INTRODUCING PHONOLOGY

Peter Hawkins

ROUTLEDGE LIBRARY EDITIONS:
PHONETICS AND PHONOLOGY

ROUTLEDGE

ROUTLEDGE LIBRARY EDITIONS:
PHONETICS AND PHONOLOGY

Volume 7

INTRODUCING PHONOLOGY

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

INTRODUCING PHONOLOGY

PETER HAWKINS

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published in 1984 by Hutchinson & Co. (Publishers) Ltd

This edition first published in 2019
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-138-60364-6 (Set)
ISBN: 978-0-429-43708-3 (Set) (ebk)
ISBN: 978-1-138-60405-6 (Volume 7) (hbk)
ISBN: 978-0-429-46871-1 (Volume 7) (ebk)

Introducing Phonology

Peter Hawkins

London

First published in 1984 by
Hutchinson & Co. (Publishers) Ltd

Reprinted in 1992 by
Routledge
11 New Fetter Lane, London EC4P 4EE

Set in Times Roman

Printed and bound in Great Britain by
Biddles Ltd, Guildford and King's Lynn

British Library Cataloguing in Publication Data

Hawkins, Peter
Introducing phonology.
1. English language—Phonology
I. Title
421'.5 PE1133

Library of Congress Cataloging in Publication Data

Hawkins, Peter
Introducing phonology.
Bibliography: p.
Includes index
1. Grammar, Comparative and general—Phonology.
2. English language—Phonology. I. Title.
P217.H34 1984 414 84–3826

ISBN 0–415–08397–4

Contents

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Preface

Phonology is the study of sound patterns in language. It stands alongside syntax, morphology, and semantics as one of the major branches of linguistics. While closely linked to phonetics, it emphasizes the *patterns* of sounds to be found in any particular language (and in languages generally), and the *relationships* between those sounds, rather than the *description* of the sounds in articulatory or auditory terms (which is the subject-matter of phonetics). Phonology is also a major component of several 'applied' fields such as variation between (and within) speakers of the same language ('dialects'), historical change, children's learning of their first language ('language acquisition'), and the teaching and learning of foreign languages.

Phonology is itself divided into two major components; *segmental* phonology, which is concerned with individual sounds (i.e. 'segments' of speech) and their patterns, and *suprasegmental* (or *non*-segmental, as it is sometimes called) phonology whose domain is the larger units of connected speech: words, phrases and sentences. The plan of this book is organized accordingly. Chapters 1 to 5 are concerned with segmental phonology; Chapters 6 and 7 discuss suprasegmental aspects, including stress, rhythm and intonation. The last three chapters are devoted to different fields of 'applied' phonology: Chapter 8 to dialects; Chapter 9 to historical change, i.e. the sound changes which take place within a language over a period of time; and Chapter 10 to language acquisition, with a focus on both 'normal' and 'delayed', or abnormal, development – the latter field being one which is now rapidly establishing itself, under the title of 'clinical phonology'. The last three chapters are included not just because they discuss practical and applied aspects of phonology, but also because these areas provide much of the raw material which informs (or should inform) theoretical phonology.

The book is intended as an introductory course for first or second year university students or the equivalent. It is not, however, intended

for those who have done no phonetics at all. Readers should be at least familiar with the basic terms of articulatory phonetics – terms such as *bilabial*, *fricative*, *lateral*, *voiceless*, *aspirated*, etc. – and with the *symbols* which denote these classes of sounds (for example, [k g ŋ x] as instances of the category *velar*, etc.). As there are already a number of good introductory textbooks available I decided that it would not be advisable to include chapters on articulatory phonetics, even in an introductory text. I have, however, included a *glossary* of basic terms for reference purposes, and some exercises for practising the notation used in this book for the transcription of English.

Each chapter includes a number of exercises, which readers are recommended to work through. In some cases the exercises have been incorporated within the chapter as an essential part of it: the exercise is followed by a 'discussion' section which uses the data as its raw material. The aim is to encourage readers to *do* the exercises – if they are grouped at the end of a chapter it is all too easy to ignore them!

Every author should declare his theoretical bias – particularly in linguistics which has seen, in recent years, a number of significant theoretical changes and the proliferation of various 'schools'. It is undeniable that the dominant model has been, until recently, the 'orthodox generative' phonology of the 1960s. While sympathetic towards this model, I have tried to avoid both its more dogmatic aspects and the excessive formalism it has given rise to. There is thus virtually nothing on such topics as rule ordering and the interrelationships between different phonological rules, topics which provoked endless, but essentially unproductive, theoretical argument during the 1970s. I have also avoided the excessively abstract analyses favoured by orthodox generative phonology, preferring instead a 'near-the-surface' approach which respects 'phonetic facts' (for discussion, see Chapter 5). I believe it is important to be guided not only by the phonetic data, but also by the views of ordinary native speakers of the language, as reflected where possible in their (linguistic) behaviour.

The first two chapters are (I hope) theoretically uncontroversial. Chapter 3 traces the development of distinctive features and Chapter 4 the concepts of neutralization and language universals, all of which are an accepted part of current phonology. In Chapter 5 I have faced up to the theoretical issues, looking first at the problems of 'classical' phonemics, then at the generative 'solution', and finally offering a critique of the generative view.

Theoretical phonology can (and, in my view, should) be concerned

with the description and validation of the fundamental phonological processes of language – processes which are found over and over again in different languages and at different periods of time, and which often apply in several different fields, such as dialectal variation, language acquisition, etc. These phonological processes will be the main concern of the last two chapters, but they are introduced in several earlier chapters (including Chapters 2, 4 and 6), and they form what is perhaps the work's main theoretical 'theme'.

