DYALA UNIVERSITY

ENGLISH DEPARTMENT

SECOND STAGE

PHONOLOGY

SECOND STAGE

للفصل الدراسي الثاني

English Phonetics and Phonology

A practical course

Third edition

PETER ROACH
Professor of Phonetics
University of Reading

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
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 1983
Third edition 2000
Fourth printing 2002

Printed in the United Kingdom at the University Press, Cambridge

Typeset in Times 11/14pt. *System* 3b2 [CE]

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

ISBN 0 521 78613 4 paperback
ISBN 0 521 79798 5 set of two cassettes
ISBN 0 521 79799 3 set of two audio CDs

Roach, Peter روچ، پیتر
(اینگلیش فونوتیکس اند فونولوژی...)
English Phonetics and Phonology - A practical course /
Peter Roach
تهران: جنگل، ۱۳۸۵=۲۰۰۶م.
xii، ۲۸۳ص، مصور.
فهرست نویسی بر اساس اطلاعات فیپا.
انگلیسی.
افست از روی ویراست سوم ۲۰۰۲: کمبریج.
۱. زبان انگلیسی -- آموزش زبان -- تلفظ. ۲. زبان انگلیسی --
واج شناسی. الف. عنوان: English Phonetics and Phonology- A
practical course.
PE۱۴۴۹ /۴و۹ل ۴۲۸/۱
۱۳۸۵
کتابخانه ملی ایران ۸۵-۱۷۱۱۴م

شناسنامه کتاب
نام کتاب: **English Phonetics and Phonology Third edition**
مؤلف: Peter Roach
تاریخ و نوبت چاپ: پنجم
تیراژ: ۳۰۰۰ جلد
مجتمع لیتوگرافی، چاپ و صحافی جنگل
ناشر: انتشارات جنگل (Jungle Publications)
تلفن:
021-66495275-66486115-9
0311-2212047-2239809
www.junglepub.org

8 The syllable

The syllable is a very important unit. Most people seem to believe that, even if they cannot define what a syllable is, they can count how many syllables there are in a given word or sentence. If they are asked to do this they often tap their finger as they count, which illustrates the syllable's importance in the rhythm of speech. As a matter of fact, if one tries the experiment of asking English speakers to count the syllables in, say, a recorded sentence, there is often a considerable amount of disagreement.

8.1 The nature of the syllable

When we looked at the nature of vowels and consonants in Chapter 1 it was shown that one could decide whether a particular sound was a vowel or a consonant on phonetic grounds (in relation to how much they obstructed the airflow) or on phonological grounds (vowels and consonants having different distributions). We find a similar situation with the syllable, in that it may be defined both phonetically and phonologically. Phonetically (that is, in relation to the way we produce them and the way they sound), syllables are usually described as consisting of a centre which has little or no obstruction to airflow and which sounds comparatively loud; before and after this centre (that is, at the beginning and end of the syllable), there will be greater obstruction to airflow and/or less loud sound. We will now look at some examples:

i) What we might call a **minimum syllable** would be a single vowel in isolation, e.g. the words 'are' ɑɪ, 'or' ɔɪ, 'err' ɜɪ. These are preceded and followed by silence. Isolated sounds such as m, which we sometimes produce to indicate agreement, or ʃ, to ask for silence, must also be regarded as syllables.

ii) Some syllables have an **onset** (that is, they have more than just silence preceding the centre of the syllable):

'bar' **bɑː** 'key' **kiː** 'more' **mɔː**

iii) Syllables may have no onset but have a **coda**:

'am' **æm** 'ought' **ɔːt** 'ease' **iːz**

iv) Some syllables have onset and coda:

'run' **rʌn** 'sat' **sæt** 'fill' **fɪl**

This is one way of looking at syllables. Looking at them from the phonological point of view is quite different. What this involves is looking at the possible combinations of English phonemes; the study of the possible phoneme combinations of a language is called **phonotactics**. It is simplest to start by looking at what can occur in initial position – in other words, what can occur at the beginning of the first word when we begin to speak after a pause. We find that the word can begin with a vowel, or with one, two or three consonants. No word begins with more than three consonants. In the same way, we can look at how a word ends when it is the last word spoken before a pause; it can end with a vowel, or with one, two, three or (in a small number of cases) four consonants. No word ends with more than four consonants.

8.2 The structure of the English syllable

Let us now look in more detail at syllable onsets. If the first syllable of the word in question begins with a vowel (any vowel may occur, though **ʊ** is rare) we say that this initial syllable has a **zero onset**. If the syllable begins with one consonant, that initial consonant may be any consonant phoneme except **ŋ**; **ʒ** is rare. We now look at syllables beginning with two consonants. When we have two or more consonants together we call them a **consonant cluster**.

Initial two-consonant clusters are of two sorts in English. One sort is composed of **s** followed by one of a small set of consonants; examples of such clusters are found in words such as 'sting' **stɪŋ**, 'sway' **sweɪ**, 'smoke' **sməʊk**. The **s** in these clusters is called the **pre-initial** consonant and the other consonant (**t**, **w**, **m** in the above examples) the **initial** consonant. These clusters are shown in Table 2.

Table 2 *Two-consonant clusters with pre-initial* **s**

Pre-initial	Initial																	
s *plus*	**p**	**t**	**k**	**b**	**d**	**g**	**f**	**θ**	**s**	**ʃ**	**h**	**v**	**ð**	**z**	**ʒ**	**m**	**n**	**ŋ**
	spɪn	**stɪk**	**skɪn**	–	–	–	**sfɪə**	–	–	–	–	–	–	–	–	**smel**	**snəʊ**	–

Note: Two-consonant clusters of **s** plus **l**, **w**, **j** are also possible (e.g. **slɪp**, **swɪŋ**, **sjuː**), and even perhaps **sr** in 'syringe' **srɪndʒ** for some speakers. These clusters can be analysed *either* as pre-initial **s** plus initial **l**, **w**, **j**, **r** *or* initial **s** plus post-initial **l**, **w**, **j**, **r**. There is no clear answer to the question of which analysis is better; here they are treated in the latter way, and appear in Table 3.

The other sort begins with one of a set of about fifteen consonants, followed by one of the set **l**, **r**, **w**, **j** as in, for example, 'play' **pleɪ**, 'try' **traɪ**, 'quick' **kwɪk**, 'few' **fjuː**. We call the first consonant of these clusters the **initial consonant** and the second the **post-initial**. There are some restrictions on which consonants can occur together. This can best be shown in table form, as shown in Table 3. When we look at three-consonant clusters we can recognise a clear relationship between them and the two sorts of two-consonant cluster described above; examples of three-consonant initial clusters are: 'split' **splɪt**, 'stream' **striːm**, 'square' **skweə**. The **s** is the pre-initial consonant, the **p**, **t** and **k** that follow **s** in the three example words are the initial consonant and the **l**, **r** and **w** are post-initial. In fact, the number of possible initial three-consonant clusters is quite small and they can be set out in full (words given in spelling form): AU8, Ex 2

		Post-initial			
		l	**r**	**w**	**j**
s plus initial	**p**	'splay'	'spray'	–	'spew'
	t	–	'string'	–	'stew'
	k	'sclerosis'	'screen'	'squeak'	'skewer'

AU8, Exs 3 & 4

We now have a similar task to do in studying final consonant clusters. Here we find the possibility of up to four consonants at the end of a word. If there is no final consonant we say that there is a **zero coda**. When there is one consonant only, this is called the **final** consonant. Any consonant may be a final consonant except **h**, **r**, **w**, **j**. There are two sorts of two-consonant final cluster, one being a final consonant preceded by a **pre-final** consonant and the other a final consonant followed by a **post-final** consonant. The pre-final consonants form a small set: **m**, **n**, **ŋ**, **l**, **s**. We can see these in 'bump' **bʌmp**, 'bent' **bent**, 'bank' **bæŋk**, 'belt' **belt**, 'ask' **ɑːsk**. The post-final consonants also form a small set: **s**, **z**, **t**, **d**, **θ**; example words are: 'bets' **bets**, 'beds' **bedz**, 'backed' **bækt**, 'bagged' **bægd**, 'eighth' **eɪtθ**. These post-final consonants can often be identified as separate morphemes (although not always, e.g. 'axe' **æks** is a single

Table 3 *Two-consonant clusters with post-initial* **l, r, w, j**

Post-initial	p	t	k	b	d	g	f	θ	s	ʃ	h	v	ð	z	ʒ	m	n	ŋ	l	r	w	j
l	**pleɪ**	–	**kleɪ**	**blæk**	–	**gluː**	**flaɪ**	–	**slɪp**	–	–	–	–	–	–	–	–	–	–	–	–	–
r	**preɪ**	**treɪ**	**kraɪ**	**brɪŋ**	**ˈdrɪp**	**grɪn**	**fraɪ**	**θrəʊ**	?[1]	**ʃruː**	–	–	–	–	–	–	–	–	–	–	–	–
w	–	**twɪn**	**kwɪk**	–	**dwel**	?[2]	–	**θwɔːt**	**swɪm**	?[3]	–	–	–	–	–	–	–	–	–	–	–	–
j	**pjɔː**	**tjuːn**	**kjuː**	**bjuːti**	**djuː**	?[4]	**fjuː**	?[5]	**sjuː**	–	**hjuːdʒ**	**vjuː**	–	–	–	**mjuːz**	**njuːz**	–	**ljuːd**	–	–	–

Notes on doubtful cases:

1 Some people pronounce the word 'syringe' as **srɪndʒ**; there are no other cases of **sr**, unless one counts foreign place names (e.g. Sri Lanka).

2 Many Welsh names (including some well-known outside Wales) – such as girls' names like Gwen and place names like the county of Gwent – have initial **gw** and English speakers seem to find them perfectly easy to pronounce.

3 Two cases make **ʃw** seem familiar: the vowel name 'schwa', and the name of the soft drinks brand Schweppes. This is, however, a very infrequent consonant cluster for English.

4 The only possible occurrence of **gj** would be in the archaic (heraldic) word 'gules', which is in very few people's vocabulary.

5 **θj** occurs in the archaic word 'thew' only.

morpheme and its final s has no separate meaning). A point of pronunciation can be pointed out here: the release of the first plosive of a plosive-plus-plosive cluster such as the g (of gd) in bægd or the k (of kt) in bækt is usually without plosion and is therefore practically inaudible.

AU8, Ex 5

There are two types of final three-consonant cluster; the first is pre-final plus final plus post-final, as set out in the following table:

		Pre-final	Final	Post-final
'helped'	he	l	p	t
'banks'	bæ	ŋ	k	s
'bonds'	bɒ	n	d	z
'twelfth'	twe	l	f	θ

The second type shows that more than one post-final consonant can occur in a final cluster: final plus post-final 1 plus post-final 2. Post-final 2 is again one of s, z, t, d, θ.

		Pre-final	Final	Post-final 1	Post-final 2
'fifths'	fɪ	–	f	θ	s
'next'	ne	–	k	s	t
'lapsed'	læ	–	p	s	t

Most four-consonant clusters can be analysed as consisting of a final consonant preceded by a pre-final and followed by post-final 1 and post-final 2, as shown below:

		Pre-final	Final	Post-final 1	Post-final 2
'twelfths'	twe	l	f	θ	s
'prompts'	prɒ	m	p	t	s

A small number of cases seem to require different analysis, as

consisting of a final consonant with no pre-final but three post-finals:

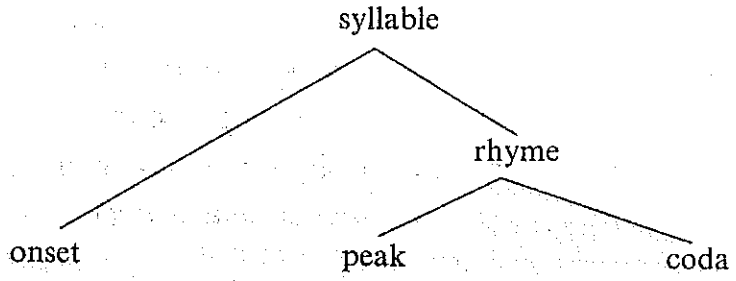
		Pre-final	Final	Post-final 1	Post-final 2	Post-final 3
'sixths'	sɪ	–	k	s	θ	s
'texts'	te	–	k	s	t	s

To sum up, we may describe the English syllable as having the following maximum phonological structure:

pre-initial	initial	post-initial	VOWEL	pre-final	final	post-final 1	post-final 2	post-final 3
Onset				Coda				

It will be noticed that there must be a vowel in the centre of the syllable. There is a special case, that of **syllabic consonants** (which are introduced in Chapter 9); we do not, for example, analyse the word 'students' **stjuːdnts** as consisting of one syllable with the three-consonant cluster **stj** for its onset and ending with a four-consonant cluster **dnts**. To fit in with what English speakers feel, we say that the word contains two syllables, with the consonant **d** dividing them and the second syllable ending with the cluster **nts**; in other words, we treat the word as though there was a vowel between **d** and **n**, although a vowel only occurs here in very slow, careful pronunciation. This phonological problem will be discussed in Chapter 13.

Recent work in phonology makes use of a rather more refined analysis of the syllable in which the vowel and the coda (if there is one) are known as the **rhyme**; if you think of rhyming English verse you will see that this works by matching just that part of the last syllable of a line. The rhyme is divided into the **peak** (normally the vowel) and the **coda** (but note that this is optional: the rhyme may have no coda, as in a word like 'me'). As we have seen, the syllable may also have an onset, but this is not obligatory. The structure is thus the following:



8.3 Syllable division

There are still problems with this phonetic description of the syllable: an unanswered question is how we decide on the division between syllables when we find a connected sequence of them as we usually do in normal speech. We will look at two words that are good examples of this difficulty. Most English speakers feel that the word 'going' **gəʊɪŋ** consists of two syllables; we could decide on phonetic grounds that the **ʊ** in the middle is the dividing point between the two syllables, since the articulation is slightly closer to obstructing airflow than the vowels next to it. This would not answer the question of whether the **ʊ** belongs to the first or to the second syllable; of course, we know that the **ʊ** is part of the **əʊ** diphthong phoneme, but this is a fact of phonology, not of the phonetic structure of the syllable. Another difficult case is the word 'extra' **ekstrə**. One problem is that by some definitions the **s** in the middle, between **k** and **t**, would be counted as a syllable, which most English speakers would reject. They feel that the word has two syllables. However, the most controversial issue relates to where the two syllables are to be divided; the possibilities are (using the symbol **.** to signify a syllable boundary):

i) **e.kstrə**
ii) **ek.strə**
iii) **eks.trə**
iv) **ekst.rə**
v) **ekstr.ə**

How can we decide on the division? No single rule will tell us what to do without bringing up problems.

One of the most widely accepted guidelines is what is known as the **maximum onsets principle**. This principle states that where two syllables are to be divided, any consonants between them should be

attached to the right-hand syllable, not the left, as far as possible. If we just followed this rule, we would have to divide 'extra' as (i) **e.kstrə**, but we know that an English syllable cannot begin **kstr**. Our rule must therefore state that consonants are assigned to the right-hand syllable as far as possible *within the restrictions governing syllable onsets and codas*. This means that we must reject (i) **e.kstrə** because of its impossible onset, and (v) **ekstr.ə** because of its impossible coda. We then have to choose between (ii), (iii) and (iv). The maximum onsets rule makes us choose (ii). However, there are many problems still remaining. For example, in looking at isolated syllables, we never find one ending with one of the vowels **ɪ, e, æ, ʌ, ɒ** or **ʊ**, so we must conclude that syllables with a short vowel and no coda do not occur in English (unless the vowel is **ə**, as will be explained in Chapter 9).

How, then, should we divide words like 'better' **betə**? The maximal onsets principle tells us to put the **t** on the right-hand syllable, giving **be.tə**, but that means that the first syllable is analysed as **be**, which we have just seen is not allowed. The maximal onsets principle must therefore also be modified to allow a consonant to be assigned to the left syllable if that prevents a short vowel from occurring at the end of a syllable. We can then analyse the word as **bet.ə**, which seems more satisfactory.

There are words like 'carry' **kæri**, however, which still give us problems: if we divide the word **kæ.ri**, we get a syllable-final **æ**, but if we divide it **kær.i** we have a syllable-final **r**, and both of these are non-occurring in BBC English. We have to decide on the lesser of two evils here, and the preferable solution is to divide the word as **kær.i** on the grounds that in the many rhotic accents of English (see Section 7.3) this division would be the natural one to make.

One further possibility should be mentioned: when one consonant stands between vowels and it is difficult to assign the consonant to one syllable or the other – as in 'better' and 'carry' – we could say that the consonant belongs to *both* syllables. The term used by phonologists for a consonant in this situation is **ambisyllabic**.

8.4 Practical conclusions

Analysing syllable structure, as we have been doing in this chapter, can be useful to foreign learners of English. Obviously there are

many more limitations on possible combinations of vowels and consonants, but an understanding of the basic structures described above will help learners to become aware of precisely what type of consonant cluster presents pronunciation problems – most learners find some English clusters difficult, but few find *all* of them difficult.

Notes on problems and further reading

The study of syllable structure is a subject of considerable interest to phonologists. If you want to read further in this area, I would recommend Giegerich (1992: Chapter 6), Katamba (1989: Chapter 9), Hogg and McCully (1987: Chapter 2) and Goldsmith (1990: Chapter 3). Some writers believe that it is possible to describe the possible combinations of phonemes with little reference to the syllable as an independent unit in theoretical phonology; see, for example, Harris (1994: Section 2.3). Cruttenden (1994: Chapter 10), Section 9, and Kreidler (1989: 117–38) describe the phonotactics of English in more detail.

A paper that has had a lot of influence on more recent work is Fudge (1969). This paper brings up two ideas first discussed by earlier writers: The first is that **sp**, **st**, **sk** could be treated as individual phonemes, removing the pre-initial position from the syllable onset altogether and removing **s** from the pre-final set of consonants; the second is that since post-initial **j** only occurs before **ʊ**, **uː** and **ʊə** (which in his analysis all begin with the same vowel), one could postulate a diphthong **ju** and remove **j** from post-initial position. These are interesting proposals, but there is not enough space here to examine the arguments in full.

There are many different ways of deciding how to divide syllables. To see two different approaches, see the Introductions to the *Longman Pronunciation Dictionary* (Wells, 2000) and the *English Pronouncing Dictionary* (15th edition, eds. Jones, Roach and Hartman, 1997: xiii).

Note for teachers

The last paragraph of Chapter 8 explains why the study of syllable structure is relevant to the learner of English. English has a more complex syllable structure than most languages, and it follows from what is said in this chapter that it is advisable to discover exactly

which types of consonant cluster are difficult for learners of a particular native-language background and construct exercises to give practice in them. There is discussion of this problem in Celce-Murcia *et al.* (1996: 80–9) and Dalton and Seidlhofer (1994: 34–8).

Written exercises

Using the analysis of the word 'cramped' given below as a model, analyse the structure of the following one-syllable English words:

	Initial	Post-initial		Pre-final	Final	Post-final
'cramped'	**k**	**r**	**æ**	**m**	**p**	**t**
	Onset		Peak		Coda	

a) squealed
b) eighths
c) splash
d) texts

9 Strong and weak syllables

9.1 Strong and weak

One of the most noticeable features of English is that some of its syllables are **strong** while many others are **weak**; this is also true of many other languages, but it is necessary to study how these weak syllables are pronounced and where they occur in English. The distribution of strong and weak syllables is a subject that will be met in several later chapters. For example, we will look later at **stress**, which is very important in deciding whether a syllable is strong or weak. **Elision** is a closely related subject, and in considering **intonation** the difference between strong and weak syllables is also important. Finally, words with "strong" and "weak" forms are clearly a related matter. In this chapter we look at the general nature of weak syllables.

What do we mean by "strong" and "weak"? To begin with, we can look at how we use these terms to refer to phonetic characteristics of syllables. When we compare weak syllables with strong syllables, we find the vowel in a weak syllable tends to be shorter, of lower intensity and different in quality. For example, in the word 'father' **fɑːðə** the second syllable, which is weak, is shorter than the first, is less loud and has a vowel that cannot occur in strong syllables. In a word like 'bottle' **bɒtl̩** the weak second syllable contains no vowel at all, but consists entirely of the consonant l̩. We call this a **syllabic consonant**.

There are other ways of characterising strong and weak syllables. We could describe them partly in terms of stress (by saying, for example, that strong syllables are stressed and weak syllables unstressed) but, until we describe what "stress" *means*, such a description would not be very useful. The most important thing to note at present is that any strong syllable will have as its peak one of the vowel phonemes (or possibly a triphthong) listed in Chapter 3, but

not ə, i or u. If the vowel is short, then the strong syllable will always have a coda as well. Weak syllables, on the other hand, as they are defined here, can only have one of a very small number of possible peaks. At the end of a word, we may have a weak syllable ending with a vowel (i.e. with no coda):

i) the vowel ə ("schwa");
ii) a close front unrounded vowel in the general area of **iː** and **ɪ** (symbolised **i**);
iii) a close back rounded vowel in the general area of **uː** and **ʊ** (symbolised **u**).

Examples would be:

i) 'better' **betə** ii) 'happy' **hæpi** iii) 'thank you' **θæŋk ju**

We also find weak syllables in word-final position with a coda if the vowel is ə. For example:

i) 'open' **əʊpən** ii) 'sharpen' **ʃɑːpən**

Inside a word, we can find the above vowels acting as peaks without codas in weak syllables; for example, look at the second syllable in each of these words:

i) 'photograph' **fəʊtəgrɑːf** ii) 'radio' **reɪdiəʊ**
iii) 'influence' **mfluəns**

In addition, the vowel ɪ can act as a peak without coda if the following syllable begins with a consonant:

iv) 'architect' **ɑːkɪtekt**

In the rest of this chapter we will look at the different types of weak syllable in more detail.

