutch Final Devoicing*
$[-\text{son}] \rightarrow [-\text{voice}] \,/\, \_\_ \,]_\sigma$

(where we use $]_\sigma$ to denote 'end of syllable').

We can also find cases in Dutch in which [+sonorant] functions as a natural class. Dutch has a process of diminutive formation in which the diminutive suffix is added to a noun. The suffix has a number of different allomorphs, illustrated in (14):

(14)		*noun*				*diminutive*	
	a.	nek	'neck'	/nɛk/	[nɛk]	nekje	[nɛkjə]
		pruik	'wig'	/prœyk/	[prœyk]	pruikje	[prœykjə]
	b.	kam	'comb'	/kɑm/	[kɑm]	kammetje	[kɑmətjə]
	c.	pruim	'plum'	/prœym/	[prœym]	pruimpje	[prœympjə]
		boon	'bean'	/boːn/	[boːn]	boontje	[boːntjə]
		haring	'herring'	/haːrɪŋ/	[haːrɪŋ]	harinkje	[haːrɪŋkjə]
		beer	'bear'	/beːr/	[beːr]	beertje	[beːrtjə]
		uil	'owl'	/œyl/	[œyl]	uiltje	[œyltjə]
	d.	ui	'onion'	/œy/	[œy]	uitje	[œytjə]

The form of the allomorph of the diminutive suffix is predictable according to the phonological form of the noun to which it is attached. Crucially for our purposes here, it takes the form [[stop] + jə] only if the preceding segment is [+sonorant] (a consonant in (14c), a vowel in (d)). (Notice that in (c) the stop assimilates in place to the preceding consonant; the difference between the forms of the suffix in (b) and (c) is due to the nature of the segments preceding the final liquid or nasal.) Thus [+sonorant], just like [–sonorant], can function to identify a natural class of segments.[9]

This phenomenon also provides evidence that natural classes can be defined by a combination of two or more features. (14b), for example, is representative of the larger class in (15):

(15)	*noun*				*diminutive*	
	kam	'comb'	/kɑm/	[kɑm]	kammetje	[kɑmətjə]
	man	'man'	/mɑn/	[mɑn]	mannetje	[mɑnətjə]
	ring	'ring'	/rɪŋ/	[rɪŋ]	ringetje	[rɪŋətjə]
	nar	'fool'	/nɑr/	[nɑr]	narretje	[nɑrətjə]
	bel	'bell'	/bɛl/	[bɛl]	belletje	[bɛlətjə]

[9] In this discussion we are making no assumptions about the underlying form of the diminutive suffix in Dutch, which has been an issue of some debate (see, e.g., Ewen 1978; Trommelen 1983; van der Hulst 1984; Booij 1995 for discussion of diminutive formation). The validity of the particular argument here depends on the assumption that the allomorphs [tjə], [pjə] and [kjə] are derived, rather than underlying.

The class of segments which determine the choice of the [-ətjə] suffix, i.e. the class of nasals and liquids, is defined by the feature combination [+sonorant, +consonantal], together with the nature of the preceding vowel.

Similar evidence from natural class behaviour can be cited in the justification of the various features which we identify in what follows. However, we will only consider such evidence when it is of particular interest for the point we are making.

The next two features in (10), [continuant] and [voice], are used to make further distinctions among the various major classes (vowels, liquids, nasals, obstruents). [+continuant] sounds differ from [–continuant] sounds in not having a complete closure in the oral tract. In the obstruent category ([–sonorant]), fricatives (e.g. /f v ʃ ʒ x ɣ χ/) are [+continuant], in that the stricture of close approximation does not block the airstream entirely, while stops (/p b t d k g q/, etc.) are [–continuant]. Similarly, within the [+sonorant, +consonantal] category, nasals (/m n ŋ/) are [–continuant], in that there is again a complete obstruction of the airstream in the oral cavity (although air does of course escape through the nasal cavity, so that nasal stops can be prolonged), while liquids are [+continuant].[10]

As might be expected, [+voice] sounds are those produced with vibration of the vocal cords; [–voice] are those with no such vibration.