Examples and data in the book are taken from a variety of languages, but mainly from English for two reasons: first, because more linguistic/phonological analysis is available for English than for any other single language; and second, because the readership is aimed at native and non-native English speakers. In gathering data from other languages I have confined myself to languages of which I have first-hand experience or a very reliable source. Illustrations are thus mainly from European languages, which again are those likely to be most familiar to the readers themselves, who can thus verify the accuracy (or otherwise!) of the data.

I am indebted, in particular, to my students. Not only have they tried out most of the exercises, but their transcriptions (particularly the errors) have provided a valuable resource for observing 'native speaker intuitions', i.e. for establishing just how much a native speaker 'knows' (or does not know) of his language. I owe also a debt of gratitude to those who taught me phonetics and phonology at the University of Manchester and at University College London: in particular to Edward Carney, Professors Fry and Gimson, Messrs O'Connor and Arnold, and Dr John Wells.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Transcription symbols for English

Given below is the notation used in this book to transcribe the phonemes of English.

The transcription is based on the standard accent of English known as Received Pronunciation (RP). As a notation it will therefore be more appropriate for English (as opposed to Scottish or Irish) accents, and for RP-related accents in other parts of the English-speaking world. For Scottish, Irish and American accents, some modifications will be necessary; for example most of them would not require /iə, eə, uə/, and for most American accents the vowel of *pot* would be /ɑ/ rather than /ɒ/. Differences between the various accents of English are discussed in Chapter 8.

Apart from exceptions just mentioned (/iə, eə, uə, ɒ/), the phonemes in the list below are the 'basic' phonemes which will be recognized by virtually all speakers of English – they are in a sense, the 'universal' phonemes of English.

The notation we use corresponds very closely to that of Ladefoged (1975). It differs in minor details from other popular notations such as:

1 the one used by Gimson in his *Introduction to the Pronunciation of English* (3rd edn 1980)* and in the *English Pronouncing Dictionary* from the 14th Edition (1977) onwards;
2 the notation used by J. Windsor Lewis in his *Concise Pronouncing Dictionary* (1972) and in the *Oxford Advanced Learner's Dictionary* from the 3rd Edition (1974) onwards;
3 the notation of Longman's *Dictionary of Contemporary English.*

*Full references quoted in the text are contained in the Bibliography beginning on p. 309.

Vowels

1 Simple

/ɪ/	pit	/ʌ/	putt
/e/	pet	/ɒ/	pot
/æ/	pat	/ʊ/	put
		/ə/	*p*otato

2 Long vowels and diphthongs

/i/	beat	/iə/	beer
/eɪ/	bait	/eə/	bear
/aɪ/	bite	/uə/	tour
/ɔɪ/	boy		
/u/	boot	/ɜ/	Bert
/oʊ/	boat	/ɔ/	bought
/aʊ/	bout	/ɑ/	bart

Consonants

Plosives	/p/	pin	Fricatives	/f/	fin
	/b/	bin		/v/	vine
	/t/	tin		/θ/	thin
	/d/	din		/ð/	then
	/k/	kin		/s/	sin
	/g/	gain		/z/	zinc
				/ʃ/	shin
Affricates	/ʧ/	chin		/ʒ/	measure /'meʒə/
	/ʤ/	gin			
Nasals	/m/	ram	Approximants	/l/	lime
	/n/	ran		/r/	rhyme
	/ŋ/	rang		/j/	yet
				/w/	wet
				/h/	hen

Reading and transcription practice

Exercise 1

Read the following texts, given in the notation used in this book.

/wʌn mɔnɪŋ ə bæŋk mænədʒə ɒn ɪz weɪ tə wɜk ɪn nju jɔks wɔl strit wəz stændɪŋ niə ðɪ oʊpən dɔz əv ðə sʌbweɪ kærɪdʒ. sʌdn̩li ə kɜli heəd jʌŋ mæn dʒʌmpt ɪntə ðə kærɪdʒ, brʌʃt əgenst ɪm, ənd lept aʊt əgen. ðə bæŋkər ɪnstɪŋktɪvli felt fər ɪz wɒlət ænd, faɪndɪŋ ɪt mɪsɪŋ, hi rɪʧt aʊt ən græbd ðə jʌŋ mæn baɪ ɪz koʊt kɒlə. ðə sʌbweɪ dɔz kloʊzd wɪð ðeə rʌbər edʒəz raʊnd ðə mænədʒəz ɑmz bʌt i held ɒn ivən wen ðə kærɪdʒ stɑtəd muvɪŋ, ən i drægd ðɪ ʌðə mæn sevrəl fit əlɒŋ ðə steɪʃn̩ plætfɔm. faɪnəli ðə kɒlə tər ɒf, livɪŋ hɪm wɪð ə fju ɪntʃəz əv tætəd klɒθ ɪn ɪz hænd. stɪl æŋgri, hi mɑʧt ɪntu ɪz ɒfɪs. ðə foʊn ræŋ. ɪt wəz ɪz waɪf. baɪ ðə weɪ, diə, ʃi sed, dɪd jə noʊ jə left jɔ wɒlət ət hoʊm./

Exercise 2

Transcribe the following into the notation given. Begin and end with a slant line to denote 'phonemic transcription'.