9.2 The ə vowel ("schwa")

AU9, Ex 1

The most frequently occurring vowel in English is ə, which is always associated with weak syllables. In quality it is mid (that is, half-way between close and open) and central (that is, half-way between front and back). It is generally described as lax, that is, not articulated with much energy. Of course, the quality of this vowel is not always the same, but the variation is not important.

Not all weak syllables contain ə, though many do. Learners of English need to learn where ə is appropriate and where it is not. To do this we often have to use information that traditional phonemic theory would not accept as relevant – we must consider spelling. The question to ask is: if the speaker were to pronounce a particular weak syllable as strong instead, which vowel would it be most likely to have, according to the usual rules of English spelling? Knowing this will not tell us which syllables in a word or utterance should be weak – that is something we look at in later chapters – but it will give us a rough guide to the correct pronunciation of weak syllables. Let us look at some examples:

i) Spelt with 'a'; strong pronunciation would have **æ**
   'a'ttend' **ətend** 'char*a*cter' **kærəktə**
   'barracks' **bærəks**
ii) Spelt with 'ar'; strong pronunciation would have **ɑː**
   'p*ar*ticular' **pətɪkjələ** 'mol*ar*' **məʊlə**
   'mon*ar*chy' **mɒnəki**
iii) Adjectival endings spelt 'ate'; strong pronunciation would have **eɪ**
   'intim*ate*' **ɪntɪmət** 'accur*ate*' **ækjərət**
   'desol*ate*' **desələt** (although there are exceptions to this: 'private' is usually **praɪvɪt**)
iv) Spelt with 'o'; strong pronunciation would have **ɒ** or **əʊ**
   't*o*morrow' **təmɒrəʊ** 'p*o*tato' **pəteɪtəʊ**
   'carr*o*t' **kærət**
v) Spelt with 'or'; strong pronunciation would have **ɔː**
   'f*or*get' **fəget** 'ambassad*or*' **æmbæsədə**
   'opp*or*tunity' **ɒpətjuːnəti**
vi) Spelt with 'e'; strong pronunciation would have **e**
   'settlem*e*nt' **setlmənt** 'viol*e*t' **varələt**
   'postm*e*n' **pəʊstmən**
vii) Spelt with 'er'; strong pronunciation would have **ɜː**
   'p*er*haps' **pəhæps** 'strong*er*' **strɒŋgə**
   'sup*er*man' **suːpəmæn**
viii) Spelt with 'u'; strong pronunciation would have **ʌ**
   'Aut*u*mn' **ɔːtəm** '*su*pport' **səpɔːt**
   'halib*u*t' **hælɪbət**

ix) Spelt with 'ough' (there are many pronunciations for the letter-sequence 'ough')

'thor*ough*' θʌrə 'bor*ough*' bʌrə

ix) Spelt with 'ou'; strong pronunciation might have aʊ

'graci*ous*' greɪʃəs 'call*ous*' kæləs

9.3 Close front and close back vowels

Two other vowels are commonly found in weak syllables, one close front (in the general region of iː and ɪ) and the other close back rounded (in the general region of uː and ʊ). In strong syllables it is comparatively easy to distinguish iː from ɪ, uː from ʊ, but in weak syllables the difference is not so clear. For example, although it is easy enough to decide which vowel one hears in 'beat' or 'bit', it is much less easy to decide which vowel one hears in the second syllable of words such as, for example, 'easy' or 'busy'. There are accents of English (for example Welsh accents) in which the second syllable sounds most like the iː in the first syllable of 'easy', and others (for example Yorkshire accents) in which it sounds more like the ɪ in the first syllable of 'busy'. In present-day BBC pronunciation, however, the matter is not so clear. There is uncertainty, too, about the corresponding close back rounded vowels. If we look at the words 'good to eat' and 'food to eat', we must ask if the word 'to' is pronounced with the ʊ vowel phoneme of 'good' or the uː phoneme of 'food'. Again, which vowel comes in 'to' in 'I want to'?

One common feature is that the vowels in question are more like iː or uː when they precede another vowel, less so when they precede a consonant or pause. You should notice one further thing: with the exception of one or two very artificial examples, there is really no possibility in these contexts of a phonemic contrast between iː and ɪ, or between uː and ʊ. Effectively, then, the two distinctions, which undoubtedly exist within strong syllables, are **neutralised** in weak syllables of BBC pronunciation. How should we transcribe the words 'easy' and 'busy'? We will use the close front unrounded case as an example, since it is more straightforward. The possibilities, using our phoneme symbols, are the following:

	'easy'	'busy'
i)	**iːziː**	**bɪziː**
ii)	**iːzɪ**	**bɪzɪ**

Few speakers with a BBC accent seem to feel satisfied with any of these transcriptions. There is a possible solution to this problem, but it goes against standard phoneme theory. We can symbolise this weak vowel as **i**, that is, using the symbol for the vowel in 'beat' but without the length mark. Thus:

iːzi **bɪzi**

The **i** vowel is neither the **iː** of 'beat' nor the **ɪ** of 'bit', and is not in contrast with them. We can set up a corresponding vowel **u** that is neither the **uː** of 'shoe' nor the **ʊ** of 'book' but a weak vowel that shares the characteristics of both. If we use **i** and **u** in our transcription as well as **iː**, **ɪ**, **uː** and **ʊ**, it is no longer a true phonemic transcription in the traditional sense. However, this need not be too serious an objection, and the fact that native speakers seem to think that this transcription fits better with their feelings about the language is a good argument in its favour.

AU9, Ex 2

Let us now look at where these vowels are found, beginning with close front unrounded ones. We find **i** occurring:

i) In word-final position in words spelt with final 'y' or 'ey' (after one or more consonant letters), e.g. 'happy' **hæpi**, 'valley' **væli**, and in morpheme-final position when such words have suffixes beginning with vowels, e.g. 'happier' **hæpiə**, 'easiest' **iːziəst**, 'hurrying' **hʌriŋ**.
ii) In a prefix such as those spelt 're', 'pre', 'de' if it precedes a vowel and is unstressed, for example in 'react' **riækt**, 'preoccupied' **priɒkjəpaɪd**, 'deactivate' **diæktɪveɪt**.
iii) In the suffixes spelt 'iate', 'ious' when they have two syllables, for example in 'appreciate' **əpriːʃieɪt**, 'hilarious' **hɪleəriəs**.
iv) In the following words when unstressed: 'he', 'she', 'we', 'me', 'be' and the word 'the' when it precedes a vowel.

In most other cases of syllables containing a short close front unrounded vowel we can assign the vowel to the **ɪ** phoneme, as in the

first syllable of 'resist' **rɪzɪst**, 'inane' **meɪn**, 'enough' **ɪnʌf**, the middle syllable of 'incident' **ɪnsɪdənt**, 'orchestra' **ɔːkɪstrə**, 'artichoke' **ɑɪtɪtʃəʊk**, and the final syllable of 'swimming' **swɪmɪŋ**, 'liquid' **lɪkwɪd**, 'optic' **ɒptɪk**. It can be seen that this vowel is most often represented in spelling by the letters 'i' and 'e'.

Weak syllables with close back rounded vowels are not so commonly found. We find **u** most frequently in the words 'you', 'to', 'into', 'do', when they are unstressed and are not immediately preceding a consonant, and 'through' and 'who' in all positions when they are unstressed. This vowel is also found before another vowel within a word, as in 'evacuation' **ɪvækjueɪʃn̩**, 'influenza' **ɪnfluenzə**.

9.4 Syllabic consonants

In the above sections we have looked at vowels in weak syllables. We must also consider syllables in which no vowel is found. In this case, a consonant, either **l**, **r** or a nasal, stands as the peak of the syllable instead of the vowel, and we count these as weak syllables like the vowel examples given earlier in this chapter. It is usual to indicate that a consonant is syllabic by means of a small vertical mark (ˌ) for example 'cattle' **kætl̩**.

Syllabic l

AU9, Ex 3

Syllabic **l** is perhaps the most noticeable example of the English syllabic consonants, although it would be wrong to expect to find it in all accents. It occurs after another consonant, and the way it is produced depends to some extent on the nature of that consonant. If the preceding consonant is alveolar, as in 'bottle' **bɒtl̩**, 'muddle' **mʌdl̩**, 'tunnel' **tʌnl̩**, the articulatory movement from the preceding consonant to the syllabic **l** is quite simple. The sides of the tongue, which are raised for the preceding consonant, are lowered to allow air to escape over them (this is called **lateral release**). The tip and blade of the tongue do not move until the articulatory contact for the **l** is released. The **l** is a "dark l" (as explained in Chapter 7). In some accents – particularly London ones, and "Estuary English" – we often find a close back rounded vowel instead (for example 'bottle' **bɒtu**). Where do we find syllabic **l̩** in the BBC accent? It is useful to look at the spelling as a guide. The most obvious case is where we have a word ending with one or more consonant letters followed by

'le' (or, in the case of noun plurals or third person singular verb forms, 'les'). Examples are:

i) with alveolar consonant preceding
   'cattle' **kætl̩** 'bottle' **bɒtl̩**
   'wrestle' **resl̩** 'muddle' **mʌdl̩**
ii) with non-alveolar consonant preceding
   'couple' **kʌpl̩** 'trouble' **trʌbl̩**
   'struggle' **strʌgl̩** 'knuckle' **nʌkl̩**

Such words usually lose their final letter 'e' when a suffix beginning with a vowel is attached, but the l usually remains syllabic. Thus:

'bottle' – 'bottling'	**bɒtl̩** – **bɒtlɪŋ**
'muddle' – 'muddling'	**mʌdl̩** – **mʌdlɪŋ**
'struggle' – 'struggling'	**strʌgl̩** – **strʌglɪŋ**

Similar words not derived in this way do not have the syllabic l̩ – it has been pointed out that the two words 'coddling' (derived from the verb 'coddle') and 'codling' (meaning "small cod", derived by adding the diminutive suffix '-ling' to 'cod') show a contrast between syllabic and non-syllabic l: 'coddling' **kɒdl̩ɪŋ** and 'codling' **kɒdlɪŋ**. In the case of words such as 'bottle', 'muddle', 'struggle', which are quite common, it would be a mispronunciation to insert a vowel between the l and the preceding consonant in the accent described here. There are many accents of English which may do this, so that, for example, 'cattle' is pronounced **kætəl**, but this is not the case in BBC pronunciation.

We also find syllabic l in words spelt, at the end, with one or more consonant letters followed by 'al' or 'el', for example:

'panel' **pænl̩**	'papal' **peɪpl̩**
'petal' **petl̩**	'parcel' **pɑɪsl̩**
'kernel' **kɜɪnl̩**	'Babel' **beɪbl̩**
'pedal' **pedl̩**	'ducal' **djuɪkl̩**

In some less common or more technical words, it is not obligatory to pronounce syllabic l and the sequence əl may be used instead, although it is less likely: 'missal' **mɪsl̩** or **mɪsəl**; 'acquittal' **əkwɪtl̩** or **əkwɪtəl**.

Syllabic **n**

AU9, Ex 4

Of the syllabic nasals, the most frequently found and the most important is **n̩**. When should it be pronounced? A general rule could be made that weak syllables which are phonologically composed of a plosive or fricative consonant plus **ən** are uncommon except in initial position in the words. So we can find words like 'tonight' **tənaɪt**, 'canary' **kəneəri** with an **ə** before **n**, but medially and finally – as in words like 'threaten', 'threatening' – we find much more commonly a syllabic **n**: **θretn̩**, **θretn̩ɪŋ**. To pronounce a vowel before the nasal consonant would sound strange (or at best overcareful) in BBC. Syllabic **n** is most common after alveolar plosives and fricatives; in the case of **t** and **d** followed by **n** the plosive is nasally released by lowering the soft palate, so that in the word 'eaten' **iːtn̩**, for example, the tongue does not move in the **tn̩** sequence but the soft palate is lowered at the end of **t** so that compressed air escapes through the nose. We do not find **n̩** after **l** or **tʃ**, **dʒ**, so that for example 'sullen' must be pronounced **sʌlən**, 'Christian' as **krɪstʃən** (though this word may be pronounced with **t** plus **i** or **j** instead of **tʃ**) and 'pigeon' as **pɪdʒən**.

Syllabic **n̩** after non-alveolar consonants is not so widespread. In words where the syllable following a velar consonant is spelt 'an' or 'on' (for example, 'toboggan', 'wagon') it is rarely heard, the more usual pronunciation being **təbɒgən**, **wægən**. After bilabial consonants, in words like 'happen', 'happening', 'ribbon' we can consider it equally acceptable to pronounce them with syllabic **n̩** (**hæpn̩**, **hæpn̩ɪŋ**, **rɪbn̩**) or with **ən** (**hæpən**, **hæpənɪŋ**, **rɪbən**). As we will see, syllabic **m̩** is also possible in this context. In a similar way, after velar consonants in words like 'thicken', 'waken', syllabic **n̩** is possible but **ən** is also acceptable. Syllabic velar nasal **ŋ̩** is also possible in this context.

After **f** or **v**, syllabic **n̩** is more common than **ən** (except, as with the other cases described, in word-initial syllables). Thus 'seven', 'heaven', 'often' are more usually **sevn̩**, **hevn̩**, **ɒfn̩** than **sevən**, **hevən**, **ɒfən**.

In all the examples given so far the syllabic **n̩** has been following another consonant; sometimes it is possible for another consonant to precede that consonant, but in this case a syllabic consonant is less likely to occur. If **n** is preceded by **l** and a plosive, as in 'Wilton', the

pronunciation **wɪltn̩** is possible, but **wɪltən** is also found regularly. If s precedes, as in 'Boston', a final syllabic nasal is less frequent, while clusters formed by nasal + plosive + syllabic nasal are very unusual: thus 'Minton', 'lantern', 'London', 'abandon' will normally have ə in the last syllable and be pronounced **mɪntən**, **læntən**, **lʌndən**, **əbændən**. Other nasals also discourage a following plosive plus syllabic nasal, so that for example 'Camden' is normally pronounced **kæmdən**.

Syllabics m, ŋ

We will not spend much time on the syllabic pronunciation of these consonants. Both can occur as syllabic, but only as a result of processes such as assimilation and elision that I have not yet described. We find them sometimes in words like 'happen', which can be pronounced **hæpm̩**, though **hæpn̩** and **hæpən** are equally acceptable, and 'uppermost', which could be pronounced as **ʌpm̩əust** though **ʌpəməust** would be more usual. Examples of possible syllabic velar nasals would be 'thicken' **θɪkŋ̩** (where **θɪkən** and **θɪkn̩** are also possible), and 'broken key' **brəukŋ̩ kiː**, where the nasal consonant occurs between velar consonants (again, **n̩** or **ən** could be substituted for **ŋ̩**).

A note about symbols: the usual convention for the syllabic mark is that it should be placed below symbols that do not come below the line, for example **m̩**, **n̩** but above a symbol that does come below the line, for example **ŋ̍**. In this course, however, it is felt preferable to put the mark underneath the symbol in all cases of syllabic consonants.

Syllabic r

In many accents of the type called "rhotic" (as explained in Chapter 7), such as most American accents, syllabic **r̩** is very common. The word 'particular', for example, would probably be pronounced **pr̩tɪkələr̩** in careful speech by most Americans, while BBC speakers would pronounce this word **pətɪkjələ**. Syllabic **r̩** is less common in BBC, and in most cases where it occurs there are perfectly acceptable alternative pronunciations without the syllabic consonant.

There are a few pairs of words (minimal pairs) in which a difference in meaning appears to depend on whether a particular **r** is syllabic or not, for example:

'Hungary' **hʌŋgr̩i** 'hungry' **hʌŋgri**

But we find no case of syllabic **r̩** where it would not be possible to substitute either non-syllabic **r** or **ər**; in the example above, 'Hungary' could equally well be pronounced **hʌŋgəri**.

Combinations of syllabic consonants

It is not unusual to find two syllabic consonants together. Examples are: 'national' **næʃn̩l̩**, 'literal' **lɪtr̩l̩**, 'visionary' **vɪʒn̩r̩i**, 'veteran' **vetr̩n̩**. It is important to remember that it is often not possible to say with certainty whether a speaker has pronounced a syllabic consonant, a non-syllabic consonant or a non-syllabic consonant plus **ə**. For example, the word 'veteran' given above could be pronounced in other ways than **vetr̩n̩**. A BBC speaker might instead say **vetrən**, **vetərn̩** or **vetərən**. The transcription makes it look as if the difference between these words is clear; it is not. In examining colloquial English it is often more or less a matter of arbitrary choice how one transcribes such a word. Transcription has the unfortunate tendency to make things seem simpler and more clear-cut than they really are.

Notes on problems and further reading

9.1 I have at this point tried to bring in some preliminary notions of stress and prominence without giving a full explanation. By this stage in the course it is important to be getting familiar with the difference between stressed and unstressed syllables, and the nature of "schwa". However, the subject of stress is such a large one that I have felt it best to leave its main treatment until later. On the subject of schwa, see Jones (1975: Sections 355–72); Cruttenden (1994: Section 8.9.12).

9.2 The introduction of **i** and **u** is a relatively recent idea, but it is now widely accepted as a convention in influential dictionaries such as the *Longman Dictionary of Contemporary English* (Summers, 1987), the *Longman Pronunciation Dictionary* (Wells, 2000), the *Cambridge International Dictionary of English* (Procter, 1995) and the *Daniel Jones English Pronouncing Dictionary* (15th edition; eds.

Roach and Hartman, 1997; see p. xiv of the Introduction to that dictionary). Since I mention native speakers' feelings in this connection, and since I am elsewhere rather sceptical about appeals to native speakers' feelings, I had better explain that in this case my evidence comes from the native speakers of English I have taught in practical classes on transcription over many years. A substantial number of these students have either been speakers with BBC pronunciation or had accents only slightly different from it, and their usual reaction to being told to use ɪ for the vowel at the end of 'easy', 'busy' has been one of puzzlement and frustration; like them, I cannot equate this vowel with the vowel of 'bit'. I am, however, reluctant to use iː, which suggests a stronger vowel than should be pronounced (like the final vowel in 'evacuee', 'Tennessee'). I must emphasise that the vowels **i** and **u** are not included in the set of English phonemes but are simply additional symbols to make the writing and reading of transcription easier. The Introduction to the *Daniel Jones English Pronouncing Dictionary* (eds. Roach and Hartman, 1997) discusses some of the issues involved in syllabic consonants and weak syllables: see pp. xiv–xv.

Notes for teachers

Introduction of the "schwa" vowel has been deliberately delayed until this chapter, since I wanted it to be presented in the context of weak syllables in general. Since students should by now be comparatively well informed about basic segmental phonetics, it is very important that their production and recognition of this vowel should be good before moving on to the following chapters.

This chapter is in a sense a crucial point in the course: although the segmental material of the preceding chapters is important as a foundation, the relationship between strong and weak syllables and the overall prosodic characteristics of words and sentences are essential to intelligibility, and most of the remaining chapters of the course are concerned with such matters.

Written exercises

The following sentences have been partially transcribed, but the vowels have been left blank. Fill in the vowels, taking care to identify which vowels are weak; put no vowel at all if you think a syllabic

consonant is appropriate, but put a syllabic mark beneath the syllabic consonant.

1 A particular problem of the boat was a leak
p t kj l pr bl m v ð b t w z l k

2 Opening the bottle presented no difficulty
p n ŋ ð b t l pr z nt d n d f k lt

3 There is no alternative to the Government's proposal
ð r z n lt n t v t ð g v nm nt spr p zl

4 We ought to make a collection to cover the expenses
w tt m k k l kʃ nt k v ð ksp ns z

5 Finally they arrived at a harbour at the edge of the mountains
f n l ð r v d t h b r t ð dʒ v ð m nt nz

10 Stress in simple words

10.1 The nature of stress

AU10, Ex 1

Stress has been mentioned several times already in this course without any attempt to define what the word means. The nature of stress is simple enough: practically everyone would agree that the first syllable of words like 'father', 'open', 'camera' is stressed, that the middle syllable is stressed in 'potato', 'apartment', 'relation' and that the final syllable is stressed in 'about', 'receive', 'perhaps'. Also, most people feel they have some sort of idea of what the difference is between stressed and unstressed syllables, although they might explain it in many different ways.

We will mark a stressed syllable in transcription by placing a small vertical line (') high up, just before the syllable it relates to; the words quoted above will thus be transcribed as follows:

'faɪðə	**pə'teɪtəʊ**	**ə'baʊt**
'əʊpən	**ə'pɑːtmənt**	**rɪ'siːv**
'kæmr̩ə	**rɪ'leɪʃn̩**	**pə'hæps**

What are the characteristics of stressed syllables that enable us to identify them? It is important to understand that there are two different ways of approaching this question. One is to consider what the speaker does in producing stressed syllables and the other is to consider what characteristics of sound make a syllable seem to a listener to be stressed. In other words we can study stress from the point of view of **production** and of **perception**; the two are obviously closely related, but are not identical. The production of stress is generally believed to depend on the speaker using more muscular energy than is used for unstressed syllables. Measuring muscular effort is difficult, but it seems possible, according to experimental studies, that when we produce stressed syllables, the muscles that we

use to expel air from the lungs are often more active, producing higher subglottal pressure. It seems probable that similar things happen with muscles in other parts of our speech apparatus.