In this discussion, we have implicitly assumed a grouping of features ([consonantal] with [sonorant], which together define 'major classes'; [voice] with [continuant], involved in characterising 'manner of articulation') which is in no way reflected in the matrix in (10). Indeed, the internal structure of the feature-matrix seems quite irrelevant – for example, as we have seen, changing the order in which the features occur in the matrix does not yield a segment which is different in any way. There is, however, a great deal of evidence that grouping of this kind is phonologically relevant: the sets of segments characterised by combinations of particular values of the features within these 'groups' are typically – and recurrently – appealed to in phonological

[10] Notice, though, that it has been claimed that lateral liquids are [–continuant], because, although there is a stricture of open approximation at the sides of the tongue, they also display complete *central* closure. This claim is given weight by the fact that there appear to be processes in some languages in which the lateral liquid forms a natural class with the nasals, as opposed to the non-lateral liquid. Thus Ó Dochartaigh (1978) notes that in some dialects of Scottish Gaelic short vowels diphthongise before /l n/, but lengthen before /r/. Similarly, Clements (1989) demonstrates that in some rhotic dialects of English words like *prince* and *false* may be realised with an epenthetic or inserted [t], i.e. as [prɪnᵗs] and [fɔːlᵗs]. However, between an /r/ and an /s/, as in *nurse* ([nʌrs]), insertion is not possible, again showing that the lateral forms a natural class with the nasal, rather than with the non-lateral liquid. Behaviour like this would support the point of view that laterals may be [–continuant]. In other processes, however, they clearly form a natural class with the non-lateral liquids (e.g. /r/), and so appear to be [+continuant]. We will not be concerned here with how our feature system should capture this apparent anomaly.

processes. Consider again the features [consonantal] and [sonorant], which, it will be recalled, divide up the 'major classes' of segments as in (16):[11]

(16)		O	N/L	V
	[son]	−	+	+
	[cons]	+	+	−

('O' is obstruent, 'N' nasal, 'L' liquid, 'V' vowel). The interaction of these two features is relevant to a number of phonological phenomena. In other words, various combinations of the two features define classes which occur frequently in phonological processes. In addition, the ordering of elements within a syllable is typically determined by these features, so that a vowel ([+sonorant, –consonantal]) forms the peak of a syllable, and an obstruent ([–sonorant, +consonantal]) the margin, with any liquids or nasals ([+sonorant, +consonantal]) being intermediate. Thus /prɪns/ is a well-formed English syllable, while */rpɪsn/ is not.

It is often claimed that the features [sonorant] and [consonantal] determine a **sonority hierarchy** (or sonority scale), and that this hierarchy is reflected in the behaviour of segments in the syllable: the higher the sonority of a segment, the closer it is to the peak of the syllable (see, e.g., Vennemann 1972; Hooper 1976; Kiparsky 1981; Clements 1990). Such hierarchies have a more widespread role, and sometimes involve the other two features already discussed, [continuant] and [voice]. Although we will not discuss this in detail at this point, [+continuant] segments are higher on the sonority hierarchy than [–continuant], and [+voice] segments are higher than [–voice]. This can be established with respect to processes such as the historical 'weakening' or lenition of stops to sonorants in intervocalic position, which involves the gradual assimilation of the features of the stop to those of the surrounding vowels, as illustrated by the development from Pre-Old English to Modern English of the word *own* (from Lass and Anderson 1975: 158):

(17) Pre-OE *[aagan] > OE [aaɣan] > ME [ɔɔwən] > lME [ɔɔn] > MdE /oon/ (/oʊn/, /əʊn/, etc.) 'own'[12]

Each of the changes in (17) represents a step along a lenition hierarchy, which for velars in intervocalic position involves the steps in (18):