I take it you already know
Of though and bough and cough and dough?
Others may stumble, but not you
On hiccough, thorough, rough and through
Well done! And now you wish, perhaps,
To learn of less familiar traps?
Beware of heard, a dreadful word
That looks like beard and sounds like bird,
And dead: it's said like bed, not bead.
For goodness sake don't call it 'deed'!
Watch out for meat and great and threat,
(They rhyme with suite and straight and debt);
A moth is not as 'moth' in mother,
Nor both in bother, broth in brother.
And here is not a match for there,
Nor dear and fear for bear and pear,
And then there's dose and rose and lose –
Just look these up – and goose and choose.

And cork and work, and card and ward,
And font and front and word and sword,
And do and go, and thwart and cart,
Come, come I've hardly made a start!
A dreadful language? Man alive,
I'd mastered it when I was five!

1 Phonemic and phonetic

Phoneme, allophone

When we talk about the sounds of a language, the term 'sound' can be interpreted in two rather different ways. In the first place, we can say that [p] and [b] are two different sounds in English, and we can illustrate this by showing how they contrast with each other to make a difference of meaning in a large number of pairs, such as *pit* vs *bit*, *rip* vs *rib*, etc. But on the other hand, if we listen carefully to the [k] of *key* and compare it with the [k] of *car*, or if we compare the vowel of *chew* with the vowel of *coo*, we can hear that the two sounds are not the same; the [k] of *key* is more 'fronted' ('palatalized') in its articulation and has a higher pitch, while the [k] of *car* is articulated further back and has a lower pitch. Similarly, the vowel [u] is fronted when it follows the palato-alveolar [ʧ], but retracted after the velar [k̠]. In each case we have two different 'sounds', one fronted and the other retracted. Yet it is clear that this sense of 'sound' differs from the first sense, for we could also say that fronted and retracted [k] are both variants of the one 'sound' k (and likewise, both forms of [u] are variants of the 'same' vowel). To avoid this ambiguity, the linguist uses two separate terms: *phoneme* is used to mean 'sound' in the former (i.e. the contrastive) sense, and *allophone* is used for sounds which are variants of a phoneme: sounds which differ, but which do not contrast. We would thus say that fronted [k] and retracted [k] are *allophones* (i.e. variants) of the *phoneme* k, and to make the distinction quite clear in writing, we enclose allophones in square brackets, [ ], and phonemes in slants, / /. Using this notation we can now write [k̟] (fronted k) and [k̠] (retracted k) as allophones of /k/.

Further examples of phonemes/allophones are readily available.

1 The [n] of *tenth* differs from the [n] of *ten*; in *tenth* the sound is dental, [n̪], while in *ten* it is the 'ordinary' English alveolar [n] (compare the two sounds by looking into a mirror as you pronounce

them). [n̪] and [n] are allophones of the phoneme /n/; the dental allophone is found whenever another dental ([θ, ð]) immediately follows.

2 The [l] of *lip* differs, in many accents, from the [l] of *pill*. In *pill*, the [l] is accompanied by a raising of the back of the tongue: it is *velarized*. This sound, known as 'dark l', and written [ɫ], differs from the 'clear l' of *lip*, usually written simply as [l]. These two sounds are allophones of the phoneme /l/. There is a corresponding difference of distribution: [l] occurs *before* a vowel, [ɫ] *after* one.

3 The [s] of *seep* and the [s] of *soup* are not identical. The phoneme /s/ has a plain, unrounded allophone, [s], in *seep*, and a rounded, or *labialized*, allophone [s͍] in *soup*. As we might expect, the rounded allophone is used when a rounded vowel follows, and the unrounded allophone elsewhere.

Contrastive function

The native speaker[1]* is quite readily aware of the phonemes of his language but much less aware of the allophones: it is possible, in fact, that he will not hear the difference between two allophones like [k̟] and [k̠] even when a distinction is pointed out; a certain amount of ear-training may be needed. The reason is that the phonemes have an important function in the language: they differentiate words like *pit* and *bit* from each other, and to be able to hear and produce phonemic differences is part of what it means to be a competent speaker of the language. Allophones, on the other hand, have no such function: they usually occur in different positions in the word (i.e. in different *environments*) and hence cannot contrast with each other, nor be used to make meaningful distinctions.

For example, as noted above [ɫ] ('dark l') occurs *following* a vowel as in *pill*, *cold*, *school*, but is not found before a vowel, whereas [l] ('clear l') only occurs *before* a vowel, as in *lip*, *late*, *like*. These two sounds therefore cannot contrast with each other in the way that /l/ contrasts with /r/ in *lip* vs *rip* or *lake* vs *rake*; there are no pairs of words which differ only in that one has [l], the other [ɫ].

Predictability

The difference between phoneme and allophone can be seen in terms

* Superior figures refer to the Notes beginning on p. 301.

of predictability. Allophones are predictable in that we can say: in environment P we find allophone A; in environment Q, allophone B; in environment R, allophone C, etc. Thus, given the sequence -Vl# or -VlC, where V stands for (any) vowel, C for (any) consonant, and # for word-boundary, we can predict that the allophone of /l/ occurring here will be [ɬ] and not [l].

Phonemic contrasts are not predictable in this way. Given the environment /–ip/, we have a choice between /p/ (*peep*) vs /l, w, s/, etc. From the phonetic environment there is no way of predicting which phoneme will occur. We shall see below that the predictability of the allophones can be expressed more formally by means of rules.