Many experiments have been carried out on the perception of stress, and it is clear that many different sound characteristics are important in making a syllable recognisably stressed. From the perceptual point of view, all stressed syllables have one characteristic in common, and that is **prominence**. Stressed syllables are recognised as stressed because they are more prominent than unstressed syllables. What makes a syllable prominent? At least four different factors are important.

i) Most people seem to feel that stressed syllables are **louder** than unstressed; in other words, loudness is a component of prominence. In a sequence of identical syllables (e.g. **baːbaːbaːbaː**), if one syllable is made louder than the others, it will be heard as stressed. However, it is important to realise that it is very difficult for a speaker to make a syllable louder without changing other characteristics of the syllable such as those explained below (ii–iv); if one literally changes *only* the loudness, the perceptual effect is not very strong.

ii) The **length** of syllables has an important part to play in prominence. If one of the syllables in our "nonsense word" **baːbaːbaːbaː** is made longer than the others, there is quite a strong tendency for that syllable to be heard as stressed.

iii) Every voiced syllable is said on some **pitch**; pitch in speech is closely related to the frequency of vibration of the vocal folds and to the musical notion of low- and high-pitched notes. It is essentially a *perceptual* characteristic of speech. If one syllable of our "nonsense word" is said with a pitch that is noticeably different from that of the others, this will have a strong tendency to produce the effect of prominence. For example, if all syllables are said with low pitch except for one said with high pitch, then the high-pitched syllable will be heard as stressed and the others as unstressed. To place some **movement** of pitch (e.g. rising or falling) on a syllable is even more effective.

iv) A syllable will tend to be prominent if it contains a vowel that is different in **quality** from neighbouring vowels. If we change one

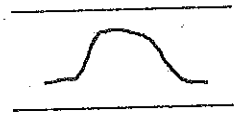
of the vowels in our "nonsense word" (e.g. **baːbiːbaːbaː**) the "odd" syllable **biː** will tend to be heard as stressed. This effect is not very powerful nor very important, but there is one particular way in which it is relevant in English: the previous chapter explained how the most frequently encountered vowels in weak syllables are **i**, **u**, **ɪ** and **ə** (syllabic consonants are also quite common). We can look on stressed syllables as occurring against a "background" of these weak syllables, so that their prominence is increased by contrast with these background qualities.

Prominence, then, is produced by four main factors: (i) loudness, (ii) length, (iii) pitch and (iv) quality. Generally these four factors work together in combination, although syllables may sometimes be made prominent by means of only one or two of them. Experimental work has shown that these factors are not equally important; the strongest effect is produced by pitch, and length is also a powerful factor. Loudness and quality have much less effect.

10.2 Levels of stress

Up to this point we have talked about stress as though there were a simple distinction between "stressed" and "unstressed" syllables with no intermediate levels; such a treatment would be a **two-level** analysis of stress. Usually, however, we have to recognise one or more intermediate levels. It should be remembered that in this chapter we are dealing only with stress *within the word*; this means that we are looking at words as they are said in isolation, which is a rather artificial situation: we do not often say words in isolation, except for a few such as 'yes', 'no', 'possibly', 'please' and interrogative words such as 'what', 'who', etc. However, looking at words in isolation does help us to see stress placement and stress levels more clearly than studying them in the context of continuous speech.

Let us begin by looking at the word 'around' **əˈraʊnd**, where the stress always falls clearly on the last syllable and the first syllable is weak. From the point of view of stress, the most important fact about the way we pronounce this word is that on the second syllable the pitch of the voice does not remain level, but usually falls from a higher to a lower pitch. We might diagram the pitch movement as shown below, where the two parallel lines represent the speaker's high and low pitch level:



The prominence that results from this pitch movement, or tone, gives the strongest type of stress; this is called **primary stress**.

In some words, we can observe a type of stress that is weaker than primary stress but stronger than that of the first syllable of 'around'; for example, in the first syllables of the words 'photographic' **fəʊtəgræfɪk**, 'anthropology' **ænθrəpɒlədʒi**. The stress in these words is called **secondary stress**. It is sometimes represented in transcription with a low mark (ˌ) so that the examples could be transcribed as **ˌfəʊtə'græfɪk**, **ˌænθrə'pɒlədʒi**. This convention will only be used where necessary in this course.

We have now identified two levels of stress: primary and secondary; this also implies a third level which can be called **unstressed** and is regarded as being the absence of any recognisable amount of prominence. These are the three levels that we will use in describing English stress. However, it is worth noting that unstressed syllables containing **ə**, **ɪ**, **i** or **u**, or a syllabic consonant will sound less prominent than an unstressed syllable containing some other vowel. For example, the first syllable of 'poetic' **pəʊ'etɪk** is more prominent than the first syllable of 'pathetic' **pə'θetɪk**. This *could* be used as a basis for a further division of stress levels, giving us a third ("tertiary") and fourth level. It is also possible to suggest a tertiary level of stress in some polysyllabic words. To take an example, it has been suggested that the word 'indivisibility' shows four different levels: the syllable **bɪl** is strongest (carrying primary stress), the initial syllable **ɪn** has secondary stress, while the third syllable **vɪz** has a level of stress which is weaker than those two but stronger than the second, fourth, sixth and seventh (which are all unstressed). Using the symbol ˌ to mark this tertiary stress, the word could be represented like this: **ˌɪndɪˌvɪzɪ'bɪlətɪ**. While this may be a phonetically correct account of some pronunciations, the introduction of tertiary stress seems to introduce an unnecessary degree of complexity.

10.3 Placement of stress within the word

We now come to a question that causes a great deal of difficulty, particularly to foreign learners (who cannot simply dismiss it as an

academic question): how can one select the correct syllable or syllables to stress in an English word? As is well known, English is not one of those languages where word stress can be decided simply in relation to the syllables of the word, as can be done in French (where the last syllable is usually stressed), Polish (where the syllable before the last – the penultimate syllable – is usually stressed) or Czech (where the first syllable is usually stressed). Many writers have said that English word stress is so difficult to predict that it is best to treat stress placement as a property of the individual word, to be learned when the word itself is learned. Certainly anyone who tries to analyse English stress placement has to recognise that it is a highly complex matter. However, it must also be recognised that in most cases (though certainly not all), when English speakers come across an unfamiliar word, they can pronounce it with the correct stress; in principle, it should be possible to discover what it is that the English speaker knows and to write it in the form of rules. The following summary of ideas on stress placement in nouns, verbs and adjectives is an attempt to present a few rules in the simplest possible form. Nevertheless, practically all the rules have exceptions and readers may feel that the rules are so complex that it would be easier to go back to the idea of learning the stress for each word individually.

In order to decide on stress placement, it is necessary to make use of some or all of the following information:

i) Whether the word is morphologically **simple**, or whether it is **complex** as a result either of containing one or more affixes (that is, prefixes or suffixes) or of being a compound word.
ii) What the grammatical category of the word is (noun, verb, adjective, etc.).
iii) How many syllables the word has.
iv) What the phonological structure of those syllables is.

It is sometimes difficult to make the decision referred to in (i). The rules for complex words are different from those for simple words and these will be dealt with in Chapter 11. Obviously, single-syllable words present no problems: if they are pronounced in isolation they are said with primary stress.

Point (iv) above is something that should be dealt with right away, since it affects many of the other rules that we will look at later. It is

possible to divide syllables into two basic categories: **strong** and **weak**. We saw in Chapter 8 that one component of a syllable is the **rhyme**, which contains the syllable peak and the coda. A strong syllable has a rhyme which *either* has a syllable peak which is a long vowel or diphthong, *or* a vowel followed by a coda (that is, one or more consonants). Weak syllables have a syllable peak which is a short vowel, and no coda unless the syllable peak is the schwa vowel ə or (in some circumstances) ɪ. Examples of strong syllables are:

'die' **daɪ** 'heart' **hɑɪt** 'bat' **bæt**

Examples of weak syllables (with syllable divisions shown) are:

're' in 'reduce' **rɪ.djuːs** 'bi' in 'herbicide' **hɜɪ.bɪ.saɪd** 'pen' in 'open' **əʊ.pən**

The important point to remember is that, although we do find unstressed strong syllables (as in the last syllable of 'dialect' **'daɪəlekt**), *only* strong syllables can be stressed. Weak syllables are always unstressed. This piece of knowledge does not by any means solve all the problems of how to place English stress, but it does help in some cases.

Two-syllable words

AU10, Ex 3

Here the choice is still simple: either the first or the second syllable will be stressed – not both. We will look first at verbs. The basic rule is that if the second syllable of the verb is a strong syllable, then that second syllable is stressed. Thus:

'apply' **ə'plaɪ** 'attract' **ə'trækt**
'arrive' **ə'raɪv** 'assist' **ə'sɪst**

If the final syllable is weak, then the first syllable is stressed. Thus:

'enter' **'entə** 'open' **'əʊpən**
'envy' **'envi** 'equal' **'iːkwəl**

A final syllable is also unstressed if it contains **əʊ** (e.g. 'follow' **'fɒləʊ**, 'borrow' **'bɒrəʊ**).

Two-syllable simple adjectives are stressed according to the same rule, giving:

'lovely' **'lʌvli** 'divine' **dɪ'vaɪn**
'even' **'iːvn̩** 'correct' **kə'rekt**
'hollow' **'hɒləʊ** 'alive' **ə'laɪv**

As with most stress rules, there are exceptions; for example, 'honest' **'ɒnɪst**, 'perfect' **'pɜɪfɪkt** (or **'pɜɪfekt**), both of which end with strong syllables but are stressed on the first syllable.

Nouns require a different rule: if the second syllable contains a short vowel, then the stress will usually come on the first syllable. Otherwise it will be on the second syllable.

'money' **'mʌni** 'estate' **ɪ'steɪt**
'product' **'prɒdʌkt** 'balloon' **bə'luːn**
'larynx' **'lærɪŋks** 'design' **dɪ'zaɪn**

Other two-syllable words such as adverbs and prepositions seem to behave like verbs and adjectives.

Three-syllable words

Here we find a more complicated picture. In verbs, if the final syllable is strong, then it will be stressed. Thus:

'entertain' **entə'teɪn** 'resurrect' **rezə'rekt**

If the last syllable is weak, then it will be unstressed, and stress will be placed on the preceding (penultimate) syllable if that syllable is strong. Thus:

'encounter' **ɪŋ'kaʊntə** 'determine' **dɪ'tɜɪmɪn**

If both the second and third syllable are weak, then the stress falls on the initial syllable:

'parody' **'pærədi**

Nouns require a slightly different rule. Here, if the final syllable is weak, or ends with **əʊ**, then it is unstressed; if the syllable preceding this final syllable is strong, then that middle syllable will be stressed. Thus:

'mimosa' **mɪ'məʊzə** 'disaster' **dɪ'zɑɪstə**
'potato' **pə'teɪtəʊ** 'synopsis' **sɪ'nɒpsɪs**

If the second and third syllables are both weak, then the first syllable is stressed:

'quantity' **'kwɒntəti** 'emperor' **'empərə**
'cinema' **'sɪnəmə** 'custody' **'kʌstədi**

Most of the above rules show stress tending to go on strong syllables. However, three-syllable simple nouns are different. Even if the final

syllable is strong, the stress will usually be placed on the first syllable. The last syllable is usually quite prominent, so that in some cases it could be said to have secondary stress.

'intellect' **'ɪntəlekt**	'marigold' **'mærɪgəʊld**
'alkali' **'ælkəlaɪ**	'stalactite' **'stæləktaɪt**

Adjectives seem to need the same rule, to produce stress patterns such as:

'opportune' **'ɒpətjuːn**	'insolent' **'ɪnsələnt**
'derelict' **'derəlɪkt**	'anthropoid' **'ænθrəpɔɪd**

The above rules do not, of course, cover all English words. They apply only to major categories of lexical words (nouns, verbs and adjectives in this chapter), not to function words such as articles and prepositions. There is not enough space in this course to deal with simple words of more than three syllables, nor with special cases of loan words (words brought into the language from other languages comparatively recently). Complex and compound words are dealt with in Chapter 11. One problem that we must also leave until Chapter 11 is the fact that there are many cases of English words with alternative possible stress patterns (e.g. 'controversy' as either **'kɒntrəvɜːsi** or **kən'trɒvəsi**). Other words – which we will look at in studying connected speech – change their stress pattern according to the context they occur in. Above all, there is not space to discuss the many exceptions to the above rules. Despite the exceptions, it seems better in many ways to attempt to produce some stress rules (even if they are rather crude and inaccurate) than to claim that there is no rule or regularity in English word stress.

Notes on problems and further reading

The subject of English stress has received a large amount of attention, and the references given here are only a small selection from an enormous number. As I implied in the notes on the previous chapter, incorrect stress placement is a major cause of intelligibility problems for foreign learners, and is therefore a subject that needs to be treated very seriously.

10.1 I have deliberately avoided using the term *accent*, which is found widely in the literature on stress. This is for three main reasons:

i) It increases the complexity of the description without, in my view, contributing much to its value.
ii) Different writers do not agree with each other about the way the term should be used.
iii) The word 'accent' is used elsewhere to refer to different varieties of pronunciation (e.g. "a foreign accent"); it is confusing to use it for a quite different purpose – to a lesser extent we also have this problem with the word 'stress', which can be used to refer to psychological tension.

There is a good discussion of the confusing nature of the terms 'stress' and 'accent' in Clark and Yallop (1995: 41–2). Their Section 9.6 on stress in English (pp. 348–57) is a useful summary. For a review of the production and perception of stress, see Laver (1994: 512–17).

10.2 On the question of the number of levels of stress, in addition to Laver (1994: 516), see also Wells (2000).

10.3 It is said in this chapter that one may take one of two positions. One is that stress is not predictable by rule and must be learned word by word (see, for example, Jones 1975: Sections 920–1). The second (which I prefer) is to say that, difficult though the task is, one must try to find a way of writing rules that express what native speakers naturally tend to do in placing stress (while acknowledging that there will always be a substantial residue of cases which appear to follow no regular rules). A very thorough treatment is given by Fudge (1984). More recently, Giegerich (1992) has presented a clear analysis of English word stress (including a useful explanation of strong and weak syllables) that is well worth reading; see p. 146 and Chapter 7.

There is another approach to English stress rules which is radically different. This is based on **generative phonology**, an analysis which was first presented in Chomsky and Halle (1968) and has been followed by a large number of works exploring the same field. To anyone not familiar with this type of treatment, the presentation will seem difficult or even unintelligible; within the generative approach, many different theories, all with different names, tend to come and

go with changes in fashion. The following paragraph is an attempt to summarise the main characteristics of basic generative phonology, and recommends some further reading for those interested in learning about it in detail.

The level of phonology is very abstract in this theory. An old-fashioned view of speech communication would be that what the speaker intends to say is coded – or *represented* – as a string of phonemes just like a phonemic transcription, and what a hearer hears is also converted by the brain from sound waves into a similar string of phonemes. A generative phonologist, however, would say that this phonemic representation is not accurate; the representation in the brain of the speaker or hearer is much more abstract and is often quite different from the 'real' sounds recognisable in the sound wave. You may hear the word 'football' pronounced as **fʊpbɔːl**, but your brain recognises the word as made up of 'foot' and 'ball' and interprets it phonologically as **fʊtbɔːl**. You may hear **ə** in the first syllable of 'photography', in the second syllable of 'photograph' and in the third syllable of 'photographer', but the brain recognises links between these **ə** vowels and **əʊ**, **ɒ** and **æ** respectively, and supplies the *underlying* vowels. In speaking, underlying segments may be realised as different sounds as the stress pattern changes. These vowel changes are brought about by *rules* – not the sort of rules that one might teach to language learners, but more like the instructions that one might build into a machine or write into a computer program. According to Chomsky and Halle, at the abstract phonological level words do not possess stress; stress (of many different levels) is the result of the application of phonological rules, which are simple enough in theory but highly complex in practice. The principles of these rules are explained first on pp. 15–43 of Chomsky and Halle (1968), and in greater detail on pp. 69–162.

There is a clear and thorough introductory account of generative phonology in Clark and Yallop (1995: Chapter 5), and they present a brief account of the generative treatment of stress on pp. 355–7. A briefer review is given in Katamba (1989: Chapter 11, Section l).

Notes for teachers

It should be clear from what is said above that from the purely practical classroom point of view, explaining English word stress in

terms of generative phonology could well create confusion for learners. Finding practice and testing material for word stress is very simple: any modern English dictionary shows word stress patterns as part of word entries, and lists of these can be made either with stress marks for student to read from (as in Exercise 2 of Audio Unit 10), or without stress marks for students to put their own marks on (as in Exercise 1 of the same Audio Unit).

Written exercises (mainly for foreign learners)
Mark the stress on the following words:

1 verbs

a) protect	*e)* bellow
b) clamber	*f)* menace
c) festoon	*g)* disconnect
d) detest	*h)* entering

2 nouns

a) language	*e)* event
b) captain	*f)* jonquil
c) career	*g)* injury
d) paper	*h)* connection

(Native speakers of English should transcribe the words phonemically as well as marking stress.)

11 Complex word stress

11.1 Complex words

In Chapter 10 the nature of stress was explained and some broad general rules were given for deciding which syllable in a word should receive primary stress. The words that were described were called "simple" words; "simple" in this context means "not composed of more than one grammatical unit", so that, for example, the word 'care' is simple while 'careful' and 'careless' (being composed of two grammatical units each) are complex; 'carefully' and 'carelessness' are also complex, and are composed of three grammatical units each. Unfortunately, as was suggested in Chapter 10, it is often difficult to decide on whether a word should be treated as complex or simple. The majority of English words of more than one syllable (**polysyllabic words**) have come from other languages whose way of constructing words is easily recognisable; for example, we can see how combining 'mit' with the prefixes 'per-', 'sub-', 'com-' produced 'permit', 'submit', 'commit', words which have come into English from Latin. Similarly, Greek has given us 'catalogue', 'analogue', 'dialogue', 'monologue', in which the prefixes 'cata-', 'ana-', 'dia-', 'mono-' are recognisable. But we cannot automatically treat the separate grammatical units of other languages as separate grammatical units of English. If we did, we would not be able to study English morphology without first studying the morphology of five or six other languages, and we would be forced into ridiculous analyses such as that the English word 'parallelepiped' is composed of four or five grammatical units (which is the case in Ancient Greek). We must accept, then, that the distinction between "simple" and "complex" words is difficult to draw, and is therefore not always useful.

Complex words are of two major types:

i) words made from a basic word form (which we will call the **stem**), with the addition of an **affix**; and
ii) **compound words**, which are made of two (or occasionally more) independent English words (e.g. 'ice-cream', 'armchair').

We will look first at the words made with affixes. Affixes are of two sorts in English: **prefixes**, which come before the stem (e.g. prefix 'un-' + stem 'pleasant' → 'unpleasant') and **suffixes**, which come after the stem (e.g. stem 'good' + suffix '-ness' → 'goodness').

Affixes have one of three possible effects on word stress:

i) The affix itself receives the primary stress (e.g. 'semi-' + 'circle' **'sɜːkl̩** → 'semicircle' **'semɪsɜːkl̩**; '-ality' + 'person' **'pɜːsn̩** 'personality' **pɜːsn̩'ælətɪ**).
ii) The word is stressed just as if the affix were not there (e.g. 'pleasant' **'plezn̩t**, 'unpleasant' **ʌn'plezn̩t**; 'market' **'mɑːkɪt**, 'marketing' **'mɑːkɪtɪŋ**).
iii) The stress remains on the stem, not the affix, but is shifted to a different syllable (e.g. 'magnet' **'mægnət**, 'magnetic' **mæg'netɪk**).

11.2 Suffixes

There are so many suffixes that it will only be possible here to examine a small proportion of them; we will concentrate on those which are common and **productive** (that is, are applied to a considerable number of stems and could be applied to more to make new English words). In the case of the others, foreign learners would probably be better advised to learn the 'stem + affix' combination as an individual item.

One of the problems that will be encountered is that we may find words which are obviously complex but which, when we divide them into stem + affix, turn out to have a stem that is difficult to imagine as an English word. For example, the word 'audacity' seems to be a complex word – but what is its stem? Another problem is that it is difficult in some cases to know whether a word has one, or more than one, suffix (e.g. should we analyse 'personality', from the point of view of stress assignment, as **pɜsn̩ + ælətɪ** or as **pɜːsn̩ + æl + ətɪ**?). In the study of English word formation at a deeper level than we can go into here, it is necessary for such reasons to distinguish between a stem (which is what remains when affixes are removed), and a **root**,

which is the smallest piece of lexical material that a stem can be reduced to. So, in 'personality', we could say that the *suffix* '-ity' is attached to the *stem* 'personal' which contains the *root* 'person'. We will not spend more time here on looking at these problems, but go on to look at some generalisations about suffixes and stress, using only the term 'stem' for the sake of simplicity. The suffixes are referred to in their spelling form.

Suffixes carrying primary stress themselves

AU11, Ex 1

In the examples given, which seem to be the most common, the primary stress is on the first syllable of the suffix. If the stem consists of more than one syllable there will be a secondary stress on one of the syllables of the stem. This cannot fall on the last syllable of the stem and is, if necessary, moved to an earlier syllable. For example, in 'Japan' **dʒə'pæn** the primary stress is on the last syllable, but when we add the stress-carrying suffix '-ese' the primary stress is on the suffix and the secondary stress is placed not on the second syllable but on the first: 'Japanese' **ˌdʒæpə'niːz**.