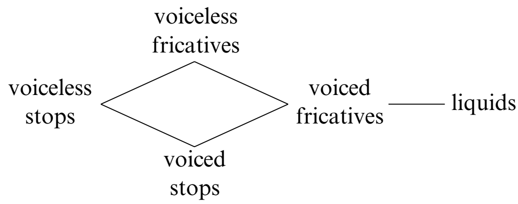
[11] It will be observed that the combination [–sonorant, −consonantal] is also formally possible. Given the definition of [sonorant], it is difficult to see what class of segments might be assigned this representation; [–consonantal] segments (vowels) appear to be inherently sonorant. Chomsky and Halle (1968) in fact assign the combination to [ʔ h], but it is not clear from the phonological behaviour of this set of segments that they should be treated as non-consonantal. We consider an alternative account of [h] in §1.3.5; we will assume here that there are no [–sonorant, –consonantal] segments.

[12] OE = Old English; ME = Middle English; lME = late Middle English; MdE = Modern English.

(18) k > (x *or* g) > ɣ > w

On the basis of processes such as these, Lass and Anderson (1975: 150) establish the general lenition hierarchy in (19):[13]

(19)



Notice that nasals typically do not participate in intervocalic lenition processes, for reasons that need not concern us at present; however, their status with respect to sonority within the class of sonorant consonants can be established on the basis of their behaviour in syllable structure: liquids ([+continuant]) are closer to the syllabic element than nasals ([–continuant]), as is evidenced by syllables such as English *kiln* and *barn* (/bɑrn/ in postvocalic *r*-pronouncing (i.e. rhotic) dialects of English), as opposed to the unacceptable syllables */kɪnl/ and */bɑnr/.

We have now shown that the four features considered so far together form a group with respect to which phonological regularities can be uncovered, as they, and they alone, distinguish the classes involved in sonority-based phenomena, as in (20):

(20)

	voiceless stops	voiced stops	voiceless frics	voiced frics	N	L	V
[son]	−	−	−	−	+	+	+
[cons]	+	+	+	+	+	+	−
[cont]	−	−	+	+	−	+	+
[voice]	−	+	−	+	+	+	+

(Here we ignore oppositions between voiced and voiceless sonorants, i.e. nasals, liquids and vowels.)

1.3.2 Vowel features

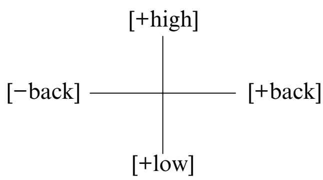
Let us now consider the remaining four features in (10), [high], [low], [back] and [round]. In *SPE*, these features, used primarily to distinguish the vowels of a language, are defined in terms of the position of the highest point of the tongue in the production of a vowel (for [high], [low] and [back]), and the

[13] Lass and Anderson in fact claim that lenition ultimately yields deletion, as illustrated by the development of *own* in (17), possibly via a vowel stage. For the moment, however, we confine the discussion to the consonant-types involved in such processes.

presence or absence of lip-rounding (for [round]). The definitions of the first three features refer to a 'neutral' position for the tongue (roughly the position for [ɛ]), such that [+high] sounds have their closest constriction higher than the neutral position, whereas [−high] sounds do not, and similarly for [+low] vs [−low] and [+back] vs [−back].

A system of this sort essentially treats the vowel features as interpretations of two axes, as in (21):

(21)



The claim inherent in the set of features given above is that languages typically make a three-way opposition on the vertical axis, but only a two-way opposition on the horizontal axis. Thus there is no separate feature [±front] beside [±back], but we do have a separate feature [±low] alongside [±high]. The only way of characterising a three-way opposition on the high–low axis within a binary feature framework is to postulate two features, giving the following possibilities:

(22)

	[−back]		[+back]	
	[−round]	[+round]	[−round]	[+round]
[+high, −low]	i	y	ɯ	u
[−high, −low]	e,ɛ	ø,œ	ɤ,ʌ	o,ɔ
[−high, +low]	a	ɶ	ɑ	ɒ

(For illustration we use here the Cardinal Vowel symbols (originally proposed by Daniel Jones; see e.g. Abercrombie 1967: ch. 10), rather than the vowels of any particular language.) Notice that the definitions of the features exclude the fourth logically possible combination of the two features defining the high–low axis, i.e. [+high, +low].