Most of the examples given so far have involved consonants. A well-known case of allophonic variation in vowels is the length difference, which is determined by the type of consonant following. Vowels are shorter in duration before voiceless consonants such as p, t, k, s, f, and longer before voiced sounds, such as b, d, g, z, v. The vowel of *lock* is shorter than that of *log*, and the vowel of *rice* shorter than that of *rise*. Using the length mark [ː], we can show two allophones for each vowel: [ɒ, aɪ], short, and [ɑː, aɪː], long (for details of relative durations, cf. Umeda and Coker (1975: 552)). This example illustrates the points we made above; first, that the native speaker is not normally aware of allophonic differences (it requires careful listening to detect the difference between the shorter vowel of *lock* and the longer vowel of *log*), and second, that the allophones are predictable in terms of their phonetic environment: short before voiceless sounds, long before voiced ones.

Perception of allophones

We are, as noted, not normally aware of allophonic variation in our own language; in listening to or learning another language, however, we may be able to perceive allophones which native speakers of that language are themselves unaware of, particularly if the allophones correspond to a *phonemic* distinction in our own language. An example is the relationship between voiced plosives ([b, d, g]) and voiced fricatives ([β, ð, ɣ]) in English and Spanish respectively. In Spanish there are three pairs of allophones, [b, β], [d, ð] and [g, ɣ], distributed as illustrated:

[b]	[beso]	'kiss'	[β]	[aβa]	'bean'
	[bala]	'ball'		[deβer]	'owe'
	[bomba]	'bomb'			

[d]	[dama]	'lady'	[ð]	[boða]	'wedding'
	[duro]	'hard'		[toðo]	'all'
	[donde]	'where'		[θjuðað]	'city'
[g]	[gato]	'cat'	[ɣ]	[aɣosto]	'August'
	[gaŋga]	'bargain'		[laɣo]	'lake'
	[grande]	'big'			

For a Spaniard, [b] and [β] are 'the same sound', as are [d] and [ð], [g] and [ɣ]; each pair are allophones of a single phoneme, the distribution being determined by the phonetic environment, namely:

> *plosives* occur word-initially and after nasals (and laterals)
> *fricatives* occur between vowels ('intervocalic') and (for [ð] only) finally

The English speaker has little difficulty in distinguishing [d] and [ð] since the difference is *phonemic* in English: compare *breathe* and *breed*, *riding* and *writhing*. Similarly, [b, β] are readily distinguished, since although [β] is not an English sound, it is sufficiently similar to [v] to be heard as such; and /b, v/ is a phonemic distinction in English, like /d ð/. But [g, ɣ] are different; English has no /ɣ/, nor anything acoustically similar to it. In fact, [ɣ] is sometimes an allophone of /g/ in English: it is often found as a 'reduced' or casual pronunciation of /g/ between vowels, as in *again*, *Agatha*, *eagle*. Since [g, ɣ] are allophones in English as well as in Spanish, the perceptions of speakers in both languages tend to be similar: they are not normally aware of a difference.

In German, the 'ch' sound has two allophones; a velar fricative [x] after back vowels, and a palatal fricative [ç] after front vowels. English speakers are more aware of this difference than Germans, because they associate [x] with k-like sounds (i.e. with other velars), and [ç] with the acoustically similar [ʃ], which is a different phoneme. Germans perceive [x, ç] as 'the same' sound, since they constitute a single phoneme.

Often, therefore, we become aware of allophonic variation in our own language as a result of observations from speakers of other languages. Sapir, for example, found that his Haida (an American-Indian language) informants could readily distinguish the [t^h] of English *top* from the unaspirated [t] of *stop*, a difference noted only with difficulty by English speakers themselves. The reason is that, in Haida, [st–] contrasts phonemically with [st^h–], i.e. the sound-

difference is used to make differences of meaning. On the other hand these informants could hardly distinguish [t] from [d] (as in *steer* and *dear*), because Haida has no voiced–voiceless contrast (Sapir 1970 (1921): 55).

Similarly, Jones (1950: 37) reports the case of the speaker of Syriac, who noticed a difference between the [t^h] of English *ten* and the [t] of *letter*, a difference which is phonemic in Syriac but only allophonic in English. In Kikuyu, [d ð θ] are all allophones of one phoneme. When a writing system (an *orthography*) was provided for that language, it was devised by English speakers, who, hearing a difference between [d] and [θ, ð], allocated two separate symbols, 'd' for [d] and 'th' for [θ, ð]. They did not realize that, as far as the Kikuyu is concerned, all three are for him the 'same' sound, for which a single symbol would be appropriate.

The phonetician is trained to listen carefully for all possible differences of sound regardless of the particular language they occur in. But even phoneticians are native speakers of some language, and may be more familiar with, say, European languages in general than with any others. This tends to colour their interpretation of the sounds they hear in unfamiliar languages, and acts as a bias, however unconscious. As Bloomfield says, the phonetician's equipment is 'personal and accidental: he hears those features which are discriminated in the languages he has observed' (1933: 84). For this reason, the phonetician does not rely too heavily on his own phonetic transcription of what he hears: he consults native speakers of the language as early as possible, to ascertain the distinctions they are readily aware of, and to make use of their intuitions and insights for interpreting the data.