- '-ee': 'refugee' **ˌrefjʊ'dʒiː**; 'evacuee' **ɪˌvækju'iː**
- '-eer': 'mountaineer' **ˌmaʊntɪ'nɪə**; 'volunteer' **ˌvɒlən'tɪə**
- '-ese': 'Portuguese' **ˌpɔːtʃə'giːz**; 'journalese' **ˌdʒɜːnl̩'iːz**
- '-ette': 'cigarette' **ˌsɪgɹ̩'et**; 'launderette' **ˌlɔːndɹ̩'et**
- '-esque': 'picturesque' **ˌpɪktʃɹ̩'esk**

Suffixes that do not affect stress placement

AU11, Ex 2

- '-able': 'comfort' **'kʌmfət**; 'comfortable' **'kʌmftəbl̩**
- '-age': 'anchor' **'æŋkə**; 'anchorage' **'æŋkərɪdʒ**
- '-al': 'refuse' (verb) **rɪ'fjuːz**; 'refusal' **rɪ'fjuːzl̩**
- '-en': 'wide' **waɪd**; 'widen' **'waɪdn̩**
- '-ful': 'wonder' **'wʌndə**; 'wonderful' **'wʌndəfl̩**
- '-ing': 'amaze' **ə'meɪz**; 'amazing' **ə'meɪzɪŋ**
- '-ish': 'devil' **'devl̩**; 'devilish' **'devl̩ɪʃ**

(This is the rule for adjectives; verbs with stems of more than one syllable always have the stress on the syllable immediately preceding 'ish', e.g. 'replenish' **rɪ'plenɪʃ**, 'demolish' **dɪ'mɒlɪʃ**.)

- '-like': 'bird' **'bɜːd**; 'birdlike' **'bɜːdlaɪk**
- '-less': 'power' **'paʊə**; 'powerless' **'paʊələs**
- '-ly': 'hurried' **'hʌrɪd**; 'hurriedly' **'hʌrɪdli**
- '-ment' (noun): 'punish' **'pʌnɪʃ**; 'punishment' **'pʌnɪʃmənt**
- '-ness': 'yellow' **'jeləʊ**; 'yellowness' **'jeləʊnəs**
- '-ous': 'poison' **'pɔɪzn̩**; 'poisonous' **'pɔɪzn̩əs**
- '-fy': 'glory' **'glɔːri**; 'glorify' **'glɔːrɪfaɪ**
- '-wise': 'other' **'ʌðə**; 'otherwise' **'ʌðəwaɪz**
- '-y' (adjective or noun): 'fun' **'fʌn**; 'funny' **'fʌni**

Suffixes that influence stress in the stem

AU11, Ex 3

In these examples primary stress is on the last syllable of the stem.

- '-eous': 'advantage' **əd'vɑːntɪdʒ**; 'advantageous' **ˌædvən'teɪdʒəs**
- '-graphy': 'photo' **'fəʊtəʊ**; 'photography' **fə'tɒgrəfi**
- '-ial': 'proverb' **'prɒvɜːb**; 'proverbial' **prə'vɜːbiəl**
- '-ic': 'climate' **'klaɪmɪt**; 'climatic' **klaɪ'mætɪk**
- '-ion': 'perfect' **'pɜːfɪkt**; 'perfection' **pə'fekʃn̩**
- '-ious': 'injure' **'ɪndʒə**; 'injurious' **ɪn'dʒɔːriəs**
- '-ty': 'tranquil' **'træŋkwɪl**; 'tranquillity' **træŋ'kwɪlətɪ**
- '-ive': 'reflex' **'riːfleks**; 'reflexive' **rɪ'fleksɪv**

Finally, when the suffixes '-ance', '-ant' and '-ary' are attached to single-syllable stems, the stress is almost always placed on the stem. When the stem has more than one syllable, the stress is on one of the syllables in the stem. To explain this we need to use a rule based on syllable structure, as was done for simple words in the previous chapter. If the final syllable of the stem is strong, that syllable receives the stress. For example: 'importance' **ɪm'pɔːtn̩s**; 'centenary' **sen'tiːnr̩i**. Otherwise the syllable *before* the last one receives the stress: 'inheritance' **ɪn'herɪtəns**; 'military' **'mɪlɪtri**.

11.3 Prefixes

We will only deal briefly with prefixes. Their effect on stress does not have the comparative regularity, independence and predictability of suffixes, and there is no prefix of one or two syllables that always carries primary stress. Consequently, the best treatment seems to be to say that stress in words with prefixes is governed by the same rules as those for words without prefixes.

11.4 Compound words

AU11, Ex 4

The words discussed so far in this chapter have all consisted of a stem plus an affix. We now pass on to another type of word. This is called **compound**, and its main characteristic is that it can be analysed into two words, both of which can exist independently as English words. (Some compounds are made of more than two words, but we will not consider these.) As with many of the distinctions being made in connection with stress, there are areas of uncertainty. For example, it could be argued that 'photograph' may be divided into two independent words, 'photo' and 'graph'; yet we usually do not regard it as a compound, but as a simple word. (If, however, someone drew a graph displaying numerical information about photos, this would perhaps be called a 'photo-graph' and the word would be regarded as a compound.) Compounds are written in different ways; sometimes they are written as one word – e.g. 'armchair', 'sunflower' – sometimes with the words separated by a hyphen – e.g. 'gear-change', 'fruit-cake' – and sometimes with two words separated by a space – e.g. 'desk lamp', 'battery charger'. In this last case there would, of course, be no indication to the foreign learner that the pair of words was to be treated as a compound. There is no clear dividing line between two-word compounds and pairs of words that simply happen to occur together quite frequently.

As far as stress is concerned, the question is quite simple. When is primary stress placed on the first constituent word of the compound and when on the second? Both patterns are found. A few rules can be given, although these are not completely reliable. Words which do not receive primary stress normally have secondary stress, although for the sake of simplicity this is not marked here. Perhaps the most familiar type of compound is the one which combines two nouns and which normally has the stress on the first element, as in:

'typewriter' **'taɪpraɪtə**
'car-ferry' **'kɑːferi**
'sunrise' **'sʌnraɪz**
'suitcase' **'suːtkeɪs**
'tea-cup' **'tiːkʌp**

It is probably safest to assume that stress will normally fall in this way on other compounds; however, a variety of compounds receive stress instead on the second element. For example, compounds with

an adjectival first element and the *-ed* morpheme at the end have this pattern (given in spelling only):

bad-'tempered
half-'timbered
heavy-'handed

Compounds in which the first element is a number in some form also tend to have final stress:

three-'wheeler
second-'class
five-'finger

Compounds functioning as adverbs are usually final-stressed:

head-'first
North-'East
down'stream

Finally, compounds which function as verbs and have an adverbial first element take final stress:

down-'grade
back-'pedal
ill-'treat

11.5 Variable stress

It would be wrong to imagine that the stress pattern is always fixed and unchanging in English words. Stress position may vary for one of two reasons: either as a result of the stress on other words occurring next to the word in question, or because not all speakers agree on the placement of stress in some words. The former case is an aspect of connected speech that will be encountered again in Chapter 14: the main effect is that the stress on a final-stressed compound tends to move to a preceding syllable if the following word begins with a strongly stressed syllable. Thus (using some examples from the previous section):

bad-'tempered	but	a 'bad-tempered 'teacher
half-'timbered	but	a 'half-timbered 'house
heavy-'handed	but	a 'heavy-handed 'sentence

The second is not a serious problem, but is one that foreign learners should be aware of. A well-known example is 'controversy', which is pronounced by some speakers as **'kɒntrəvɜɪsi** and by others as **kən'trɒvəsi**; it would be quite wrong to say that one version was correct and one incorrect. Other examples of different possibilities are 'ice-cream' (either **ˌaɪs'kriːm** or **'aɪskriːm**), 'kilometre' (either **'kɪləmiːtə** or **kɪ'lɒmɪtə**) and 'formidable' (**'fɔːmɪdəbl̩** or **fɔɪ'mɪdəbl̩**).

11.6 Word-class pairs

AU11, Ex 5

One aspect of word stress is best treated as a separate issue. There are several dozen pairs of two-syllable words with identical spelling which differ from each other in stress placement, apparently according to word class (noun, verb or adjective). All appear to consist of prefix + stem. We shall treat them as a special type of word and give them the following rule: if a pair of prefix-plus-stem words exists, both members of which are spelt identically, one of which is a verb and the other of which is either a noun or an adjective, then the stress is placed on the second syllable of the verb but on the first syllable of the noun or adjective. Some common examples are given below (V = verb, A = adjective, N = noun):

'abstract'	**'æbstrækt** (A)	**æb'strækt** (V)
'conduct'	**'kɒndʌkt** (N)	**kən'dʌkt** (V)
'contract'	**'kɒntrækt** (N)	**kən'trækt** (V)
'contrast'	**'kɒntrɑːst** (N)	**kən'trɑɪst** (V)
'desert'	**'dezət** (N)	**dɪ'zɜɪt** (V)
'escort'	**'eskɔɪt** (N)	**ɪs'kɔɪt** (V)
'export'	**'ekspɔɪt** (N)	**ɪk'spɔɪt** (V)
'import'	**'ɪmpɔɪt** (N)	**ɪm'pɔɪt** (V)
'insult'	**'ɪnsʌlt** (N)	**ɪn'sʌlt** (V)
'object'	**'ɒbdʒɪkt** (N)	**əb'dʒekt** (V)
'perfect'	**'pɜɪfɪkt** (A)	**pə'fekt** (V)
'permit'	**'pɜɪmɪt** (N)	**pə'mɪt** (V)
'present'	**'prezn̩t** (N, A)	**prɪ'zent** (V)
'produce'	**'prɒdjuɪs** (N)	**prə'djuɪs** (V)
'protest'	**'prəʊtest** (N)	**prə'test** (V)
'rebel'	**'rebl̩** (N)	**rɪ'bel** (V)
'record'	**'rekɔɪd** (N)	**rɪ'kɔɪd** (V)
'subject'	**'sʌbdʒɪkt** (N)	**səb'dʒekt** (V)

Notes on problems and further reading

Most of the reading recommended in the notes for the previous chapter is relevant for this one too. Looking specifically at compounds, it is worth reading Fudge (1984: Chapter 5). See also Cruttenden (1994: 207–9). If you wish to go more deeply into compound-word stress, you should first study English word formation. Recommended reading for this is Bauer (1983). On the distinction between *stem* and *root*, see Radford *et al.* (1999), pp. 67–8.

Written exercises

1 Put stress marks on the following words (try to put secondary stress marks on as well).

a) shop-fitter *b)* open-ended *c)* Javanese

d) birth-mark *e)* anti-clockwise *f)* confirmation

g) eight-sided *h)* fruit-cake *i)* defective

j) roof-timber

2 Write the words in phonemic transcription, including the stress marks.

12 Weak forms

Chapter 9 discussed the difference between strong and weak syllables in English. We have now moved on from looking at syllables to looking at words. We will now consider certain well-known English words that can be pronounced in two different ways; these are called **strong forms** and **weak forms**. As an example, the word 'that' can be pronounced **ðæt** (strong form) or **ðət** (weak form). The sentence 'I like that' is pronounced **aɪ laɪk ðæt** (strong form); the sentence 'I hope that she will' is pronounced **aɪ həʊp ðət ʃi wɪl** (weak form). There are roughly forty such words in English. It is possible to use only strong forms in speaking, and some foreigners do this. Usually they can still be understood by other speakers of English, so why is it important to learn how weak forms are used? There are two main reasons; first, most native speakers of English find an "all-strong-form" pronunciation unnatural and foreign-sounding, something that most learners would wish to avoid. Second, and more importantly, speakers who are not familiar with the use of weak forms are likely to have difficulty understanding speakers who do use weak forms; since practically all native speakers of British English use them, learners of the language need to learn about these weak forms to help them to understand what they hear.

We must distinguish between weak forms and **contracted forms**. Certain English words are shortened so severely (usually to a single phoneme) and so consistently that they are represented differently in informal writing, e.g. 'it is' – 'it's'; 'we have' – 'we've'; 'do not' – 'don't'. These contracted forms are discussed in a later chapter, and are not included here.

Almost all the words which have both a strong and weak form belong to a category that may be called **function words** – words that do not have a dictionary meaning in the way that we normally expect

nouns, verbs, adjectives and adverbs to have. These function words are words such as auxiliary verbs, prepositions, conjunctions, etc., all of which are in certain circumstances pronounced in their strong forms but which are more frequently pronounced in their weak forms. It is important to remember that there are certain contexts where only the strong form is acceptable, and others where the weak form is the normal pronunciation. There are some fairly simple rules; we can say that the strong form is used in the following cases:

i) For many weak-form words, when they occur at the end of a sentence; for example, the word 'of' has the weak form **əv** in the following sentence:

'I'm fond of chips' **aɪm 'fɒnd əv 'tʃɪps**

However, when it comes at the end of the sentence, as in the following example, it has the strong form **ɒv**:

'Chips are what I'm fond of' **'tʃɪps ə 'wɒt aɪm 'fɒnd ɒv**

Many of the words given below (particularly the first nine) never occur at the end of a sentence, e.g. 'the', 'your'. Some words (particularly the pronouns numbered 10–14 below) do occur in their weak forms in final position.

ii) When a weak-form word is being contrasted with another word; for example:

'The letter's *from* him, not *to* him' **ðə 'letəz 'frɒm ɪm nɒt 'tuː ɪm**

A similar case is what we might call a **co-ordinated** use of prepositions:

'I travel to and from London a lot' **aɪ 'trævl̩ 'tuː ən 'frɒm 'lʌndən ə 'lɒt**
'A work of and about literature' **ə 'wɜɪk 'ɒv ən ə'baʊt 'lɪtrɪtʃə**

iii) When a weak-form word is given stress for the purpose of emphasis; for example:

'You *must* give me more money' **ju 'mʌst 'gɪv mi 'mɔː 'mʌni**

iv) When a weak-form word is being "cited" or "quoted"; for example:

'You shouldn't put "and" at the end of a sentence'
ju 'ʃʊdn̩t pʊt 'ænd ət ði 'end əv ə 'sentəns

Another point to remember is that when weak-form words whose spelling begins with 'h' (e.g. 'her', 'have') occur at the beginning of a sentence, the pronunciation is with initial **h**, even though this is usually omitted in other contexts.

AU12, Exs 1–4

In the rest of this chapter, the most common weak-form words will be introduced.

1 'the'
Weak forms: **ðə** (before consonants)
'Shut the door' **'ʃʌt ðə 'dɔː**
ði (before vowels)
'Wait for the end' **'weɪt fə ði 'end**

2 'a', 'an'
Weak forms: **ə** (before consonants)
'Read a book' **'riɪd ə 'bʊk**
ən (before vowels)
'Eat an apple' **'iɪt ən 'æpl̩**

3 'and'
Weak form: **ən** (sometimes **n̩ after t, d, s, z, ʃ**)
'Come and see' **'kʌm ən 'siː**
'Fish and chips' **'fɪʃ n̩ 'tʃɪps**

4 'but'
Weak form: **bət**
'It's good but expensive' **ɪts 'gʊd bət ɪk'spensɪv**

5 'that'
This word only has a weak form when used in a relative clause; when used with a demonstrative sense it is always pronounced in its strong form.
Weak form: **ðət**
'The price is the thing that annoys me' **ðə 'praɪs ɪz ðə 'θɪŋ ðət ə'nɔɪz mi**

6 'than'
Weak form: **ðən**
'Better than ever' **'betə ðən 'evə**

7 'his' (when it occurs before a noun)
Weak form: **ɪz** (**hɪz** at the beginning of a sentence)
'Take his name' **'teɪk ɪz 'neɪm**
(Another sense of 'his', as in 'it was his', or 'his was late', always has the strong form.)

8 'her'
When used with possessive sense, preceding a noun; as an object pronoun, this can also occur at the end of a sentence.
Weak forms: **ə** (before consonants)
'Take her home' **'teɪk ə 'həʊm**
ər (before vowels)
'Take her out' **'teɪk ər 'aʊt**

9 'your'
Weak forms: **jə** (before consonants)
'Take your time' **'teɪk jə 'taɪm**
jər (before vowels)
'On your own' **'ɒn jər 'əʊn**

10 'she', 'he', 'we', 'you'
This group of pronouns has weak forms pronounced with weaker vowels than the **iː** and **uː** of their strong forms. I use the symbols **i** and **u** (in preference to **ɪ** and **ʊ**) to represent them. There is little difference in the pronunciation in different places in the sentence, except in the case of 'he'.
Weak forms:

a) 'she' **ʃi**
'Why did she read it?' **'waɪ dɪd ʃi 'riːd ɪt**
'Who *is* she?' **'huː 'ɪz ʃi**

b) 'he' **i** (the weak form is usually pronounced without **h** except at the beginning of a sentence)
'Which did he choose?' **'wɪtʃ dɪd i 'tʃuːz**
'He was late, wasn't he?' **hi wəz 'leɪt 'wɒzn̩t i**

c) 'we' **wi**
'How can we get there?' **'haʊ kən wi 'get ðeə**
'We need that, don't we?' **wi 'niːd ðæt 'dəʊnt wi**

d) 'you' **ju**
'What do you think?' **'wɒt də ju 'θɪŋk**
'You like it, do you?' **ju 'laɪk ɪt 'duː ju**

11 'him'
Weak form: **ɪm**
'Leave him alone' **'liːv ɪm ə'ləʊn**
'I've seen him' **aɪv 'siːn ɪm**

12 'her'
Weak form: **ə** (**hə** when sentence-initial)
'Ask her to come' **'ɑːsk ə tə 'kʌm**
'I've met her' **aɪv 'met ə**

13 'them'
Weak form: **ðəm**
'Leave them here' **'liːv ðəm 'hɪə**
'Eat them' **'iːt ðəm**

14 'us'
Weak form: **əs**
'Write us a letter' **'raɪt əs ə 'letə**
'They invited all of us' **ðeɪ ɪn'vaɪtɪd 'ɔːl əv əs**

The next group of words (some prepositions and other function words) occur in their strong forms when they are final in a sentence; examples of this are given. (Example 19 is a partial exception.)

15 'at'
Weak form: **ət**
'I'll see you at lunch' **aɪl 'siː juː ət 'lʌnʃ**
In final position: **æt**
'What's he shooting at?' **'wɒts i 'ʃuːtɪŋ æt**

16 'for'
Weak form: **fə** (before consonants)
'Tea for two' **'tiː fə 'tuː**
fər (before vowels)
'Thanks for asking' **'θæŋks fər 'ɑːskɪŋ**
In final position: **fɔː**
'What's that for?' **'wɒts 'ðæt fɔː**

17. 'from'
Weak form: **frəm**
'I'm home from work' **aɪm 'həʊm frəm 'wɜːk**
In final position: **frɒm**
'Here's where it came from' **'hɪəz weər ɪt 'keɪm frɒm**

18 'of'

Weak form: **əv**

'Most of all' **'məʊst əv 'ɔɪl**

In final position: **ɒv**

'Someone I've heard of' **'sʌmwʌn aɪv 'hɜɪd ɒv**

19 'to'

Weak forms: **tə** (before consonants)

'Try to stop' **'traɪ tə 'stɒp**

tu (before vowels)

'Time to eat' **'taɪm tu 'iːt**

In final position: **tu** (It is not usual to use the strong form **tuː**, and the pre-consonantal weak form **tə** is never used.)

'I don't want to' **aɪ 'dəʊnt 'wɒnt tu**

20 'as'

Weak form: **əz**

'As much as possible' **əz 'mʌtʃ əz 'pɒsəbl̩**

In final position: **æz**

'That's what it was sold as' **'ðæts 'wɒt ɪt wəz 'səʊld æz**

21 'some'

This word is used in two different ways. In one sense (typically, when it occurs before a countable noun, meaning "an unknown individual") it has the strong form:

'I think some animal broke it' **aɪ 'θɪŋk sʌm 'ænɪml̩ 'brəʊk ɪt**

It is also used before uncountable nouns (meaning "an unspecified amount of") and before other nouns in the plural (meaning "an unspecified number of"); in such uses it has the weak form **səm**:

'Have some more tea' **'hæv səm 'mɔː 'tiː**

In final position: **sʌm**

'I've got some' **aɪv 'gɒt sʌm**

22 'there'

When this word has a demonstrative function, it always occurs in its strong form **ðeə** (**ðeər** before vowels), e.g:

'There it is' **'ðeər ɪt 'ɪz**

'Put it there' **'pʊt ɪt 'ðeə**

Weak forms: **ðə** (before consonants)
'There should be a rule' **ðə 'ʃʊd bi ə 'ruːl**
ðər (before vowels)
'There is' **ðər 'ɪz**
In final position: the pronunciation may be **ðə** or **ðeə**.
'There isn't any, is there?' **ðər 'ɪzn̩t eni 'ɪz ðə**
or **ðər 'ɪzn̩t eni 'ɪz ðeə**

The remaining weak-form words are all auxiliary verbs, which are always used in conjunction with (or at least implying) another ("full") verb. It is important to remember that in their negative form (i.e. combined with 'not') they never have the weak pronunciation, and some (e.g. 'don't', 'can't') have different vowels from their non-negative strong forms.