It is clear that other features will be required to characterise the vowel space, seeing that those discussed so far apparently fail to distinguish between various pairs of [−high, −low] vowels, for example. This is an area which has been the subject of major rethinking since the publication of *SPE*, and we devote some space here to a discussion of the issues involved.

1.3.3 The vowel-height dimension and related issues

There have been a number of proposals for distinguishing the various pairs of [−high, −low] vowels in (22). These proposals can be divided into three major groups: (i) those which distinguish the members of each pair by means of a binary feature [tense]; (ii) those which try to reflect the difference in height 'directly'; and (iii) those which introduce a feature [advanced tongue root] to distinguish the members of the various pairs.

The first proposal is found in *SPE*, in which it is argued that the difference is one of tense vs lax, with the tense vowel of each pair being 'executed with a greater deviation from the neutral or rest position' than its lax counterpart, so that 'the greater articulatory effort in the tense vowels is further manifested by their greater distinctiveness and the markedly longer duration during which the articulatory configuration remains stationary' (*SPE*: 324–5). Such a distinction is applicable to certain pairs of vowels which we have not yet considered, such as /iː/ vs /ɪ/ in RP English (e.g. *meal* /miːl/ vs *mill* /mɪl/), or in the German pairs given by Chomsky and Halle (e.g. *ihre* [īrə] 'her' vs *irre* [irə] 'err', or the similar *Huhne* [hūnə] 'chicken' vs *Hunne* [hunə] 'Hun'; notice that Chomsky and Halle distinguish tense vowels from their lax counterparts by means of a macron over the tense vowel).

This approach is also readily applicable to the [–high, –low] vowels in a system such as RP, in which the opposition between the members of each pair is again not just one of tongue-height (quality), but also of length (quantity). Thus the distinction between the two vowels in *beat* /iː/ and *bit* /ɪ/ is one which can be interpreted as tense vs lax, as is that between *mane* /eɪ/ and *men* /ɛ/, as well as in *coat* /əʊ/ vs *cot* /ɒ/.[14] The crucial claim that is being made here, then, is that the *type* of phonological opposition between, say, the two [–high, –low] vowels ([e] and [ɛ]) is different from that holding between, say, the high vowel [i] and the higher of the two mid vowels [e], or between the lower of the two mid vowels [ɛ] and the low vowel [a].

As we have seen, this account seems appropriate for a system like RP, but it has encountered criticism from those who believe that there are vowel systems in which it is reasonable to speak of (at least) *four* distinct vowel heights. The front-vowel system in (23), for example, is that of some dialects of Scots English:

(23)	*beat*	[bit]
	bit	[bɪt]
	bait	[bet]
	bet	[bɛt]
	bat	[bat]

[14] We are here following Chomsky and Halle's position that the phonological distinction between these pairs is one of quality, not length. Notice too that the diphthongs of *main* and *coat* are tense, just like the monophthong of *beat*.

Here the various vowels are apparently distinguished *only* by vowel height, with no apparent difference in length, and any appeal to the notion of tense vs lax, as defined by Chomsky and Halle, seems inappropriate.

The existence of systems like this has led some phonologists to propose systems which reflect the height dimension more directly. Thus Wang (1968) replaces the feature [low] by [mid], which allows the expression of four heights, rather than the three of *SPE*:

(24)
$$\begin{bmatrix}\text{+high}\\ \text{–mid}\end{bmatrix} \quad \begin{bmatrix}\text{+high}\\ \text{+mid}\end{bmatrix} \quad \begin{bmatrix}\text{–high}\\ \text{+mid}\end{bmatrix} \quad \begin{bmatrix}\text{–high}\\ \text{–mid}\end{bmatrix}$$

/i/	/e/	/ɛ/	/æ/

Such a formulation certainly allows the expression of four heights, but notice that there is still a fundamental problem associated with the expression of vowel height by means of binary features: the fact that we have to use (at least) two binary features to express what appears to be a *single* phonetic dimension. That is, a sequence such as /i/–/e/–/ɛ/–/æ/ can be seen as a set of points on a single scale, and this has led some phonologists, e.g. Ladefoged (1975) and Williamson (1977), to abandon binary features for the expression of vowel height, and to introduce a **multivalued scalar** feature, as in (25):