Methods of phonemic analysis

It is obviously important to have some means of determining, for an unfamiliar language, which differences of sound are phonemic and which are allophonic. A number of principles have been established for ascertaining the phonemic structure of a language, and to illustrate how these work we can take some hypothetical examples involving the sounds [s] and [z]. Let us assume that a field-worker has collected a list of words, with their meanings, in a hitherto unfamiliar language, X. This language contains the following words:

[s]	[z]
[sak]	[zapi]

[setu]	[zen]
[suni]	[zudok]

It has been established that each word has a different meaning, and the sounds have been transcribed according to the IPA conventions: thus [s] represents a voiceless (and [z] a voiced) alveolar fricative. This is a *phonetic* transcription: we do not yet know whether the [s] and [z] are contrastive in this language, or whether they are merely variants of one phoneme. Hence square brackets are used, and the notation gives as much detail as possible. Since we do not know yet whether [s] and [z] are allophones or phonemes, we must call them simply *phones*. A phone is hence 'a sound whose status in the language is hitherto undetermined'.[2]

From this small set of words we cannot definitely decide whether [s] and [z] are separate phonemes, but we can hazard a guess: both sounds appear before the vowel [e] (setu, zen), before [a] (sak, zapi), and before [u] (suni, zudok); both sounds occur at the beginning of the word (*word-initially*, to use the technical term). Since they occur in similar environments like this, it is most probable that they are contrastive, and hence separate phonemes. If they were allophones, we would have to find differences in their environments, which in this case would be difficult. We might propose that [z] occurs only in one-syllable words (as in *zen*) and [s] only in polysyllabic words (for example, *setu*) but this is contradicted immediately by the examples *zapi* and *sak*. On the existing evidence, therefore, [s] and [z] are separate phonemes.

To strengthen this conclusion, we might be lucky enough to find what is called a 'minimal pair': two words with different meanings which differ *only* in that one word has [s], the other [z], such as [sed] vs [zed]. This would be strong evidence that [s] and [z] are contrastive in this language and hence can be regarded as two phonemes, /s/ and /z/. Minimal pairs are useful for establishing quickly and simply the phonemes of a language; for English, the frame /–ɪp/ immediately gives us /p, t, k, d, s, z, ʃ, l, n, r, h, w/ as separate phonemes. But not all languages afford examples of minimal pairs as readily as English.

Let us now suppose that in another unfamiliar language, Y, the following set of words occurs:

[s]	[z]
[sato]	[bazo]
[sip]	[okezi]

[nis]	[lozak]
[pos]	[retizen]

From this small amount of data, it appears that the two sounds occur in different environments: [s] is found word-initially (*sato*, *sip*) and word-finally (*nis*, *pos*), while [z] occurs between vowels. This distribution of [s] and [z] makes good phonetic sense, for we find the voiced sound between other voiced sounds (vowels), while voiceless [s] is either preceded or followed by silence, i.e. by a state of voicelessness.

The data suggests, then that [s] and [z] here are not separate phonemes: they are allophones of one phoneme, and we could expect that further data would show, analogously, that [s] is found only in a 'voiceless' environment (for example, before a voiceless consonant) and [z] only in a 'voiced' one. The *distributions* of [s] and [z] are different; and they complement each other, for [z] is not found in [s]-type environments, nor [s] in [z]-type. When this happens, the two sounds are said to be in *complementary distribution*. In English [l] and [ɬ] are in complementary distribution, because their environments are always different.

Since [s] and [z] are allophones of one phoneme in this language, we shall most probably find that native speakers of it are not normally aware of a difference between these two sounds, and that the difference would only be appreciated with difficulty, if at all. This would come as a surprise to native speakers of English, who find it hard to imagine that anyone could fail to hear the difference between [s] and [z]. It is a good illustration of the fact that, where languages are concerned, everything is relative. English speakers hear the difference because it has an important function in the language, distinguishing many pairs of words. Speakers of language Y do not hear the difference because it is of no importance in the communication process; the allophones are entirely predictable according to the phonetic environment, and thus carry no useful information.

Symbols

If [s] and [z] are allophones of one phoneme, what symbol are we to use for the phoneme? Is it to be represented as /s/, or as /z/? Clearly, either symbol could be used: the choice is arbitrary. But whichever we choose, we must be consistent; we shall transcribe *all* the words above with the same symbol. If we choose /s/, then [bazo] and [okezi] will be

phonemically transcribed as /baso/ and /okesi/. This seems wrong to us as speakers of English, but it will be entirely in accord with the feelings of a native speaker of Y, for whom [s] and [z] are 'the same sound'. If *sato* is transcribed with /s/, he will want to use /s/ for *bazo* and *okezi* as well, just as in English we are happy to use the same symbol, /l/ for both the [l] of *leap* and the [ɬ] of *peel*; cf. also the Kikuyu example, mentioned above.

Native speaker knowledge

Let us digress for a moment to consider an important theoretical point.

The fact that this analysis of [s] and [z] accounts for the intuitions of the native speaker is the real reason why we adopt it, for one of the aims of a grammar is to explain and take account of native speaker feelings about his language, as far as this is possible, and this applies to the phonological part of the grammar just as much as to the area of syntax. That [s] and [z] are in complementary distribution is a fact about the language, but it would be of little interest in itself and this is not the real reason why we treat these two sounds as one phoneme; the real reason is that the native speaker *feels* they are one phoneme. This shift of emphasis parallels, to some extent, changes which have taken place in linguistic theory, for the 'structuralist' model of languages which flourished from the 1930s to the 1950s emphasized the 'facts about the language' approach and regarded the phoneme as something to be justified in terms of complementary distribution and the other principles we are presenting in this chapter. Thus the fact that [s] and [z] were similar phonetically and were found to be in complementary distribution would have been sufficient justification, for the structuralists, to treat them as one phoneme.