23 'can', 'could'
Weak forms: **kən, kəd**
'They can wait' **'ðeɪ kən 'weɪt**
'He could do it' **'hiː kəd 'duː ɪt**
In final position: **kæn, kʊd**
'I think we can' **aɪ 'θɪŋk wi 'kæn**
'Most of them could' **'məʊst əv ðəm kʊd**

24 'have', 'has', 'had'
Weak forms: **əv, əz, əd** (with initial **h** in initial position)
'Which have you seen?' **'wɪtʃ əv ju 'siːn**
'Which has been best?' **'wɪtʃ əz biːn 'best**
'Most had gone home' **'məʊst əd 'gɒn 'həʊm**
In final position: **hæv, hæz, hæd**
'Yes, we have' **'jes wi 'hæv**
'I think she has' **aɪ 'θɪŋk ʃi 'hæz**
'I thought we had' **aɪ 'θɔːt wi 'hæd**

25 'shall', 'should'
Weak forms: **ʃəl** or **ʃl̩**; **ʃəd**
'We shall need to hurry' **wi ʃl̩ 'niːd tə 'hʌri**
'I should forget it' **'aɪ ʃəd fə'get ɪt**
In final position: **ʃæl, ʃʊd**
'I think we shall' **aɪ 'θɪŋk wi 'ʃæl**
'So you should' **'səʊ ju 'ʃʊd**

26 'must'

This word is sometimes used with the sense of forming a conclusion or deduction, e.g. 'she left at 8 o'clock, so she must have arrived by now'; when 'must' is used in this way, it is rather less likely to occur in its weak form than when it is being used in its more familiar sense of "obligation".

Weak forms: **məs** (before consonants)
'You must try harder' **ju məs 'traɪ 'hɑɪdə**
məst (before vowels)
'He must eat more' **hi məst 'iɪt 'mɔɪ**

In final position: **mʌst**
'She certainly must' **ʃi 'sɜɪtn̩li 'mʌst**

27 'do', 'does'

Weak forms:

'do' **də** (before consonants)
'Why do they like it?' **'waɪ də ðeɪ 'laɪk ɪt**
du (before vowels)
'Why do all the cars stop?' **'waɪ du ɔɪl ðə 'kɑɪz 'stɒp**

'does' **dəz**
'When does it arrive?' **'wen dəz ɪt ə'raɪv**

In final position: **duɪ, dʌz**
'We don't smoke, but some people do' **'wiɪ dəunt 'smәuk bət 'sʌm 'piɪpl̩ 'duɪ**
'I think John does' **aɪ 'θɪŋk 'dʒɒn dʌz**

28 'am', 'are', 'was', 'were'

Weak forms: **əm**
'Why am I here?' **'waɪ əm aɪ 'hɪə**
ə (before consonants)
'Here are the plates' **'hɪər ə ðə 'pleɪts**
ər (before vowels)
'The coats are in there' **ðə 'kəuts ər ɪn 'ðeə**
wəz
'He was here a minute ago' **hi wəz hɪər ə 'mɪnɪt ə'gəu**
wə (before consonants)
'The papers were late' **ðə 'peɪpəz wə 'leɪt**
wər (before vowels)

'The questions were easy' **ðə 'kwestʃənz wər iːzi**

In final position: **æm, ɑː, wɒz, wɜː**

'She's not as old as I am' **ʃiz 'nɒt əz 'əʊld əz 'aɪ æm**

'I know the Smiths are' **aɪ 'nəʊ ðə 'smɪθs ɑː**

'The last record was' **ðə 'lɑɪst 'rekɔɪd wɒz**

'They weren't as cold as we were' **ðeɪ 'wɜɪnt əz 'kəʊld əz 'wiː wɜː**

Notes on problems and further reading

This chapter is almost entirely practical. All books about English pronunciation devote a lot of attention to these words. Some of them give a great deal of importance to using weak forms, but do not stress the importance of also knowing when to use the strong forms, something which I feel is very important; see Mortimer, 1984.

Written exercises

In the following sentences, the transcription for the weak-form words is left blank. Fill in the blanks, taking care to use the appropriate form (weak or strong).

1 I want her to park that car over there.
aɪ wɒnt pɑɪk kɑːr əʊvə

2 Of all the proposals, the one that you made is the silliest.
ɔɪl prəpəʊzl̩z wʌn meɪd ɪz sɪliəst

3 Jane and Bill could have driven them to and from the party.
dʒem bɪl drɪvn̩ pɑɪti

4 To come to the point, what shall we do for the rest of the week?
kʌm pɔɪnt wɒt rest wiːk

5 Has anyone got an idea where it came from?
enɪwʌn gɒt aɪdɪə weər ɪt keɪm

6 Pedestrians must always use the crossings provided.
pədestriənz ɔːlwɪz juːz krɒsɪŋz prəvaɪdɪd

7 Each one was a perfect example of the art that had been developed there.
iːtʃ wʌn pɜɪfɪkt ɪgzɑɪmpl̩ ɑɪt biːn drveləpt

13 Problems in phonemic analysis

The concept of the phoneme was introduced in Chapter 5, and a few theoretical problems connected with phonemic analysis have been mentioned in other chapters. The general assumption (as in most phonetics books) has been that speech is composed of phonemes and that usually whenever a speech sound is produced by a speaker it is possible to identify which phoneme that sound belongs to. While this is often true, we must recognise that there are exceptions which make us consider some quite serious theoretical problems. From the comparatively simple point of view of learning pronunciation, these problems are not particularly important. However, from the point of view of learning about the phonology of English they are too important to ignore.

There are problems of different types. In some cases, we have difficulty in deciding on the overall phonemic system of the accent we are studying, while in others we are concerned about how a particular sound fits into this system. A number of such problems are discussed below.

13.1 Affricates

The affricates **tʃ** and **dʒ** are, phonetically, composed of a plosive followed by a fricative, as explained in Chapter 6. It is possible to treat each of the pair **tʃ, dʒ** as a single consonant phoneme; we will call this the **one-phoneme analysis** of **tʃ, dʒ**. It is also possible to say that they are composed of two phonemes each – **t** plus **ʃ** and **d** plus **ʒ** respectively – all of which are already established as independent phonemes of English; this will be called the **two-phoneme analysis** of **tʃ** and **dʒ**. If we adopted the two-phoneme analysis, the words 'church' and 'judge' would be composed of five phonemes each, like this:

t – ʃ – ɜɪ – t – ʃ d – ʒ – ʌ – d – ʒ

instead of the three phonemes that result from the one-phoneme analysis:

tʃ – ɜɪ – tʃ dʒ – ʌ – dʒ

and there would be no separate **tʃ** and **dʒ** phonemes. But how can we decide which analysis is preferable? The two-phoneme analysis has one main advantage: if there are no separate **tʃ** and **dʒ** phonemes, then our total set of English consonants is smaller. Many phonologists have claimed that one should prefer the analysis which is the most "economical" in the number of phonemes it results in. The argument for this might be based on the claim that when we speak to someone we are using a "code", and the most efficient codes do not use unnecessary symbols. Further, it can be claimed that a phonological analysis is a type of scientific theory, and a scientific theory should be stated as economically as possible. However, it is the one-phoneme analysis that is generally chosen by phonologists. Why is this? There are several arguments; no single one of them is conclusive, but added together they are felt to make the one-phoneme analysis seem preferable. We will look briefly at some of these arguments.

i) One argument could be called "phonetic" or "allophonic": if it could be shown that the phonetic quality of the **t** and **ʃ** (or **d** and **ʒ**) in **tʃ**, **dʒ** is clearly different from realisations of **t**, **ʃ**, **d**, **ʒ** found elsewhere in similar contexts, this would support the analysis of **tʃ**, **dʒ** as separate phonemes. As an example, it might be claimed that **ʃ** in 'hutch' **hʌtʃ** was different (perhaps in having shorter duration) from **ʃ** in 'hush' or 'Welsh' **hʌʃ**, **welʃ**; or it might be claimed that the place of articulation of **t** in 'watch apes' **wɒtʃ eɪps** was different from that of **t** in 'what shapes' **wɒt ʃeɪps**. This argument is a weak one: there is no clear evidence that such phonetic differences exist, and even if there were such evidence, it would be easy to produce explanations for the differences that did not depend on phonemic analyses (e.g. the position of the word boundary in 'watch apes', 'what shapes').
ii) It could be argued that the proposed phonemes **tʃ** and **dʒ** (if one were arguing for the one-phoneme analysis) have distributions similar to other consonants, while other combinations of plosive

plus fricative do not. It can easily be shown that **tʃ** and **dʒ** are found initially, medially and finally, and that no other combination (e.g. **pf**, **dz**, **tθ**) has such a wide distribution. However, several consonants *are* generally accepted as phonemes of the BBC accent despite *not* being free to occur in all positions (e.g. **r**, **w**, **j**, **h**, **ŋ**, **ʒ**), so this argument, although supporting the one-phoneme analysis, does not actually *prove* that **tʃ**, **dʒ** must be classed with other single consonant phonemes.

iii) If **tʃ** and **dʒ** were able to combine quite freely with other consonants to form consonant clusters, this would support the one-phoneme analysis. In initial position, however, **tʃ** and **dʒ** never occur in clusters with other consonants. In final position in the syllable, we find that **tʃ** can be followed by **t** (e.g. 'watched' **wɒtʃt**) and **dʒ** by **d** (e.g. 'wedged' **wedʒd**). Final **tʃ** and **dʒ** can be preceded by **l** (e.g. 'squelch' **skweltʃ**, 'bulge' **bʌldʒ**); **ʒ** is never preceded by **l**, and **ʃ** is preceded by **l** only in a few words and names, e.g. 'Welsh', 'Walsh' **welʃ**, **wɒlʃ**. A fairly similar situation is found if we ask if **n** can precede **tʃ** and **dʒ**; some BBC speakers have **ntʃ** in 'lunch', 'French', etc., and never pronounce the sequence **nʃ** within a syllable, while other speakers always have **nʃ** in these contexts and never **ntʃ**. It seems, then, that no contrast between syllable-final **lʃ** and **ltʃ** exists in the BBC accent, and the same appears to be true in relation to **nʃ** and **ntʃ** and to **nʒ** and **ndʒ**. There are no other possibilities for final consonant clusters containing **tʃ** and **dʒ**, except that the pre-final **l** or **n** may occur in combination with post-final **t**, **d** as in 'squelched' **skweltʃt**, 'hinged' **hɪndʒd**. It could not, then, be said that **tʃ** and **dʒ** combine freely with other consonants in forming consonant clusters; this is particularly noticeable in initial position.

How would the two-phoneme analysis affect the syllable-structure framework that was introduced in Chapter 8? Initial **tʃ**, **dʒ** would have to be interpreted as initial **t**, **d** plus post-initial **ʃ**, **ʒ**, with the result that the post-initial set of consonants would have to contain **l**, **r**, **w**, **j** and also **ʃ**, **ʒ** – consonants which are rather different from the other four and which could only combine with **t**, **d**. (The only alternative would be to put **t**, **d** with **s** in the pre-initial category, again with very limited possibilities of combining with another consonant.)

iv) Finally, it has been suggested that if native speakers of English who have not been taught phonetics feel that **tʃ** and **dʒ** are each "one sound", we should be guided by their intuitions and prefer the one-phoneme analysis. The problem with this is that discovering what untrained (or "naive") speakers feel about their own language is not as easy as it might sound. It would be necessary to ask questions like this: "Would you say that the word 'chip' begins with one sound – like 'tip' and 'sip' – or with two sounds – like 'trip' and 'skip'?" But the results would be distorted by the fact that two consonant letters are used in the spelling; to do the test properly one should use illiterate subjects, which raises many further problems.

This rather long discussion of the phonemic status of **tʃ** and **dʒ** shows how difficult it can be to reach a conclusion in phonemic analysis.

For the rest of this chapter a number of other phonological problems will be discussed comparatively briefly. I have already mentioned (in Chapter 6) problems of analysis in connection with the sounds usually transcribed **hw** and **hj**. The velar nasal **ŋ**, described in Chapter 7, also raises a lot of analysis problems; many writers have suggested that the correct analysis is one in which there is no **ŋ** phoneme, and this sound is treated as an allophone of the phoneme **n** that occurs when it precedes the phoneme **g**. It was explained in Chapter 7 that in certain contexts no **g** is pronounced, but it can be claimed that at an abstract level there *is* a **g** phoneme, though in certain contexts the **g** is not actually pronounced. The sound **ŋ** is therefore, according to this theory, an allophone of **n**.

13.2 The English vowel system

The analysis of the English vowel system presented in Chapters 2 and 3 contains a large number of phonemes, and it is not surprising that some phonologists (who believe in the importance of keeping the total number of phonemes small) propose different analyses which contain less than ten vowel phonemes and treat all long vowels and diphthongs as composed of two phonemes each. There are different ways of doing this: one way is to treat long vowels and diphthongs as

composed of two vowel phonemes. Starting with a set of basic or "simple" vowel phonemes ɪ, e, æ, ʌ, ɒ, ʊ, ə it is possible to make up long vowels by using short vowels twice. Our usual transcription is given in brackets:

ɪɪ (iː) ææ (ɑː) ɒɒ (ɔː) ʊʊ (uː) əə (ɜː)

This can be made to look less unusual by choosing different symbols for the basic vowels. In this approach, diphthongs are made from a simple vowel phoneme followed by one of ɪ, ʊ, ə, and triphthongs are made from a basic vowel plus one of ɪ, ʊ followed by ə, and are therefore composed of three phonemes.

Another way of doing this kind of analysis is to treat long vowels and diphthongs as composed of a vowel plus a consonant; this may seem a less obvious way of proceeding, but it was for many years the choice of most American phonologists. The idea is that long vowels and diphthongs are composed of a basic vowel phoneme followed by one of **j**, **w**, **h** (we should add **r** for rhotic accents). Thus the diphthongs could be made up like this (our usual transcription is given in brackets):

ej (eɪ)	**əw** (əʊ)	**ɪh** (ɪə)
æj (aɪ)	**æw** (aʊ)	**eh** (eə)
ɒj (ɔɪ)		**ʊh** (ʊə)

Long vowels:

ɪj (iː) **æh** (ɑː) **ɒh** (ɔː) **əh** (ɜː) **ʊw** (uː)

Diphthongs and long vowels are now of exactly the same phonological composition. An important point about this analysis is that **j**, **w**, **h** do not otherwise occur finally in the syllable. In this analysis, the inequality of distribution is corrected.

In Chapter 9 we saw how, although ɪ and **iː** are clearly distinct in most contexts, there are other contexts where we find a sound which cannot clearly be said to belong to one or other of these two phonemes. The suggested solution to this problem was to use the symbol **i**, which does not represent any single phoneme; a similar proposal was made for **u**. We use the term **neutralisation** for cases where contrasts between phonemes which exist in other places in the language disappear in particular contexts. There are many other ways of analysing the very complex vowel system of English, some of

which are extremely ingenious. Each has its own advantages and disadvantages.

13.3 Syllabic consonants

A final analysis problem that we will consider is that mentioned at the end of Chapter 8: how to deal with syllabic consonants. It has to be recognised that syllabic consonants are a problem: they *are* phonologically different from their non-syllabic counterparts. How do we account for the following minimal pairs, which were given in Chapter 9?

Syllabic	*Non-syllabic*
'coddling' **kɒd̩lɪŋ**	'codling' **kɒdlɪŋ**
'Hungary' **hʌŋg̩ri**	'hungry' **hʌŋgri**

One possibility is to add new consonant phonemes to our list. We could invent the phonemes **l̩**, **r̩**, **n̩**, etc. The distribution of these consonants would be rather limited, but the main problem would be fitting them into the pattern of syllable structure. For a word like 'button' **bʌtn̩** or 'bottle' **bɒtl̩**, it would be necessary to add **n̩** and **l̩** to the first post-final set; the argument would be extended to include the **r̩** in 'Hungary'. But if these consonants now form part of a syllable-final consonant cluster, how do we account for the fact that English speakers hear the consonants as extra syllables? The question might be answered by saying that the new phonemes are to be classed as vowels. Another possibility is to set up a phoneme that we might name **syllabicity**, symbolised with the mark ̩. Then the word 'codling' would consist of the following six phonemes: **k · ɒ · d · l · ɪ · ŋ**, while the word 'coddling' would consist of the following *seven* phonemes: **k · ɒ · d · (l** and simultaneously ̩**) · ɪ · ŋ**. This is superficially an attractive theory, but the proposed phoneme is nothing like the other phonemes we have identified up to this point – putting it simply, it doesn't have any sound.

Some phonologists maintain that a syllabic consonant is really a case of a vowel and a consonant that have become combined. Let us suppose that the vowel is **ə**. We could then say that, for example, 'Hungary' is phonemically **hʌŋgəri** while 'hungry' is **hʌŋgri**; it would then be necessary to say that the vowel phoneme in the phonemic

representation is not pronounced as a vowel, but instead causes the following consonant to become syllabic. This is an example of the abstract view of phonology where the way a word is represented phonologically may be significantly different from the actual sequence of sounds heard, so that the phonetic and the phonemic levels are quite widely separated.

13.4 Clusters of s plus plosives

Words like 'spill', 'still', 'skill' are usually represented with the phonemes **p**, **t**, **k** following the **s**. But, as many writers have pointed out, it would be quite reasonable to transcribe them with **b**, **d**, **g** instead. For example, **b**, **d**, **g** are unaspirated while **p**, **t**, **k** in syllable-initial position are usually aspirated. However, in **sp**, **st**, **sk** we find an unaspirated plosive, and there could be a strong argument for transcribing them as **sb**, **sd**, **sg**. We do not do this, perhaps because of the spelling, but it is important to remember that the contrasts between **p** and **b**, between **t** and **d** and between **k** and **g** are neutralised in this context.

13.5 Schwa (ə)

It has been suggested that there is not really a contrast between ə and ʌ, since ə only occurs in weak syllables and no minimal pairs can be found to show a clear contrast between ʌ and ə in unstressed syllables (although there have been some ingenious attempts). This has resulted in a proposal that one phoneme symbol (e.g. ə) be used for representing any occurrence of ə or ʌ, so that 'cup' would be transcribed **'kəp** and 'upper' as **'əpə**. This new ə phoneme would have two allophones, one being [ə] and the other [ʌ]; the stress mark would indicate the [ʌ] allophone and in syllables not marked for stress it would be more likely that [ə] would be pronounced.

Other phonologists have suggested that ə is an allophone of several other vowels; for example, compare the middle two syllables in the words 'economy' **ɪ'kɒnəmi** and 'economic' **ˌiːkə'nɒmɪk** – it appears that when the stress moves away from the syllable containing ɒ the vowel becomes ə. Similarly, compare 'Germanic' **dʒɜː'mænɪk** with 'German' **'dʒɜːmən** – when the stress is taken

away from the syllable **mæn**, the vowel weakens to ə. Many similar examples could be constructed with other vowels; some possibilities may be suggested by the list of words given in Section 9.2 to show the different spellings that can be pronounced with ə. The conclusion that could be drawn from this argument is that ə is not a phoneme of English, but is an allophone of several different vowel phonemes when those phonemes occur in an unstressed syllable. The argument is in some ways quite an attractive one, but since it leads to a rather complex and abstract phonemic analysis it is not adopted for this course.

13.6 Distinctive features

Many references have been made to phonology in this course, with the purpose of making use of the concepts and analytical techniques of that subject to help explain various facts about English pronunciation as efficiently as possible. One might call this "applied phonology"; however, the phonological analysis of different languages raises a great number of difficult and interesting theoretical problems, and for a long time the study of phonology "for its own sake" has been regarded as an important area of theoretical linguistics. Within this area of what could be called "pure phonology", problems are examined with little or no reference to their relevance to the language learner. Many different theoretical approaches have been developed, and no area of phonology has been free from critical examination. The very fundamental notion of the phoneme, for example, has been treated in many different ways. One approach that has been given a lot of importance is **distinctive feature analysis**, which is based on the principle that phonemes should be regarded not as independent and indivisible units, but instead as combinations of different features. For example, if we consider the English **d** phoneme, it is easy to show that it differs from the plosives **b** and **g** in its place of articulation (alveolar), from **t** in being lenis, from **s** and **z** in not being fricative, from **n** in not being nasal, and so on. If we look at each of the consonants just mentioned and see which of the features each one has, we get a table like this, where + means that a phoneme does possess that feature and – means that it does not:

	d	b	g	t	s	z	n
alveolar	+	–	–	+	+	+	+
bilabial	–	+	–	–	–	–	–
velar	–	–	+	–	–	–	–
lenis	+	+	+	–	–	+	(+)*
plosive	+	+	+	+	–	–	–
fricative	–	–	–	–	+	+	–
nasal	–	–	–	–	–	–	+

* Since there is no fortis/lenis contrast among nasals this could be left blank

If you look carefully at this table, you will see that the combination of + and – values for each phoneme is different; if two sounds were represented by exactly the same +'s and –'s, then by definition they could not be different phonemes. In the case of the limited set of phonemes used for this example, not all the features are needed: if one wished, it would be possible to dispense with, for example, the feature "velar" and the feature "nasal". The **g** phoneme would still be distinguished from **b** and **d** by being neither alveolar nor bilabial, and **n** would be distinct from plosives and fricatives simply by being neither plosive nor fricative. To produce a complete analysis of all the phonemes of English, other features would be needed for representing other types of consonant, and for vowels and diphthongs. In distinctive feature analysis the features themselves thus become important components of the phonology.