(25)	/i/	[4 high]
	/e/	[3 high]
	/ɛ/	[2 high]
	/æ/	[1 high]

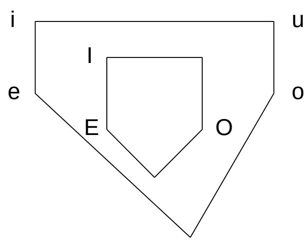
We do not pursue this approach at this point (but see our discussion of features in §2.1). Rather, we now consider the third type of proposal that has been made in this area, the introduction of a feature [advanced tongue root] (henceforth [ATR]).

It has been observed that one of the articulatory correlates of the difference between a pair of vowels such as [e] and [ɛ] typically involves the position of the tongue root: for [e] the tongue root is further forward, while for [ɛ] it is further retracted. A similar relationship holds between [i] and [ɪ]. Thus the difference between the two vowels does not relate exclusively to the relative height of the *body* of the tongue, as is suggested by (25), but involves additional phonetic parameters.

Clearly, the choice amongst the three alternatives just outlined – we refer to them here as the tense/lax, height and ATR approaches – depends on whether we can show that one of them more successfully predicts what actually happens in phonological systems and processes than the others. That is, if we find processes which show that the relationship between [i] and [e] is phonologically the same as that between [e] and [ɛ], this would provide evidence in favour of a multivalued feature [high], as in (24).

In fact, we believe that all three approaches are required in phonological theory. That is to say, we believe that vowel systems may be organised along any one of the three lines suggested by the approaches just discussed, so that the nature of the phonetic parameters which play a role in a particular sound-system is reflected in the phonology of the language in question.

Let us illustrate this with a further consideration of [tense] and [ATR]. We have already seen that the feature [tense] plays a role in the phonology of RP, for example. Thus /ɪ/ in *bit* and /ʊ/ in *look* are the [–tense] counterparts of /iː/ in *beat* and /uː/ in *Luke*, with which they are otherwise identical in terms of their feature make-up. The fact that the [–tense] vowels of RP form a class (which includes, besides /ɪ ʊ/, also /ɛ æ ʌ ɒ ə/) is shown by the fact that just this set of vowels cannot occur in final position in a stressed syllable, while the set of [+tense] vowels can (cf. /biː/ *bee* and */bɪ/, for example).[15] Similarly, they can occur before /ŋ/, while the [+tense] vowels cannot (e.g. *bang* /bæŋ/, but **boong* /buːŋ/).[16] Thus a feature such as [tense] is required in the analysis of systems such as RP. Crucially, at least with respect to the oppositions between /iː/ and /ɪ/ and between /uː/ and /ʊ/, the vowel system is organised in terms of 'central' vs 'peripheral' vowels. (26) gives the representation of a ten-vowel system such as this, consisting of a peripheral ([+tense]) set /i u e o ɑ/ and a central ([–tense]) set /ɪ ʊ ɛ ɔ ə/. Notice that for the low vowels peripherality is often manifested as greater pharyngeal constriction, so that peripheral /ɑ/ is considerably retracted:

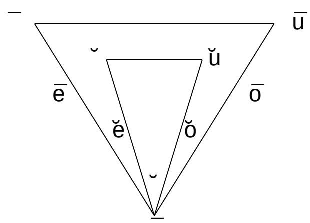
(26)


[15] It is also possible to deal with these restrictions in terms of vowel-length, rather than the apparently qualitative distinction of tense vs lax. We discuss vowel-length in Chapter 3.

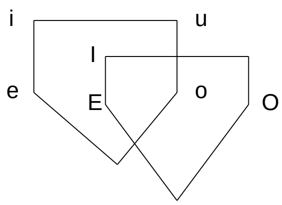
[16] The RP vowel /ʊ/ does not in fact occur before /ŋ/ in the native vocabulary of English. However, in loanwords such as *Jung* from German, we find /ʊ/, not /uː/ (cf. Collins and Mees 1996: 97), which further demonstrates the validity of the analysis.