The 'generative' model of language, which appeared during the 1960s, put greater emphasis on the need for a linguistic analysis to have explanatory power, that is, to explain adequately what the native speaker intuitively 'knows' about his language. Hence the phonemes of a language are 'discovered' in consultation with native speakers and not by the rigid application of the methods outlined here.[3] These methods are still useful, however, because they point to the reasons why the native speaker feels as he does; thus, [s] and [z] are felt to be one phoneme because their distribution is complementary, and because they are never used to make meaningful contrasts in the language.

Realization rules

To return to the main argument: we showed above that the choice of symbol for the s/z phoneme was partly arbitrary; we could write it either as /s/ or as /z/. Alternatively, we could invent a new symbol, such as /≤/, or we could take a completely abstract symbol like the number /10/. All we would need is some specification to state that phoneme number 10 is pronounced [s] in some environments and [z] in others, but – and this is our next major observation – such a specification is needed anyway, *whatever* symbol we choose. For even if we represent the phoneme as /s/, we still need to know that it is pronounced [s] in some environments and [z] in others, if we are to have a full account of the language.

These specifications, then, will be our guide as to how exactly each phoneme in the language is actually pronounced. We shall call them *realization rules*,[4] because they describe how each phoneme is actually realized as physical sound. For the s/z phoneme, which we shall arbitrarily write as /s/ rather than as /≤/, /10/, etc., we shall need a realization rule something like the following:

/s/ → [s] word-initially, word-finally
→ [z] between vowels

The arrow here means 'is realized as', 'is pronounced as'. In recent years various notations have been developed for phonological rules of this kind; using some of the conventions that have been widely adopted, we could express the above rules as:

/s/ → [s] /# __, __ #
[z] /V __ V

The slant line here separates the operation of the rule (to the left of it) from the environment (to the right), and # represents a word boundary. The dash (__) represents the position of the sound itself in relation to its environment, so # __ means 'word-initially', V __ V means 'between two vowels', etc. Hence this rule can be read as: /s/ is pronounced [s] in two environments: first, immediately *after* a word-boundary (# __); and second, immediately *before* one (__ #). It is pronounced [z] when it occurs *between vowels* (V __ V).

Symbol values

Notice how the symbol *s* has two quite different 'values'; in square brackets it gives a fairly precise description of a physical sound: a voiceless alveolar fricative. In slant brackets it is much less precise: it stands for more than one sound, and we therefore cannot know how this /s/ is actually *pronounced* until we have looked at the realization rule. This means that if one is given a *phonemic* transcription of words or sentences in an unfamiliar language, they cannot be read aloud because there is no way of knowing what value each phoneme has in each environment. /s/ could be pronounced [s], but it might equally be [z], or [ʃ], or [ʂ], etc. Reading aloud can only be done from a *phonetic* transcription, or alternatively from a phonemic one with the aid of the realization rules, which amounts to the same thing. This is a point worth making, because often a language may be written with an alphabetic notation in which each symbol corresponds to one phoneme; one is tempted to believe that, even though the language is unfamiliar, the words can be pronounced simply by reading what is written. Only a native speaker of the language can do that, because he knows the realization rules intuitively. A foreigner requires the realization rules as well as the phonemic transcription.

A realization rule for English vowel length might be written as follows:

$$/\mathrm{V}/ \rightarrow [\mathrm{V}] \quad /\_\_ \quad [-\text{ voice}]$$

$$[\mathrm{V}\text{:}] \quad /\_\_ \quad \left\{ \begin{matrix} [+\text{ voice}] \\ \# \end{matrix} \right\}$$

The first part of the rule means: a vowel (any vowel) is pronounced (relatively) short when followed by a *voiceless* ([− voice]) sound; the second part, that vowels are relatively long when preceding *either* a voiced sound *or* a word-boundary. The curly brackets denote the 'either–or' relationship.

Rules like this are only expressing, by a convenient notation, something that can equally be stated less formally. The advantage of the formal statement is a gain in economy; there may also be gains in precision, because sometimes facts are brought to light which might not have been noticed from the less formal statement alone.

Free variation

The data we have used so far has illustrated two possibilities: that [s] and [z] are separate phonemes (when their distribution is *contrastive*), and that [s] and [z] are allophones of one phoneme (when their distribution is *complementary*). There is, however, a third possibility; namely, that the sounds [s] and [z] both occur in the language but speakers are inconsistent in the way they use them. Let us suppose that the language has the words

[silap] [snud] [dislik] [ras]

but we find variants of these words: [zilap] has the same meaning as [silap]; [znud] occurs, but is regarded as 'the same word' as [snud], etc. Clearly [s] and [z] cannot be contrastive phonemes, but nor are they allophones in the usual sense, for their distribution is parallel. We must then take them as *free variants* of a single phoneme, but since the situation seems somewhat unusual we would take some trouble to find the *reason* for the variation: it is most unlikely that variation of this kind could be completely 'free', i.e. completely arbitrary. If the variation occurs *between* speakers, then we shall try to explain it on the basis of *dialect*: dialects can be regional (for example, speakers from town X might use [s], those from city Y might use [z]) or social (when usage will be defined according to social class, level of education, etc.). If the variation is found *within* individual speakers, then the language situation must be studied more closely: it could be that [z] is a 'prestige' form which the speaker uses when he is consciously 'monitoring' what he says, while [s] is the pronunciation found in casual or less formal speech. If all other explanations fail then we have truly 'free' variation, but this label is, in a sense, an admission that the critical factors at work have not yet been explained.

It is interesting to compare speakers' reactions to each of the three possibilities discussed so far.