It has been claimed by some writers that distinctive feature analysis is relevant to the study of language learning, and that pronunciation difficulties experienced by learners are better seen as due to the need to learn a particular feature or combination of features than as the absence of particular phonemes. For example, English speakers learning French or German have to learn to produce front rounded vowels. In English it is not necessary to be able to consider vowels which are [+front, +round], whereas this is necessary for French and German; it could be said that the major task for the English-speaking learner of French or German in this case is to learn the combination

of these features, not to learn the individual vowels y, ø and (in French) œ.*

English, on the other hand, has to be able to distinguish dental from labiodental and alveolar places of articulation, for θ to be distinct from f and s and for ð to be distinct from v and z. This requires an additional feature that most languages do not make use of, and learning this could be seen as a specific task for the learner of English. Distinctive feature phonologists have also claimed that when children are learning their first language, they acquire features rather than individual phonemes.

13.7 Conclusion

This chapter is intended to show that there are many ways of analysing the English phonemic system, each with its own advantages and disadvantages. We need to consider the practical goal of teaching or learning about English pronunciation, and for this purpose a very abstract analysis would be unsuitable. This is one criterion for judging the value of an analysis; unless one believes in carrying out phonological analysis for purely aesthetic reasons, the only other important criterion is whether the analysis is likely to correspond to the representation of sounds in the human brain. We do not yet know much about this, but the brain is so powerful and complex that it is very unlikely that any of the analyses proposed so far bear much resemblance to this reality; they are too heavily influenced by the theoretician's preoccupation with economy, elegance and simplicity.

Notes on problems and further reading

The analysis of tʃ and dʒ is one of the most intractable problems. The general principles in traditional phoneme theory have been set out by Trubetzkoy (1939); the German original has been translated into English (1969), and the relevant section of this translation is reprinted in Fudge (1973: 65–70). The problem is also discussed in Cruttenden (1994, 157–60). The phonemic analysis of the velar nasal has already been discussed above (see Notes in Chapter 7). The 'double vowel' interpretation of English long vowels was put forward

* The phonetic symbols represent the following sounds: y is a close front rounded vowel (e.g. the vowel in French *tu*, German *Bühne*); ø is a close-mid front rounded vowel (e.g. French *peu*, German *schön*); œ is an open-mid front rounded vowel (e.g. French *oeuf*).

by MacCarthy (1952) and is used by Kreidler (1989). The 'vowel-plus-semivowel' interpretation of long vowels and diphthongs was almost universally accepted by American (and some British) writers from the 1940s to the 1960s, and still pervades contemporary descriptions. It has the advantage of being economical on phonemes and very 'neat and tidy'. The analysis in this form is presented in Trager and Smith (1951). This work was claimed to provide an analysis that could produce a phonologically distinct representation for all English short vowels, long vowels and diphthongs in all accents. An early attack on this view was made by Sledd (1955: 316–24). In generative phonology the claim is that, at the abstract level, English vowels are simply tense or lax. If they are lax they are realised as short vowels, if tense as diphthongs (this category including what I have been calling long vowels). The quality of the first element of the diphthongs/long vowels is modified by some phonological rules, while other rules supply the second element automatically. This is set out in Chomsky and Halle (1968: 178–87). There is a valuable discussion of the interpretation of the English vowel system with reference to several different accents in Giegerich (1992: Chapter 3), followed by an explanation of the distinctive feature analysis of the English vowel system (Chapter 4) and the consonant system (Chapter 5). A more wide-ranging discussion of distinctive features is given in Clark and Yallop (1995: Chapter 10).

On the interpretation of **sp**, **st**, **sk**, see Davidsen-Nielsen (1969). There is an interesting discussion of the ə – ʌ contrast in Wells (1970: 233–5). The idea that ə is an allophone of many English vowels is not a new one. In generative phonology, ə results from vowel reduction in vowels which have never received stress in the process of the application of stress rules. This is explained – in rather difficult terms – in Chomsky and Halle (1968: 110–26). A clearer treatment of the schwa problem is in Giegerich (1992: 68–9 and 285–7).

Note for teachers

Since this is a theoretical chapter it is difficult to come up with practical work. I do not feel that it is helpful for students to do exercises on using different ways of transcribing phonemes – just learning one set of conventions is difficult enough. Some books on phonology give exercises on the phonemic analysis of other languages

(e.g. Katamba, 1989; Roca and Johnson, 1999), but although these are useful, I do not feel that it would be appropriate in this book to divert attention from English. The exercises given below therefore concentrate on bits of phonetically transcribed English which involve problems when a phonemic representation is required.

Written exercises

In this exercise you must look at phonetically transcribed material from different English accents and decide on the best way to interpret and transcribe them phonemically. Information is given where necessary about the meaning of the phonetic symbols.

1 a) 'thing' **[θɪŋg]**
b) 'think' **[θɪŋk]**
c) 'thinking' **[θɪŋkɪŋg]**
d) 'finger' **[fɪŋgə]**
e) 'singer' **[sɪŋgə]**
f) 'singing' **[sɪŋgɪŋg]**

2 It often happens in rapid speech that a nasal consonant disappears when it comes between a vowel and another consonant (for example, this may happen to the 'n' in 'front': when this happens the preceding vowel becomes **nasalised** (some of the air escapes through the nose). We symbolise a nasalised vowel in phonetic transcription by putting the ~ diacritic above it; for example, the word 'front' may be pronounced **[frʌ̃t]**. Nasalised vowels are found in the words given in phonetic transcription below. Transcribe them phonemically.

a) **[sãʊ̃d]**
b) **[æ̃gə]**
c) **[kɑ̃ːt]**
d) **[kæ̃pə]**
e) **[bɒ̃d]**

3 When the t phoneme occurs between vowels it is sometimes pronounced as a "tap" (the tongue blade strikes the alveolar ridge sharply, producing a very brief voiced plosive: the phonetic symbol is ɾ); this is very common in American English, and is also found in a number of accents in England: think of a typical

American pronunciation of "getting better" [**gerɪŋ berə**]. Look at the transcriptions of a number of words given below and see if you can work out (for the accent in question) the environment in which [**ɾ**] is found.

a) 'betting' [**berɪŋ**]
b) 'bedding' [**bedɪŋ**]
c) 'attend' [**ətʰend**]
d) 'attitude' [**ærɪtuːd**]
e) 'time' [**tʰaɪm**]
f) 'tight' [**tʰaɪt**]

4 Distinctive feature analysis looks at different properties of segments and classes of segments. In the following exercise you must mark the value of each feature in the table for each segment listed on the top row with either a '+' or '–'; you will probably find it useful to look at the IPA chart on p. xi.

	p	**d**	**s**	**m**	**z**
Continuant					
Alveolar					
Voiced					

5 In all the following sets of segments (*a–f*), all segments in the set possess some characteristic feature which they have in common and which may distinguish them from other segments. Can you identify what this common feature might be for each set?

a) English **iː, ɪ, uː, ʊ**; Cardinal vowels [**i**], [**e**], [**u**], [**o**]
b) **t d n l s tʃ dʒ ʃ ʒ r**
c) **b f v k g h**
d) **p t k f θ s ʃ tʃ**
e) **uː ɔɪ əʊ aʊ**
f) **l r w j**

14 Aspects of connected speech

Many years ago scientists tried to develop machines that produced speech from a vocabulary of pre-recorded words; the machines were designed to join these words together to form sentences. For very limited messages, such as those of a "talking clock", this technique was usable, but for other purposes the quality of the speech was so unnatural that it was practically unintelligible. In recent years, developments in computer technology have led to big improvements in this way of producing speech, but the inadequacy of the original "mechanical speech" approach has many lessons to teach us about pronunciation teaching and learning. In looking at connected speech it is useful to bear in mind the difference between the way humans speak and what would be found in "mechanical speech".

14.1 Rhythm

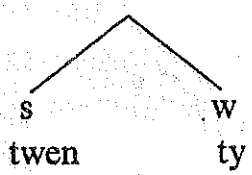
The notion of **rhythm** involves some noticeable event happening at regular intervals of time; one can detect the rhythm of a heart-beat, of a flashing light or of a piece of music. It has often been claimed that English speech is rhythmical, and that the rhythm is detectable in the regular occurrence of stressed syllables; of course, it is not suggested that the timing is as regular as a clock: the regularity of occurrence is only relative. The theory that English has **stress-timed rhythm** implies that stressed syllables will tend to occur at relatively regular intervals whether they are separated by unstressed syllables or not; this would not be the case in "mechanical speech". An example is given below. In this sentence, the stressed syllables are given numbers: syllables 1 and 2 are not separated by any unstressed syllables, 2 and 3 are separated by one unstressed syllable, 3 and 4 by two and 4 and 5 by three.

1	2		3		4			5
'Walk	'down	the	'path	to	the	'end	of	the	ca 'nal

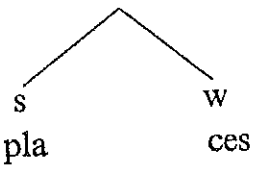
The stress-timed rhythm theory states that the times from each stressed syllable to the next will tend to be the same, irrespective of the number of intervening unstressed syllables. The theory also claims that while some languages (e.g. Russian and Arabic) have stress-timed rhythm similar to that of English, others (such as French, Telugu and Yoruba) have a different rhythmical structure called **syllable-timed rhythm**; in these languages, all syllables, whether stressed or unstressed, tend to occur at regular time-intervals and the time between stressed syllables will be shorter or longer in proportion to the number of unstressed syllables. Some writers have developed theories of English rhythm in which a unit of rhythm, the **foot**, is used (with an obvious parallel in the metrical analysis of verse); the foot begins with a stressed syllable and includes all following unstressed syllables up to (but not including) the following stressed syllable. The example sentence given above would be divided into feet as follows:

1	2	3	4	5
'Walk	'down the	'path to the	'end of the ca	'nal

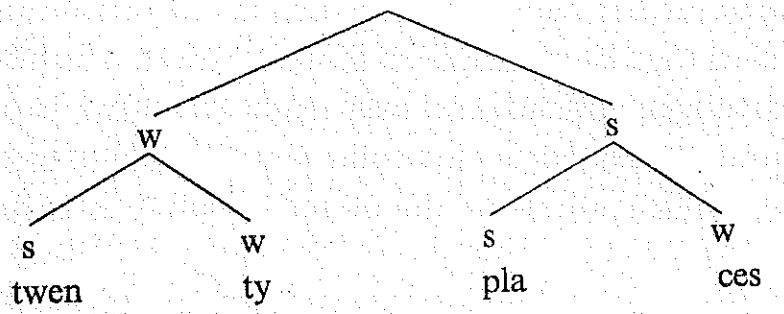
Some theories of rhythm go further than this, and point to the fact that some feet are stronger than others, producing strong–weak patterns in larger pieces of speech above the level of the foot. To understand how this could be done, let's start with a simple example: the word 'twenty' has one strong and one weak syllable, forming one foot. A diagram of its rhythmical structure can be made, where s stands for "strong" and w stands for "weak".



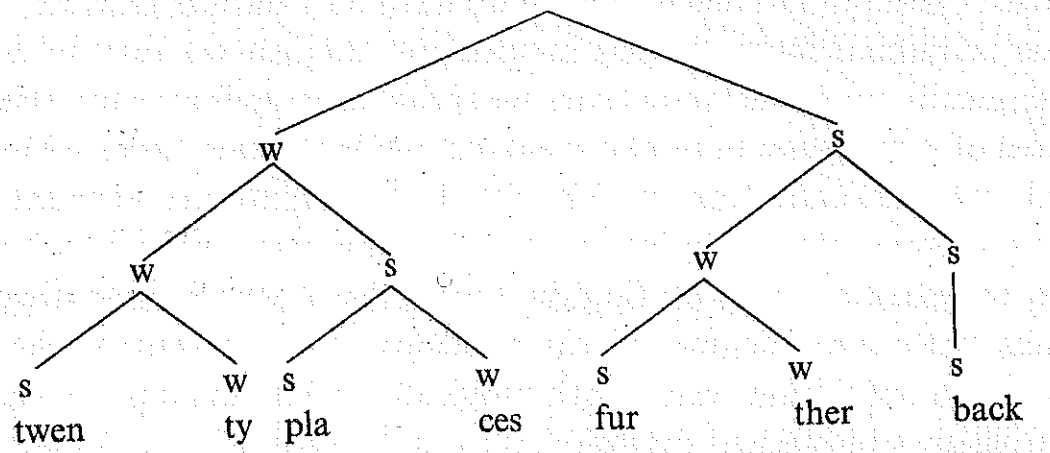
The word 'places' has the same form:



Now consider the phrase 'twenty places', where 'places' normally carries stronger stress than 'twenty', i.e. is rhythmically stronger. We can make our "tree diagram" grow to look like this:



If we then look at this phrase in the context of a longer phrase 'twenty places further back', and build up the 'further back' part in a similar way, we would end up with an even more elaborate structure:



By analysing speech in this way we are able to show the relationships between strong and weak elements, and the different levels of stress that we find. The strength of any particular syllable can be measured by counting up the number of times an **s** symbol occurs above it; the levels in the sentence shown above can be diagrammed like this (leaving out syllables that have never received stress at any level):

```
                                                    s
                    s               s               s
s                   s               s               s
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
twen        ty      pla     ces     fur     ther    back
```

The above "**metrical grid**" may be correct for very slow speech, but we must now look at what happens to the rhythm in normal speech: many English speakers would feel that, although in 'twenty places' the right-hand foot is the stronger, the word 'twenty' is stronger than 'places' in 'twenty places further back' when spoken in conversational style. It is widely claimed that English speech tends towards a regular alternation between stronger and weaker, and tends to adjust stress levels to bring this about. The effect is particularly noticeable in cases such as the following, which all show the effect of what is called **stress-shift**:

compact (adjective)	**kəm'pækt**	*but* compact disc	**'kɒmpækt 'dɪsk**
thirteen	**θɜː'tiːn**	*but* thirteenth place	**'θɜːtiːnθ 'pleɪs**
Westminster	**west'mɪnstə**	*but* Westminster Abbey	**'westmɪnstər 'æbi**

In brief, it seems that stresses are altered according to context: we need to be able to explain how and why this happens, but this is a difficult question and one for which we have only partial answers.

An additional factor is that in speaking English we vary in how rhythmically we speak: sometimes we speak very rhythmically (this is typical of some styles of public speaking) while at other times we may speak arhythmically (that is, without rhythm) if we are hesitant or nervous. Stress-timed rhythm is thus perhaps characteristic of one style of speaking, not of English speech as a whole; one always speaks with *some* degree of rhythmicality, but the degree varies between a minimum value (arhythmical) and a maximum value (completely stress-timed rhythm).

It follows from what was stated earlier that in a stress-timed language all the feet are supposed to be of roughly the same duration. Many foreign learners of English are made to practise speaking English with a regular rhythm, often with the teacher beating time or clapping hands on the stressed syllables. It must be pointed out, however, that the evidence for the existence of truly stress-timed rhythm is not strong. There are many laboratory techniques for

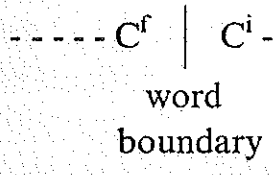
measuring time in speech, and measurement of the time intervals between stressed syllables in connected English speech has not shown the expected regularity; moreover, using the same measuring techniques on different languages, it has not been possible to show a real difference between "stress-timed" and "syllable-timed" languages. Experiments have shown that we tend to hear speech as more rhythmical than it actually is, and one suspects that this is what the proponents of the stress-timed rhythm theory have been led to do in their auditory analysis of English rhythm. However, one ought to keep an open mind on the subject, remembering that the large-scale, objective study of suprasegmental aspects of real speech is difficult to carry out, and much research remains to be done.

What, then, is the practical value of the traditional "rhythm exercise" for foreign learners? The argument about rhythm should not make us forget the very important difference in English between strong and weak syllables. Some languages do not have such a noticeable difference (which may, perhaps, explain the subjective impression of "syllable-timing"), and for native speakers of such languages who are learning English it can be helpful to practise repeating strongly rhythmical utterances since this forces the speaker to concentrate on making unstressed syllables weak. Speakers of languages like Japanese, Hungarian and Spanish – which do not have weak syllables to anything like the same extent as English does – may well find such exercises of some value (as long as they are not overdone to the point where learners feel they have to speak English as though they were reciting verse).

14.2 Assimilation

The device mentioned earlier that produces "mechanical speech" would contain all the words of English, each having been recorded in isolation. A significant difference in natural connected speech is the way that sounds belonging to one word can cause changes in sounds belonging to neighbouring words. Assuming that we know how the phonemes of a particular word would be realised when the word is pronounced in isolation, in cases where we find a phoneme realised differently as a result of being near some other phoneme belonging to a neighbouring word we call this an instance of **assimilation**. Assimilation is something which varies in extent according to speaking rate

and style: it is more likely to be found in rapid, casual speech and less likely in slow, careful speech. Sometimes the difference caused by assimilation is very noticeable, and sometimes it is very slight. Generally speaking, the cases that have most often been described are assimilations affecting consonants. As an example, consider a case where two words are combined, the first of which ends with a single final consonant (which we will call C^f) and the second of which starts with a single initial consonant (which we will call C^i); we can construct a diagram like this:



If C^f changes to become like C^i in some way, then the assimilation is called **regressive** (the phoneme that comes first is affected by the one that comes after it); if C^i changes to become like C^f in some way, then the assimilation is called **progressive**. In what ways can a consonant change? We have seen that the main differences between consonants are of three types:

i) differences in place of articulation;
ii) differences in manner of articulation;
iii) differences in voicing.

In parallel with this, we can identify assimilation of place, of manner and of voicing in consonants. Assimilation of place is most clearly observable in some cases where a final consonant (C^f) with alveolar place of articulation is followed by an initial consonant (C^i) with a place of articulation that is *not* alveolar. For example, the final consonant in 'that' **ðæt** is alveolar **t**. In rapid, casual speech the **t** will become **p** before a bilabial consonant, as in: 'that person' **ðæp pɜɪsn̩**; 'light blue' **laɪp bluː**; 'meat pie' **miːp paɪ**. Before a dental consonant, **t** will change to a dental plosive, for which the symbol is **t̪**, as in: 'that thing' **ðæt̪ θɪŋ**; 'get those' **get̪ ðəʊz**; 'cut through' **kʌt̪ θruː**. Before a velar consonant, the **t** will become **k**, as in: 'that case' **ðæk keɪs**; 'bright colour' **braɪk kʌlə**; 'quite good' **kwaɪk gʊd**. In similar contexts **d** would become **b**, **d̪** and **g**, respectively, and **n** would become **m**, **n̪** and **ŋ**. However, the same is not true of the other alveolar consonants: **s** and **z** behave differently, the only noticeable change being

that **s** becomes **ʃ**, and **z** becomes **ʒ** when followed by **ʃ** or **j**, as in: 'this shoe' **ðɪʃ ʃuː**; 'those years' **ðəʊʒ jɪəz**. It is important to note that the consonants that have undergone assimilation have not disappeared; in the above examples, the duration of the consonants remains more or less what one would expect for a two-consonant cluster. Assimilation of place is only noticeable in this regressive assimilation of alveolar consonants; it is not something that foreign learners need to learn to do.

Assimilation of manner is much less noticeable, and is only found in the most rapid and casual speech; generally speaking, the tendency is again for regressive assimilation and the change in manner is most likely to be towards an "easier" consonant – one which makes less obstruction to the airflow. It is thus possible to find cases where a final plosive becomes a fricative or nasal (e.g. 'that side' **ðæs saɪd**, 'good night' **gʊn naɪt**), but most unlikely that a final fricative or nasal would become a plosive. In one particular case we find progressive assimilation of manner, when a word-initial **ð** follows a plosive or nasal at the end of a preceding word: it is very common to find that the C^i becomes identical in manner to the C^f but with dental place of articulation. For example (the arrow symbol means "becomes"):

'in the'	**ɪn ðə**	→	**ɪn̪n̪ə**
'get them'	**get ðəm**	→	**get̪t̪əm**
'read these'	**riːd ðiːz**	→	**riːd̪d̪iːz**

It seems that the **ð** phoneme frequently occurs with no discernible friction noise.

Assimilation of voice is also found, but again only in a limited way. Only regressive assimilation of voice is found across word boundaries, and then only of one type; since this matter is important for foreign learners we will look at it in some detail. If C^f is a lenis (i.e. "voiced") consonant and C^i is fortis ("voiceless") we often find that the lenis consonant has no voicing; this is not a very noticeable case of assimilation, since, as was explained in Chapter 4, initial and final lenis consonants usually have little or no voicing anyway. When C^f is fortis ("voiceless") and C^i lenis ("voiced"), a context in which in

many languages C^f would become voiced, assimilation of voice never takes place; consider the following example: 'I like that black dog' **aɪ laɪk ðæt blæk dɒg**. It is typical of many foreign learners of English that they allow regressive assimilation of voicing to change the final **k** of 'like' to **g**, the final **t** of 'that' to **d** and the final **k** of 'black' to **g**. This creates a very strong impression of a foreign accent, and is something that should obviously be avoided.