The status of /ə/ is more problematical. It again fails to occur before /ŋ/, and shows other phonological behaviour which suggests that it forms a set of its own in some respects. However, given that it patterns with the lax vowels in not taking stress in final position in a syllable, we feel justified in categorising it in this set.

Such an analysis is proposed for Classical Latin by Allen (1973: 132), who observes: ‘the tenseness is . . . responsible for the long vowels occupying a larger, more “centrifugal” perimeter of articulations’. He represents the system of Latin as in (27), where the long tense vowels are represented with a macron (¯), the short lax vowels with a breve (˘):

(27) 

Other vowel systems, however, are organised quite differently, even though they may contain more or less the same vowels as (26). In particular, many languages divide the set of vowels not into a tense and a lax subset as in (26), but rather into two subsets according to tongue-root position, i.e. into one [+ATR] set and one [–ATR] set, as in (28):

(28) 

Here the vowels are grouped into /i u e o ə/ ([+ATR]) and /ɪ ʊ ɛ ɔ ɑ/ ([–ATR]).[17] The evidence that there are ten-vowel systems which are organised in this way comes from processes involving **vowel harmony**, i.e. processes in which all the vowels within a particular domain, often the word, must have the same value for a particular phonological feature. One such system is that of the Asante dialect of Akan, a language spoken in Ghana, in which all the vowels within a single word must have the same value for [ATR] (see Stewart 1967, 1983; Clements 1981). Stewart (1967: 186) gives the following examples:

[17] The relation between the symbols used for the non-low back vowels in (28) is to some extent arbitrary, although conventional. Thus /ʊ/ is used in (28) for a vowel which is apparently more peripheral on the front–back dimension than /u/, whereas in (26) this relationship is reversed. The opposite choice would be equally arbitrary, however, in that the same discrepancy would hold, but on the high–low dimension, rather than the front–back.

(29) [–ATR] [+ATR]

[–ATR]		[+ATR]	
/wʊbɛnʊmʔ/	'you will drink it'	/wubenumʔ/	'you will suck it'
/ɔːbɛtʊʔ/	'it is going to lay'	/oːbetu/	'he is going to pull it out'
/mɪkjɪrɛ/	'I show'	/mitie/	'I listen'

The forms in (29), involving the non-low vowels of Akan,[18] show the harmony process in operation: all the vowels in the left-hand column are [–ATR], while those in the right-hand column are [+ATR]. Forms with a mixture of [–ATR] and [+ATR] vowels, such as */wʊbenumʔ/, are ill formed.

We do not at this point pursue the question of how such harmony processes are to be analysed, an issue which we return to in some detail in §1.4.2 in relation to Turkish; however, the phenomena just outlined provide us with sufficient reason to claim that systems such as (26) and (28) are both to be found in languages of the world, and hence that the tense/lax and ATR features must *both* form part of our feature system.

We can also find evidence to support the point of view that the vowels of languages may be organised in terms of relative height, as suggested by the Scots English data in (23). Such evidence can be adduced from processes in which some or all of the vowels of a language move one 'step' up or down. For example, the effect of the English Great Vowel Shift was to move non-high long vowels up one step, giving the changes in (30) (from Lass 1987: 130), which shows various Middle English and early Modern English forms (*c*. 1600):

(30)

	ME		1600
beet	eː	>	iː
beat	ɛː	>	eː
mate	aː	>	ɛː
boot	oː	>	uː
boat	ɔː	>	oː

There are also processes which make appeal to the notion 'one step lower'. Lindau (1978: 545) observes that in Scanian, a Swedish dialect spoken in Malmö, there is a process in which the diphthongisation of a long vowel yields a first element which is one step lower than the original monophthong:

[18] As is typical in [ATR] harmony systems, the low vowels of Akan behave rather differently from the non-low ones, for reasons which need not concern us at this point.