1 If [s] and [z] are separate phonemes, the native speaker has no difficulty in perceiving the difference.
2 If they are allophones of one phoneme, it is most probable that he/she fails to perceive the difference until asked to listen carefully.
3 If they are 'free variants', two possibilities arise:
   a The difference is clearly perceived ('that speaker said [zilap]

and [znud]: you can tell he comes from the North. We say [silap] and [snud] down here').

b The speaker is aware of a difference, but cannot pinpoint it ('I can tell he comes from another part of the country; it's something about his speech').

In the latter case, the speaker may become aware of the difference once it is pointed out.

Phonetic similarity

We have indicated so far that the analysis of the sounds of a language is based on such notions as *parallel* (contrastive) *distribution*, *complementary distribution*, *minimal pairs* and *free variation*. To these we must add two more concepts, those of *phonetic similarity* and *phonemic pattern*.

To illustrate phonetic similarity we can use the data from the second imaginary language, Y, where we found [s] and [z] to be allophones of one phoneme. We came to this decision because their distribution seemed to be complementary: [s] in a voiceless environment, [z] in a voiced one. But before we can justifiably regard them as one phoneme, we must be sure that they have some similarity – that they share some features in common. In this particular case the answer is obvious: we know that [s] and [z] are similar, and to support our view we can point to the shared features: both are alveolar, both are fricative – i.e. they share the critical features of both *place* and *manner* of articulation. But there are cases where the similarity is not so great, and there is one example from English which shows how the procedures outlined so far could easily be abused: this is the case of [h] and [ŋ].

[h] in English is defective in its distribution: it is only found initially in a syllable, and apart from a few words like *ahead*, *behave*, it is normally initial in the *word* as well. [ŋ] is also defective in distribution: it occurs only finally (*sing* /sɪŋ/) or before the other velar sounds /k/ and /g/ (*sink* /sɪŋk/, *finger* /fɪŋgə(r)/). It can occur between vowels only if there is a morpheme[5] boundary (for example, *sing-er* /sɪŋə(r)/). One could argue that [ŋ] and [h] both occur between vowels, but it seems that [h] when intervocalic always follows an *unstressed* vowel (*ahead*, *behave*, etc.) while [ŋ] intervocalically always follows a *stressed* one (*ringer*, *singing*, etc.). Hence, not only are there no minimal pairs to contrast [h] and [ŋ], but there isn't even the *possibility* of such a pair. The two sounds appear to be truly in

complementary distribution and could, according to our procedures, be regarded as one phoneme. To follow this argument to its rather absurd conclusion: we could invent the symbol /ɧ/ to represent this phoneme, and transcribe *head* as /ɧed/, *bang* as /bæɧ/, etc. *Hung*, which has both sounds, will be transcribed as /ɧʌɧ/. The realization rules would present no difficulty: we could have

/ɧ/ → [h] / # __ , V̆ __ V (where V represents an unstressed vowel)

[ŋ] / __ #, __ $\begin{Bmatrix} k \\ g \end{Bmatrix}$, 'V __ V (where 'V represents a stressed vowel)

While many linguists have discussed the possibility, it would be unlikely for anyone, linguist or non-linguist alike, to accept an analysis which identifies [h] with [ŋ]. We could justify our rejection on the grounds that the two sounds differ both in *place* of articulation ([h] is usually regarded as 'glottal'; [ŋ] is velar) and in *manner* ([h] is fricative; [ŋ] is a nasal), as well as in *voicing* ([h] is voiceless; [ŋ] is voiced). The sounds have nothing in common except that they are both consonants, and this is too general a feature to be sufficient. So we reject the association of [h] and [ŋ] on the grounds of (lack of) phonetic similarity.

Criteria for 'similarity'

This answer to the problem of [h] and [ŋ] is not entirely satisfactory, for there is no obvious criterion for deciding whether two sounds are sufficiently 'similar' or not; how similar must they be, in order to qualify as a single phoneme? Alternatively, how different must they be before we say, 'sounds x and y cannot possibly be variants of one phoneme'? There is no straight answer to these questions, and we can find examples from various languages to show, first, that phones which on the face of it seem quite dissimilar are in fact members of one phoneme, and are felt by native speakers to be (variants of) 'a single sound'; and second, that phones which seem remarkably *similar* can nevertheless be regarded by the native speaker as separate phonemes. To take the latter possibility first, we can look at English plosives in final position. Of three possible pronunciations, [kæp^{h}], [kæp^{o}] and [kæb^{o}], the bilabial plosive of the first is voiceless, released, and aspirated; that of the second is voiceless and unreleased, while

that of the third is voiced and unreleased (the raised circle denotes an unreleased consonant). The voicing of the latter, however, may manifest itself only as a lengthening of the preceding vowel, the plosive itself being voiceless and hence identical to [p°]. Hence, the difference between [p^h] and [p°] is greater, phonetically, than that between [p°] and [b°], yet the native speaker associates the former pair as variants of /p/ (/kæp/, *cap*), but distinguishes the latter pair as two phonemes /p/ vs /b/ (*cap* vs *cab*). Sometimes, therefore, a fine phonetic difference can have critical significance and serve to distinguish two phonemes, while a bigger phonetic difference (as between [p^h] and [p°]) can be ignored as irrelevant.

The other possibility, which is more central to our argument, can be illustrated by a well-known example from Japanese, where a single phoneme has as its variants (allophones) a bilabial fricative [ɸ], a palatal fricative [ç], and a glottal fricative [h]. The distribution of these three phones is complementary: [ɸ] occurs before /u/, as in [ɸuku] 'luck', [ç] before /i/, as in [çito] 'man', and [h] before /e, a, o/, for example [hana] 'nose' (cf. Jones 1950: 20). From an articulatory point of view these allophones range across the whole spectrum of *place* of articulation, from bilabial at the front to glottal at the back. How can such dissimilar phones be members of a single phoneme?