Up to this point we have been looking at some fairly clear cases of assimilation across word boundaries. However, similar effects are also observable across morpheme boundaries and to some extent also within the morpheme. Sometimes in the latter case it seems that the assimilation is rather different from the word-boundary examples; for example, if in a syllable-final consonant cluster a nasal consonant precedes a plosive or a fricative in the same morpheme, then the place of articulation of the nasal is always determined by the place of articulation of the other consonant; thus: 'bump' **bʌmp**, 'tenth' **teŋ̟θ**, 'hunt' **hʌnt**, 'bank' **bæŋk**. It could be said that this assimilation has become fixed as part of the phonological structure of English syllables, since exceptions are almost non-existent. A similar example of a type of assimilation that has become fixed is the progressive assimilation of voice with the suffixes **s** and **z**; when a verb carries a third person singular '-s' suffix, or a noun carries an '-s' plural suffix or an '-'s' possessive suffix, that suffix will be pronounced as **s** if the preceding consonant is fortis ("voiceless") and as **z** if the preceding consonant is lenis ("voiced"), thus:

'cats' **kæts**	'dogs' **dɒgz**
'jumps' **dʒʌmps**	'runs' **rʌnz**
'Pat's' **pæts**	'Pam's' **pæmz**

Much more could be said about assimilation but, from the point of view of learning or teaching English pronunciation, to do so would not be very useful. It is essentially a natural phenomenon that can be seen in any sort of complex physical activity, and the only important matter is to remember the restriction, specific to English, on voicing assimilation mentioned above.

Assimilation creates something of a problem for phoneme theory; when, for example, **d** in 'good' **gʊd** becomes **g** in the context '. . . girl' (**gʊg gɜːl**) or **b** in the context '. . . boy' (**gʊb bɔɪ**), should we say that

one phoneme has been substituted for another? If we do this, how do we describe the assimilation in 'good thing', where **d** becomes dental **d** (**d̪**) before the **θ** of 'thing', or in 'good food', where **d** becomes a labiodental plosive before the **f** in 'food'? English has no dental or labiodental plosive phonemes, so in these cases, although there is clearly assimilation, there could not be said to be a substitution of one phoneme for another. The alternative is to say that assimilation causes a phoneme to be realised by a different allophone; this would mean that, in the case of **gʊg gɜɪl** and **gʊb bɔɪ**, the phoneme **d** of 'good' has velar and bilabial allophones. Traditionally, phonemes were supposed not to overlap in their allophones, so that the only plosives that could have allophones with bilabial place of articulation were **p** and **b**; this restriction is no longer looked on as so important.

14.3 Elision

AU14

The nature of **elision** may be stated quite simply: under certain circumstances sounds disappear; one might express this in more technical language by saying that in certain circumstances a phoneme may be realised as **zero**, or have **zero realisation** or be **deleted**. As with assimilation, elision is typical of rapid, casual speech. Producing elisions is something which foreign learners do not need to learn to do, but it is important for them to be aware that when native speakers of English talk to each other, quite a number of phonemes that the foreigner might expect to hear are not actually pronounced. We will look at some examples, although only a small number of the many possibilities can be given here.

i) Loss of weak vowel after **p, t, k**.
In words like 'potato', 'tomato', 'canary', 'perhaps', 'today', the vowel in the first syllable may disappear; the aspiration of the initial plosive takes up the whole of the middle portion of the syllable, resulting in these pronunciations (where **h** indicates aspiration):

pʰˈteɪtəʊ tʰˈmɑːtəʊ kʰˈneəri pʰˈhæps tʰˈdeɪ

ii) Weak vowel + **n, l** or **r** becomes syllabic (consonant see Chapter 9 for details of syllabic consonants); for example:

'tonight' **tn̩aɪt**, 'police' **pl̩ɪs**, 'correct' **kr̩ekt**

iii) Avoidance of complex consonant clusters.
It has been said that no normal English speaker would ever pronounce all the consonants between the last two words of the following:

'George the Sixth's throne' **dʒɔɪdʒ ðə sɪksθs θrəʊn**

Though this is not impossible to pronounce, something like **sɪksθrəʊn** is more likely. In clusters of three plosives or two plosives plus a fricative, the middle plosive may disappear, so that the following pronunciations result:

'acts' **æks**, 'looked back' **lʊk bæk**, 'scripts' **skrɪps**

iv) Loss of final **v** in 'of' before consonants; for example:

'lots of them' **lɒts ə ðəm**, 'waste of money' **weɪst ə mʌni**

It is difficult to know whether **contractions** of grammatical words should be regarded as examples of elision or not. The fact that they are regularly represented with special spelling forms makes them seem rather different from the above examples. The best-known cases are:

- 'had', 'would': spelt 'd, pronounced **d** (after vowels), **əd** (after consonants);
- 'is', 'has': spelt 's, pronounced **s** (after fortis consonants), **z** (after lenis consonants), except that after **s, z, ʃ, ʒ, tʃ, dʒ** 'is' is pronounced **ɪz** and 'has' is pronounced **əz** in contracted form;
- 'will': spelt 'll, pronounced **l** (after vowels), **l̩** (after consonants);
- 'have': spelt 've, pronounced **v** (after vowels), **əv** (after consonants);
- 'not': spelt n't, pronounced **nt** (after vowels), **n̩t** (after consonants)
(there are also vowel changes associated with n't, e.g. 'can' **kæn** – 'can't' **kɑːnt**; 'do' **duː** – 'don't' **dəʊnt**; 'shall' **ʃæl** – 'shan't' **ʃɑːnt**);
- 'are': spelt 're, pronounced **ə** after vowels, usually with some change in the preceding vowel, e.g. 'you' **juː** – 'you're' **jʊə** or **jɔː**, 'we' **wiː** – 'we're' **wɪə**, 'they' **ðeɪ** – 'they're' **ðeə**; linking **r** is used when a vowel follows, as explained in the next section. Contracted 'are' is also pronounced as **ə** or **ər** when following a consonant.

14.4 Linking

In our hypothetical "mechanical speech" all words would be separate units placed next to each other in sequence; in real connected speech, however, we sometimes link words together in special ways. The most familiar case is the use of **linking r**; the phoneme **r** does not occur in syllable-final position in the BBC accent, but when a word's spelling suggests a final **r**, and a word beginning with a vowel follows, the usual pronunciation is to pronounce with **r**. For example:

'here' **hɪə** *but* 'here are' **hɪər ə**
'four' **fɔː** *but* 'four eggs' **fɔːr egz**

BBC speakers often use **r** in a similar way to link words ending with a vowel, even when there is no "justification" from the spelling, as in:

'Formula A' **fɔːmjələr eɪ**
'Australia all out' **ɒstreɪlɪər ɔːl aʊt**
'media event' **miːdɪər ɪvent**

This has been called **intrusive r**; some English speakers and teachers still regard this as incorrect or sub-standard pronunciation, but it is undoubtedly widespread.

"Linking" and "intrusive r" are special cases of **juncture**; this name refers to the relationship between one sound and the sounds that immediately precede and follow it, and it has been given some importance in phonological theory. If we take the two words 'my turn' **maɪ tɜːn**, the relationship between **m** and **aɪ**, between **t** and **ɜː** and between **ɜː** and **n** is said to be one of **close juncture**. The sound **m** is preceded by silence and **n** is followed by silence, and so **m** and **n** are said to be in a position of **external open juncture**. The problem lies in deciding what the relationship is between **aɪ** and **t**; since we do not usually pause between words, there is no silence (or external open juncture) to indicate word division and to justify the space left in the transcription. But if English speakers hear **maɪ tɜːn** they can usually recognise this as 'my turn' and not 'might earn'. This is where the problem of internal open juncture (usually just called "juncture" for short) becomes apparent. What is it that makes perceptible the difference between **maɪ tɜːn** and **maɪt ɜːn**? The answer is that in the one case the **t** is aspirated (initial in 'turn'), and in the other case it is not (being final in 'might'). In addition to this, the **aɪ** diphthong is

shorter in 'might', but we will ignore this for the sake of a simpler argument. If a difference in meaning is caused by the difference between aspirated and unaspirated **t**, how can we avoid the conclusion that English has a phonemic contrast between aspirated and unaspirated **t**? The answer is, of course, that the position of a word boundary has some effect on the realisation of the **t** phoneme; this is one of the many cases in which the occurrence of different allophones can only be properly explained by making reference to units of grammar (something which was for a long time disapproved of by many phonologists).

Many ingenious minimal pairs have been invented to show the significance of juncture, a few of which are given below:

- 'might rain' **maɪt reɪn** (**r** voiced when initial in 'rain', **aɪ** short), vs. 'my train' **maɪ treɪn** (**r** voiceless following **t** in 'train')
- 'all that I'm after today' **ɔːl ðət aɪm ɑɪftə tədeɪ** (**t** unaspirated when final in 'that')
  'all the time after today' **ɔɪl ðə taɪm ɑɪftə tədeɪ** (**t** aspirated when initial in 'time')
- 'he lies' **hiɪ laɪz** ("clear l" initial in 'lies')
  'heal eyes' **hiːl aɪz** ("dark l" final in 'heal')
- 'keep sticking' **kiːp stɪkɪŋ** (**t** unaspirated after **s**; **iː** short)
  'keeps ticking' **kiːps tɪkɪŋ** (**t** aspirated in 'ticking')

Of course, the context in which the words occur almost always makes it clear where the boundary comes, and the juncture information is then redundant.

It should by now be clear that there is a great deal of difference between the way words are pronounced in isolation and in the context of connected speech. It would not be practical or useful to teach all learners of English to produce assimilations; practice in making elisions is more useful, and it is clearly valuable to do exercises related to rhythm and linking. Perhaps the most important consequence of what has been described in this chapter is that learners of English must be made very clearly aware of the problems that they will meet in listening to colloquial, connected speech.

Notes on problems and further reading

14.1 English rhythm is a controversial subject on which widely differing views have been expressed. On one side there are writers such as Abercrombie (1967) and Halliday (1967) who have set out an elaborate theory of the rhythmical structure of English speech (including foot theory). On the other side there are sceptics like Crystal (1969) who reject the idea of an inherent rhythmical pattern. The distinction between physically measurable time intervals and subjective impressions of rhythmicality is discussed in Roach (1982) and Lehiste (1977). Adams (1979) presents a review and experimental study of the subject, and concludes that, despite the theoretical problems, there is practical value in teaching rhythm to learners of English. The treatment of rhythmicality as a matter of degree is presented in Crystal (1969: 161–5). The "stress-timed / syllable-timed" dichotomy is generally agreed in modern work to be an oversimplification; a more common view is that all languages display characteristics of both types of rhythm, but may be closer to one or the other; see Mitchell (1969) and Dauer (1983). Dauer's theory makes possible comparisons between different languages in terms of their relative positions on a scale from maximally stress-timed to maximally syllable-timed (see for example Dimitrova, 1997).

For some writers concerned with English language teaching, the notion of rhythm is a more practical matter of making a sufficiently clear difference between strong and weak syllables, rather than concentrating on a rigid timing pattern, as I suggest at the end of Section 14.1; see, for example, Taylor (1981); also, Mortimer (1984) contains practice material on rhythm.

The treatment of rhythmical hierarchy is based on the theory of metrical phonology. Hogg and McCully (1987) give a full explanation of this, but it is difficult material. Goldsmith (1990: Chapter 4), and Katamba (1989: Chapter 11.1) are briefer and somewhat simpler. A recent paper by Fudge (1999) discusses the relationship between syllables, words and feet. James (1988) explores the relevance of metrical phonology to language learning.

14.2 Assimilation of place with specific reference to alveolar consonants is described in Gimson (1960). It is important to realise that the traditional view of assimilation as a change from one

phoneme to another is naive; modern instrumental studies in the broader field of coarticulation show that when assimilation happens one can often show how there is some sort of combination of articulatory gestures. In 'good girl', for example, it is not a simple matter of the first word ending *either* in **d** *or* in **g**, but rather a matter of the extent to which alveolar and/or velar closures are achieved. There may be an alveolar closure immediately preceding and overlapping with a velar closure; there may be simultaneous alveolar and velar closure, or a velar closure followed by slight contact but not closure in the alveolar region. There are many other possibilities.

14.3 An essential part of acquiring fluency in English is learning to produce connected speech without gaps between words, and this is the practical importance of linking. There are several papers on "intrusive r": see Windsor Lewis (1975), Pring (1976), Windsor Lewis (1977) and Fox (1978).

An obvious question to be asked in relation to juncture is whether 'internal open juncture' can actually be heard. Jones (1931) implies that it can, but experimental work (e.g. O'Connor and Tooley, 1964) suggests that in many cases it is not perceptible unless a speaker is deliberately trying to avoid ambiguity. It is interesting to note that some phonologists of the 1950s and 1960s felt it necessary to invent a 'phoneme' of juncture in order to be able to transcribe minimal pairs like 'grey tape' / 'great ape' unambiguously without having to refer to grammatical boundaries; see, for example, Trager and Smith (1951).

Notes for teachers

There is a lot of disagreement about the importance of the various topics in this chapter from the language teacher's point of view. My feeling is that two separate matters are sometimes mixed up: the practice and study of connected speech is agreed by everyone to be of great importance, but this can sometimes result in some relatively unimportant aspects of speech (e.g. assimilation, juncture) being given more emphasis than they should.

In looking at the importance of studying aspects of speech above the segmental level some writers have claimed that learners can come

to identify an overall "feel" of the pronunciation of the language being learned. Differences between languages have been described in terms of their **articulatory settings**, that is, overall articulatory posture, by Honikman (1964). She describes such factors as lip mobility and tongue-setting for English, French and other languages. The notion seems a useful one, although it is difficult to confirm these settings scientifically.

Audio Unit 14 is liable to come as something of a surprise to students who have not had the experience of examining colloquial English speech before. The main message to get across is that concentration on selective, analytic listening will help them to recognise what is being said, and that practice usually brings confidence.

Written exercises

1 Divide the following sentences up into feet, using a single vertical line (|) as a boundary symbol. If a sentence starts with an unstressed syllable, leave it out of consideration – it doesn't belong in a foot.

   a) A bird in the hand is worth two in the bush.
   b) Over a quarter of a century has elapsed since his death.
   c) Computers consume a considerable amount of money and time.
   d) Most of them have arrived on the bus.
   e) Newspaper editors are invariably underworked.

2 Draw tree diagrams of the rhythmical structure of the following phrases.

   a) Christmas present
   b) Rolls Royce
   c) pet food dealer
   d) Rolls-Royce rally event

3 The following sentences are given in spelling and in a "slow, careful" phonemic transcription. Rewrite the phonemic transcription as a "broad phonetic" one so as to show likely assimilations, elisions and linking.

a) One cause of asthma is supposed to be allergies
wʌn kɔːz əv æsθmə ɪz səpəʊzd tə bi ælədʒiz
[ ]

b) What the urban population could use is better trains
wɒt ðiː ɜːbən pɒpjəleɪʃn̩ kʊd juːz ɪz betə treɪnz
[ ]

c) She acts particularly well in the first scene
ʃi ækts pətɪkjələli wel ɪn ðə fɜːst siːn
[ ]

15 Intonation 1

Many of the previous chapters have been concerned with the description of phonemes, and in Section 5.2 it was pointed out that the subject of phonology includes not just this aspect (which is usually called **segmental phonology**) but also several others. In Chapters 10 and 11, for example, we studied stress. Clearly, stress has linguistic importance and is therefore an aspect of the phonology of English that must be described, but it is not usually regarded as something that is related to individual segmental phonemes; normally, stress is said to be something that is applied to (or is a property of) syllables, and is therefore part of the **suprasegmental phonology** of English. Another part of suprasegmental phonology is **intonation**, and the next five chapters are devoted to this subject.

What is intonation? No definition is completely satisfactory, but any attempt at a definition must recognise that the **pitch** of the voice plays the most important part. Only in very unusual situations do we speak with fixed, unvarying pitch, and when we speak normally the pitch of our voice is constantly changing. One of the most important tasks in analysing intonation is to listen to the speaker's pitch and recognise what it is doing; this is not an easy thing to do, and it seems to be a quite different skill from that acquired in studying segmental phonetics. We describe pitch in terms of **high** and **low**, and some people find it difficult to relate what they hear in someone's voice to a scale ranging from low to high. We should remember that "high" and "low" are arbitrary choices for end-points of the pitch scale. It would be perfectly reasonable to think of pitch as ranging instead from "light" to "heavy", for example, or from "left" to "right", and people who have difficulty in "hearing" intonation patterns are generally only having difficulty in relating what they hear (which is the same as what everyone else hears) to this "pseudo-spatial" representation.

It is very important to make the point that we are not interested in all aspects of a speaker's pitch; the only things that should interest us are those which carry some linguistic information. If a speaker tries to talk while riding fast on a horse, his or her pitch will make a lot of sudden rises and falls as a result of the irregular movement; this is something which is outside the speaker's control and therefore cannot be linguistically significant. Similarly, if we take two speakers at random we will almost certainly find that one speaker typically speaks with lower pitch than the other; the difference between the two speakers is not linguistically significant because their habitual pitch level is determined by their physical structure. But an individual speaker does have control over his or her own pitch, and may choose to speak with a higher than normal pitch; this is something which *is* potentially of linguistic significance.

A word of caution is needed in connection with the word **pitch**. Strictly speaking, this should be used to refer to an auditory sensation experienced by the hearer. The rate of vibration of the vocal folds – something which is physically measurable, and which is related to activity on the part of the speaker – is the **fundamental frequency** of voiced sounds, and should not be called 'pitch'. However, as long as this distinction is understood, it is generally agreed that the term 'pitch' is a convenient one to use informally to refer both to the subjective sensation and to the objectively measurable fundamental frequency.

We have established that for pitch differences to be linguistically significant, it is a necessary condition that they should be under the speaker's control. There is another necessary condition and that is that a pitch difference must be **perceptible**; it is possible to detect differences in the **frequency** of the vibration of a speaker's voice by means of laboratory instruments, but these differences may not be great enough to be heard by a listener as differences in pitch. Finally, it should be remembered that in looking for linguistically significant aspects of speech we must always be looking for *contrasts*; one of the most important things about any unit of phonology or grammar is the set of items it contrasts with. We know how to establish which phonemes are in contrast with **b** in the context **-in**; we can substitute other phonemes (e.g. **p**, **s**) to change the identity of the word from 'bin' to 'pin' to 'sin'. Can we establish such units and contrasts in intonation?

15.1 Form and function in intonation

To summarise what was said above, we want to know the answers to two questions about English speech:

i) What can we observe when we study pitch variations?
ii) What is the linguistic importance of the phenomena we observe?

These questions might be rephrased more briefly as:

i) What is the **form** of intonation?
ii) What is the **function** of intonation?

We will begin by looking at intonation in the shortest piece of speech we can find – the single syllable. At this point a new term will be introduced: we need a name for a continuous piece of speech beginning and ending with a clear pause, and we will call this an **utterance**. In this chapter, then, we are going to look at the intonation of one-syllable utterances. These are quite common, and give us a comparatively easy introduction to the subject.

Two common one-syllable utterances are 'yes' and 'no'. The first thing to notice is that we have a choice of saying these with the pitch remaining at a constant level, or with the pitch changing from one level to another. The word we use for the overall behaviour of the pitch in these examples is **tone**; a one-syllable word can be said with either a **level tone** or a **moving tone**. If you try saying 'yes' or 'no' with a level tone (rather as though you were trying to sing them on a steady note) you may find the result does not sound natural, and indeed English speakers do not use level tones on one-syllable utterances very frequently. Moving tones are more common. If English speakers want to say 'yes' or 'no' in a definite, final manner they will probably use a **falling** tone – one which descends from a higher to a lower pitch. If they want to say 'yes?' or 'no?' in a questioning manner they may say it with a **rising** tone – a movement from a lower pitch to a higher one.

Notice that already, in talking about different tones, some idea of function has been introduced; speakers are said to select from a choice of tones according to how they want the utterance to be heard, and it is implied that the listener will hear one-syllable utterances said with different tones as sounding different in some way. During the development of modern phonetics in the twentieth

century it was for a long time hoped that scientific study of intonation would make it possible to state what the function of each different aspect of intonation was, and that foreign learners could then be taught rules to enable them to use intonation in the way that native speakers use it. Few people now believe this to be possible. It is certainly possible to produce a few general rules, and some will be given in this course, just as a few general rules for word stress were given in Chapters 10 and 11. However, these rules are certainly not adequate as a complete practical guide to how to use English intonation. My treatment of intonation is based on the belief that foreign learners of English at advanced levels who may use this course should be given training to make them better able to recognise and copy English intonation. The only really efficient way to learn to *use* the intonation of a language is the way a child acquires the intonation of its first language, and the training referred to above should help the adult learner of English to acquire English intonation in a similar (though much slower) way – through listening to and talking to English speakers. It is perhaps a discouraging thing to say, but learners of English who are not able to talk regularly with native speakers of English, or who are not able at least to listen regularly to colloquial English, are not likely to learn English intonation, although they may learn very good pronunciation of the segments and use stress correctly.