To understand what is happening here we must briefly examine speech as a communication process, and we shall see that phonetic similarity, like the discipline of phonetics itself, can be viewed from two perspectives, one concerning the speaker, the other the listener. Phonetics, the discipline, is divided into three branches, each corresponding to a different stage in the communication process. Articulatory phonetics studies the movements made by the speaker, and the resulting sounds are described in articulatory terminology such as *labio-dental*, *fricative*, *voiceless*, etc. Acoustic phonetics studies, by instrumental methods, the disturbances of air particles resulting from the speaker's articulations: it studies the transmission stage between speaker and listener. Auditory phonetics studies what the listener perceives: the impressions made by the sounds on the listener's receptive mechanisms, and his interpretation of what he hears. Auditory phonetics is a study of perception, and has connections with psychology in the same way as studies of visual perception have.[6]

Until recently, articulatory phonetics has been the dominating branch, and most descriptive work has been done in articulatory terms. Furthermore, there has appeared to be no need to alter the

balance (in favour of auditory description) in any substantial way, since in most cases there is a one-to-one correspondence between what happens at the speaker's end (articulation) and what the listener is aware of (perception): a difference in articulation (such as plosive vs fricative) has a corresponding acoustic effect and is hence perceived as different by the listener. Conversely, whatever the hearer perceives as different, requires a difference of articulation on the part of the speaker: one cannot use the same articulation to make two different sounds.

There are occasions, however, when this one-to-one correspondence breaks down, for it is possible in some cases for different articulations to result in rather similar acoustic effects, which may be perceived as identical by the listener. The Japanese example quoted above is an illustration of this, for the three sounds [ɸ], [ç], and [h], although articulated at widely differing points, nevertheless produce acoustic effects which are reasonably similar. We are, therefore, justified in assuming phonetic similarity for these sounds, in spite of the articulatory differences. Phonetic similarity in this instance is based on the auditory effect, and not on the articulation; it is listener- rather than speaker-oriented.

Examples of the same kind can also be found in English. For instance, [ʔ] ('glottal stop', used commonly in English dialects as a pronunciation of /t/ both finally, as in *wait*, and medially, as in *water*) and [t] are treated by native speakers as allophones of one phoneme, /t/. In word-final position [ʔ] is in free variation with [t°] (i.e. an unreleased [t]) or with a double articulation [ʔ͡t] in which closures are made simultaneously at the alveolum and at the glottis. It is often difficult to hear whether the /t/ of a word like *bit*, pronounced without a release, has been realized by [ʔ], [t°], or [ʔ͡t] (cf. Roach 1974). The auditory effect of these three sounds is almost identical for English-speakers, yet in articulatory terms their description differs greatly: one is alveolar, the other glottal, two points of articulation which could scarcely be further apart, and which are, moreover, separated by an intervening plosive, the velar [k].

These examples illustrate the point that phonetic similarity should not be thought of in purely articulatory terms. Sounds may differ greatly in place or manner of articulation, and yet share certain auditory properties which justify regarding them as similar. The traditional dominance of speaker-oriented description in phonetics has tended to make us think that sounds can only be similar if their articulation is similar. This is in fact often true, but not always.

Phonetic similarity can be determined *either* on an auditory *or* on an articulatory basis.

We have still not arrived at an objective answer to the question 'how similar is similar?', posed above, but we have at least widened the basis of phonetic similarity. In many cases, of course, the allophones of a phoneme resemble each other on both counts, but there are, as we have seen, some instances where a one-to-one correspondence does not hold. Looking again at the case of [h] and [ŋ] in English, we can now see that these sounds are dissimilar both in articulation *and* in perception, and therefore have no basis at all for being treated as phonetically similar. This makes it very unlikely that they could be members of a single phoneme, but the strongest argument in favour of their separateness (apart from the intuition of native speakers) rests on the concept of phonemic pattern, to which we now turn.

Phonemic pattern

A phonemic analysis needs to take account of the overall pattern of sounds, and their structural relationships, in the language concerned: phonemes are not isolated, unrelated phenomena.

To illustrate the principle in operation, take the example of [t^h] (as in *ten*) and [t] (as in *letter*). The aspirated/unaspirated distinction here can be extended to the other voiceless plosives /p, k/ in the same environments; aspirated [p^h, k^h] occur initially in a stressed syllable, unaspirated [p, k] occur intervocalically. Without the principle of phonemic pattern, we would be faced with a problem, for we can see that unaspirated [t] is in complementary distribution not only with [t^h] but also with [p^h] and [k^h], since [t] is found only intervocalically whereas [p^h, t^h, k^h] never occur in that position. Why then should we take [t] with [t^h], rather than with [p^h] or [k^h]? And similarly, why should [p] be associated with [p^h] rather than with [t^h] or [k^h]? The answer is, of course, that since we have aspirated and unaspirated variants of all three plosives we must pair them off, [p^h] with [p], etc. The problem has been raised only for the sake of demonstrating that the principle of complementary distribution, taken by itself at face value, is not enough; the overall pattern of the sounds must be taken into account.

We shall see more critical examples of the usefulness of this principle later on: note, however, that [h] and [ŋ] are not only phonetically dissimilar, but also pattern differently. [h] behaves (in