15.2. Tone and tone languages

AU15, Exs 1–3

In the preceding section we mentioned three simple possibilities for the intonation used in pronouncing the one-word utterances 'yes' and 'no'. These were: level, fall and rise. It will often be necessary to use symbols to represent tones, and for this we will use marks placed before the syllable in the following way (phonemic transcription will not be used in these examples – words are given in spelling):

Level	_yes	_no
Falling	ˌyes	ˌno
Rising	ˏyes	ˏno

Obviously, this simple system for tone transcription could be extended, if we wished, to cover a greater number of possibilities. For example, if it were important to distinguish between a high level and low level tone for English we could do it in this way:

High level	ˉyes	ˉno
Low level	ˍyes	ˍno

Although in English we do on occasions say ˉyes or ˉno and on other occasions ˍyes or ˍno, no speaker of English would say that the meaning of the words 'yes' and 'no' is different with the different tones. (As will be seen below, we will not use the symbols for high and low level tones in the description of English intonation.) But there are many languages in which the tone can determine the meaning of a word, and changing from one tone to another can completely change the meaning. For example, in Kono, a language of West Africa, we find the following (meanings given in brackets):

High level	ˉ**bɛŋ** ('uncle')	ˉ**buu** ('horn')
Low level	ˍ**bɛŋ** ('greedy')	ˍ**buu** ('to be cross')

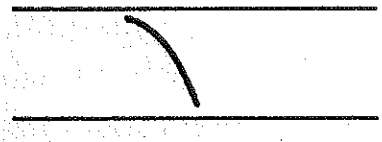
Similarly, while we can hear a difference between English ˍyes, ˌyes and ˈyes, and between ˍno, ˌno and ˈno, there is not a difference in meaning in such a clear-cut way as in Mandarin Chinese, where, for example, ˉ**ma** means 'mother', ˌ**ma** means 'hemp' and ˈ**ma** means 'scold'. Languages such as the above are called **tone languages**; although to most speakers of European languages they may seem strange and exotic, such languages are in fact spoken by a very large proportion of the world's population. In addition to the many dialects of Chinese, many other languages of South-East Asia (e.g. Thai, Vietnamese) are tone languages; so are very many African languages, particularly those of the South and West, and a considerable number of Native American languages. English, however, is not a tone language, and the function of tone is much more difficult to define than in a tone language.

15.3 Complex tones and pitch height

We have introduced three simple tones that can be used on one-syllable English utterances: level, fall and rise. However, other more complex tones are also used. One that is quite frequently found is the **fall–rise** tone, where the pitch descends and then rises again. Another complex tone, much less frequently used, is the **rise–fall** in which the pitch follows the opposite movement. We will not consider any more complex tones, since these are not often encountered and are of little importance.

One further complication should be mentioned here. Each speaker has his or her own normal pitch range: a top level which is the highest pitch normally used by the speaker, and a bottom level that the speaker's pitch normally does not go below. In ordinary speech, the intonation tends to take place within the lower part of the speaker's pitch range, but in situations where strong feelings are to be expressed it is usual to make use of extra pitch height. For example, if we represent the pitch range by drawing two parallel lines representing the highest and lowest limits of the range, then a normal unemphatic 'yes' could be diagrammed like this:

but a strong, emphatic 'yes' like this:



We will use a new symbol ↑ (a vertical upward arrow) to indicate extra pitch height, so that we can distinguish between:

ˌyes and ↑ˌyes

Any of the tones presented in this chapter may be given extra pitch height, but since this course is based on normal, unemotional speech, it will not be necessary to use the symbol very frequently.

15.4 Some functions of English tones

AU15, Ex 4

In this chapter only a very small part of English intonation has been introduced. We will now see if it is possible to state in what circumstances the different tones are used within the very limited context of the words 'yes' and 'no' said in isolation. We will look at some typical occurrences; no examples of extra pitch height will be considered here, so the examples should be thought of as being said relatively low in the speaker's pitch range.

Fall ˌyes ˌno

This is the tone about which least needs to be said, and which is usually regarded as more or less "neutral". If someone is asked a

question and replies ˌyes or ˌno it will be understood that the question is now answered and that there is nothing more to be said. The fall could be said to give an impression of "finality".

Rise ˌyes ˌno

In a variety of ways, this tone conveys an impression that something more is to follow; a typical occurrence in a dialogue between two speakers whom we shall call A and B might be the following:

A (wishing to attract B's attention): Excuse me.
B: ˌyes

(B's reply is, perhaps, equivalent to 'what do you want?') Another quite common occurrence would be:

A: Do you know John Smith?

One possible reply from B would be ˌyes, inviting A to continue with what she intends to say about John Smith after establishing that B knows him. To reply instead ˌyes would give a feeling of "finality", of "end of the conversation"; if A did have something to say about John Smith, the response with a fall would make it difficult for A to continue.

We can see similar "invitations to continue" in someone's response to a series of instructions or directions. For example:

A: You start off on the ring road . . .
B: ˌyes
A: turn left at the first roundabout . . .
B: ˌyes
A: and ours is the third house on the left.

Whatever B replies to this last utterance of A, it would be most unlikely to be ˌyes again, since A has clearly finished her instructions and it would be pointless to "prompt" her to continue.

With 'no', a similar function can be seen. For example:

A: Have you seen Ann?

If B replies ˌno, he implies quite clearly that he has no interest in continuing with that topic of conversation. But a reply of ˌno would be an invitation to A to explain why she is looking for Ann, or why she does not know where she is.

Similarly, someone may ask a question that implies readiness to present some new information. For example:

A: Do you know what the longest balloon flight was?

If B replies ˌno he is inviting A to tell him, while a response of ˏno could be taken to mean that he does not know and is not expecting to be told. This is, in fact, a common cause of misunderstanding in English conversation, when a question such as A's above might be a request for information or an offer to provide some.

Fall–rise ˬyes ˬno

The fall–rise is used a lot in English and has some rather special functions. In the present context we will only consider one fairly simple one, which could perhaps be described as "limited agreement" or "response with reservations". Examples may make this clearer:

A: I've heard that it's a good school.
B: ˬyes

B's reply would be taken to mean that he would not completely agree with what A said, and A would probably expect B to go on to explain why he was reluctant to agree. Similarly:

A: It's not really an expensive book, is it?
B: ˬno

The fall–rise in B's reply again indicates that he would not completely agree with A. Fall–rise in such contexts almost always indicates both something "given" or "conceded" and at the same time some "reservation" or "hesitation". This use of intonation will be returned to in Chapter 19.

Rise–fall ˆyes ˆno

This is used to convey rather strong feelings of approval, disapproval or surprise. It is not usually considered to be an important tone for foreign learners to acquire, although it is still useful practice to learn to distinguish it from other tones. Here are some examples:

A: You wouldn't do an awful thing like that, would you?
B: ˆno

A: Isn't the view lovely!
B: ˏyes

A: I think you said it was the best so far.
B: ˏyes

Level _yes _no

This tone is certainly used in English, but in a rather restricted context: it almost always conveys (on single-syllable utterances) a feeling of saying something routine, uninteresting or boring. A teacher calling the names of pupils from a register will often do so using a level tone on each name, and the pupils are likely to respond with _yes when their name is called. Similarly, if one is being asked a series of routine questions for some purpose – such as applying for an insurance policy – one might reply to each question of a series (like 'Have you ever been in prison?', 'Do you suffer from any serious illness?', 'Is your eyesight defective?', etc.) with _no.

A few "meanings" have been suggested for the five tones that have been introduced, but each tone may have many more such meanings. Moreover, it would be quite wrong to conclude that in the above examples only the tones given would be appropriate; it is, in fact, almost impossible to find a context where one could not substitute a different tone. This is not the same thing as saying that any tone can be used in any context: the point is that no particular tone has a unique "privilege of occurrence" in a particular context. When we come to look at more complex intonation patterns, we will see that defining intonational "meanings" does not become any easier.

Notes on problems and further reading

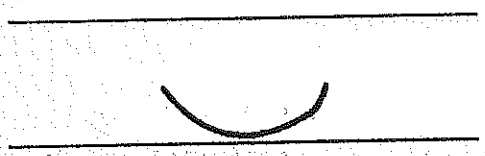
To devote five chapters to intonation may seem excessive, but I feel that this is necessary since the subject is difficult and complex, and needs to be explained at considerable length if the explanation is to be intelligible. The study of intonation went through many changes in the twentieth century. The most intensive theoretical development began during the 1940s. In the United States the theory that evolved was based on 'pitch phonemes' (Pike, 1945): four contrastive pitch levels were established and intonation was described basically in terms of a series of movements from one of these levels to another.

This approach was further developed in Trager and Smith (1951). Although this 'pitch phoneme' theory became an orthodoxy, it was consistently attacked by one American linguist, D. Bolinger (e.g. Bolinger, 1951). In Britain the 'tone-unit' or 'tonetic' approach begun by H. E. Palmer in the 1920s (Palmer, 1924) was developed by (among others) Kingdon (1958), O'Connor and Arnold (1962) and Halliday (1967). These two different theoretical approaches became gradually more elaborate; in the American case perhaps the most elaborate exposition was in Trager (1964), while O'Connor and Arnold produced an extended version of their treatment in their second edition (1973) which was very difficult to learn in full. Since the 1970s it has become clear that, despite their complexity, such frameworks are inadequate for dealing with natural spontaneous speech. In Britain the most influential work leading to this recognition was Crystal (1969). I have tried in this course to reflect some of the more recently developed ideas for dealing with intonation, although the treatment remains essentially within the conventions of the British tradition. A good introduction to the theoretical issues is Cruttenden (1997). A more difficult (though very valuable) book on theory is Ladd (1996).

15.1 The amount of time to be spent on learning about tone languages should depend to some extent on your background. Those whose native language is a tone language should be aware of the considerable linguistic importance of tone in such languages; often it is extremely difficult for people who have spoken a tone language all their life to learn to observe their own use of tone objectively. The study of tone languages when learning English is less important for native speakers of non-tone languages, but most students seem to find it an interesting subject. A good introduction is Ladefoged (1993: 226–32). The classic work on the subject is Pike (1948), while more modern treatments are Hyman (1975: 212–29, Fromkin (1978) and Katamba (1989: Chapter 10).

Many analyses within the British approach to intonation include among tones both 'high' and 'low' varieties. For example, O'Connor and Arnold (1973) distinguished between 'high fall' and 'low fall' (the former starting from a high pitch, the latter from mid), and also between 'low rise' and 'high rise' (the latter rising to a higher point

than the former). Some writers have high and low versions of all tones. Compared with the proposed establishing of a separate feature of 'extra pitch height' (which is explained more fully in Section 18.1), this is unnecessary duplication. However, if one adds extra pitch height to a tone, one has not given all possible detail about it. If we take as an example a fall–rise without extra pitch height:



then something symbolised as ↑ ˬ could be any of the following:

It would be possible to extend our framework to distinguish between these possibilities, but I do not believe it would be profitable to do so. Several writers have included in their set of tones **fall–rise–fall** and **rise–fall–rise**; I have seldom felt the need to recognise these as distinct from rise–fall and fall–rise respectively.

Note for teachers

As explained above, some students may be perfectly well able to discriminate between tones, but have difficulty in labelling them as 'fall', 'rise', etc. I find that about five per cent of the students I teach are never able to overcome this difficulty (even though they may have perfect hearing and in some cases a high level of linguistic and musical ability). Of the remainder, a few are especially gifted and cannot understand how anyone could find the task difficult, and most others eventually learn after five or ten hours of practical classes. Many students find it very helpful to work with a computer showing a real-time display of their pitch movements as they speak.

Written exercise

In the following sentences and bits of dialogue, each underlined syllable must be given an appropriate tone mark. Write a tone mark just in front of each of the syllables.

1 This train is for Leeds, York, Darlington and Durham.

2 Can you give me a lift?
Possibly Where to?

3 No! Certainly not! Go away!

4 Did you know he'd been convicted of drunken driving?
No!

5 If I give him money he goes and spends it
If I lend him the bike he loses it
He's completely unreliable

form, and you will notice that no punctuation is used; the reason for this is that intonation and stress are the vocal equivalents of written punctuation, so that when these are transcribed it would be unnecessary or even confusing to include punctuation as well.)

AU16, Exs 1 & 2

Let us begin with a one-syllable utterance:

ˌyou

We will underline syllables that carry a tone. Now consider this utterance:

is it ˌyou

The third syllable is more prominent than the other two and carries a rising tone. The other two syllables will normally be much less prominent, and be said on a level pitch. Why do we not say that each of the syllables 'is' and 'it' carries a level tone? This is a difficult question that will be examined more fully later; for the present I will answer it (rather unsatisfactorily) by saying that it is unusual for a syllable said on a level pitch to be so prominent that it would be described as carrying a level *tone*. To summarise the analysis of 'is it ˌyou' so far, it is an utterance of three syllables, consisting of one tone-unit; the only syllable that carries a tone is the third one. From now on, a syllable which carries a tone will be called a **tonic syllable**. It has been mentioned several times that tonic syllables have a high degree of prominence; prominence is, of course, a property of stressed syllables, and a tonic syllable not only carries a tone (which is something related to intonation) but also a type of stress that will be called **tonic stress**. (Some writers use the terms **nucleus** and **nuclear stress** for **tonic syllable** and **tonic stress**.)

The example can now be extended:

˪John is it ˌyou

(A fall–rise is used quite commonly in calling someone's name out.) If there is a clear pause (silence) between '˪John' and 'is it ˌyou' then, according to the definition of an utterance given in Chapter 15, there are two utterances; however, it is quite likely that a speaker would say '˪John is it ˌyou' with no pause, so that the four syllables would make up a single utterance. In spite of the absence of any pause, the

16 Intonation 2

16.1 The tone-unit

In Chapter 15 it was explained that some of the world's languages are "tone languages", in which substituting one distinctive tone for another on a particular word or morpheme can cause a change in the dictionary ("lexical") meaning of that word or morpheme, or in some aspect of its grammatical categorisation. Although tones or pitch differences are used for other purposes English is one of the languages that do not use tone in this way. Such languages are sometimes called **intonation languages**. In tone languages the main suprasegmental contrastive unit is the tone, which is usually linked to the phonological unit that we call the syllable. It could be said that someone analysing the function and distribution of tones in a tone language would be mainly occupied in examining utterances syllable by syllable, looking at each syllable as an independently variable item. In Chapter 15, five tones found on English one-syllable utterances were introduced, and if English were spoken in isolated monosyllables, the job of tonal analysis would be a rather similar one to that described for tone languages. However, when we look at continuous speech in English utterances we find that these tones can only be identified on a small number of particularly prominent syllables. For the purposes of analysing intonation a unit generally greater in size than the syllable is needed, and this unit is called the **tone-unit**; in its smallest form the tone-unit may consist of only one syllable, so it would in fact be wrong to say that it is always composed of more than one syllable. The tone-unit is difficult to define, and one or two examples may help to make it easier to understand the concept. (As explained in Chapter 15, examples used to illustrate intonation transcription are usually given in spelling

utterance would normally be regarded as divided into two tone-units: '˩John' and 'is it ˧you'. Since it is very difficult to lay down the conditions for deciding where the boundaries between tone-units exist, the discussion of this matter must wait until later.

It should be possible to see now that the tone-unit has a place in a range of phonological units that are in a **hierarchical relationship**: speech consists of a number of **utterances** (the largest unit that we shall consider); each utterance consists of one or more **tone-units**; each tone-unit consists of one or more **feet**; each foot consists of one or more **syllables**; each syllable consists of one or more **phonemes**.

16.2 The structure of the tone-unit

In Chapter 8 the structure of the English syllable was examined in some detail. Like the syllable, the tone-unit has a fairly clearly-defined internal structure, but the only component that has been mentioned so far is the tonic syllable. The first thing to be done is to make more precise the role of the tonic syllable in the tone-unit. Most tone-units are of a type that we call **simple**, and the sort that we call **compound** are not discussed in this chapter. Each simple tone-unit has one and only one tonic syllable; this means that the tonic syllable is an obligatory component of the tone-unit. (Compare the role of the vowel in the syllable.) We will now see what the other components may be.

The head

Consider the following one-syllable utterance:

˧those

We can find the same tonic syllable in a long utterance (still of one tone-unit):

'give me ˧those

The rest of the tone-unit in this example is called the **head**. Notice that the first syllable has a stress mark; this is important. A head is all of that part of a tone-unit that extends from the first stressed syllable up to (but not including) the tonic syllable. It follows that if there is no stressed syllable before the tonic syllable, there cannot be a head. In the above example, the first two syllables (words) are the head of

the tone-unit. In the following example, the head consists of the first five syllables:

'Bill 'called to 'give me ˌthese

As was said a little earlier, if there is no stressed syllable preceding the tonic syllable, there is no head. This is the case in the following example:

in an ˌhour

Neither of the two syllables preceding the tonic syllable is stressed. The syllables 'in an' form a **pre-head**, which is the next component of the tone-unit to be introduced.

The pre-head

The pre-head is composed of all the unstressed syllables in a tone-unit preceding the first stressed syllable. Thus pre-heads are found in two main environments:

i) when there is no head (i.e. no stressed syllable preceding the tonic syllable), as in this example:

in an ˌhour

ii) when there is a head, as in this example:

in a 'little 'less than an ˌhour

In this example, the pre-head consists of 'in a', the head consists of ''little 'less than an', and the tonic syllable is 'ˌhour'.

The tail

It often happens that some syllables follow the tonic syllable. Any syllables between the tonic syllable and the end of the tone-unit are called the **tail**. In the following examples, each tone-unit consists of an initial tonic syllable and a tail:

ˌlook at it ˌwhat did you say
ˌboth of them were here

When it is necessary to mark stress in a tail, we will use a special symbol, a raised dot · for reasons that will be explained later. The above examples should, then, be transcribed as follows:

,look at it ,what did you ·say
,both of them were ·here

This completes the list of tone-unit components. If we use brackets to indicate optional components (that is, components which may be present or may be absent), we can summarise tone-unit structure as follows:

(pre-head) (head) tonic syllable (tail)

or, more briefly, as:

(PH) (H) TS (T)

To illustrate this more fully, let us consider the following passage, which is transcribed from a tape-recording of spontaneous speech (the speaker is describing a picture). When we analyse longer stretches of speech, it is necessary to mark the places where tone-unit boundaries occur (that is, where one tone-unit ends and another begins, or where a tone-unit ends and is followed by a pause, or where a tone-unit begins following a pause). It was mentioned above that tone-units are sometimes separated by silent pauses and sometimes not; pause-type boundaries can be marked by double vertical lines (‖) and non-pause boundaries with a single vertical line (|). In practice it is not usually important to mark pauses at the beginning and end of a passage. In the rest of the book I put no lines on short examples and only single lines around longer ones; the boundaries within a passage are much more important.

‖ and then 'nearer to the ˌfront ‖ on the ˌleft | theres a 'bit of ˌforest | 'coming 'down to the ˌwaterside ‖ and then a 'bit of a ˌbay ‖

We can mark their structure as follows:

‖ PH	H	TS	‖ PH	TS	PH
and then	'nearer to the	ˌfront	on the	ˌleft	there's a

H	TS	T	H	TS	T	‖
'bit of	ˌfor	est	'coming 'down to the	ˌwa	terside	

PH	H	TS	‖
and then a	'bit of a	ˌbay	

The above passage contains five tone-units. Notice that in the third

tone-unit, since it is the tonic syllable rather than the word that carries the tone, it is necessary to divide the word 'forest' into two parts, 'for' for and 'est' ɪst (it could be argued that the syllables should be divided 'fo' and 'rest', but this is not important here). This example shows clearly how the units of phonological analysis can sometimes be seen to differ from those of grammatical analysis.

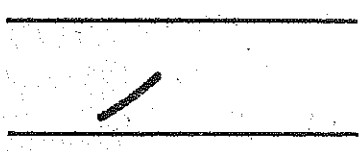
16.3 Pitch possibilities in the simple tone-unit

It has been said several times in this chapter that tone is carried by the tonic syllable, and it is now necessary to examine this statement more carefully. Before doing this, another general statement will be made (and will also need further explanation): intonation is carried by the tone-unit.

In a one-syllable utterance, the single syllable must have one of the five tones described in Chapter 15. In a tone-unit of more than one syllable, the tonic syllable must have one of those tones. If the tonic syllable is the final syllable, the tone will not sound much different from that of a corresponding one-syllable tone-unit. For example, the word 'here' will be said in much the same way in the following:

ˌhere 'shall we 'sit ˌhere

However, if there are other syllables following the tonic syllable (i.e. there is a tail), we find that the pitch movement of the tone is not completed on the tonic syllable. If a tail follows a tonic syllable that has a rising tone, it will almost always be found that the syllable or syllables of the tail will continue to move upwards from the pitch of the tonic syllable. For example, if the word 'what' is said on a rising tone, 'ˌwhat', it might have a pitch movement that could be diagrammed like this:



The four syllables in 'ˌwhat did you say' might be said like this:

these – called 'sense-groups' – bounded by places where 'pauses *may* be made' and consisting of 'a few words in close grammatical connexion'. Trim (1959)* criticises this proposal, saying that Jones' 'sense-group' is defined in semantic terms and the 'breath-group' in physiological terms, whereas we should be concerned with phonetic and phonological units and definitions. Instead, he proposes that the unit used should be the 'tone-group', defined in terms of rhythm and pitch movements, and that we should distinguish between 'major' and 'minor' tone-groups. The minor tone-group corresponds to the tone-unit used in this course, and the idea of a larger unit (the major tone-unit) is a valuable one that will be discussed further in the notes on Chapter 19.

Note for teachers

The move from tones to tone-units is a difficult one, and I feel it is advisable at this stage to use only slow, careful speech for exercises (Audio Units 15 and 16). More difficult exercises follow later (Audio Units 18 and 19).

Written exercises

1 Here is a list of single tonic syllables. Add a number of extra syllables (as specified by the number in brackets) to make a tail. Example: go (2); Answer: go for it

a) buy (3)
b) hear (1)
c) talk (2)

(The answers section gives some possible versions.)

2 Now expand the following tonic syllables by putting heads in front of them, containing the number of stressed syllables indicated in brackets. Example: (2) dark; Answer: 'John was a'fraid of the dark

a) (1) step
b) (3) train
c) (2) hot

* Not surprisingly, the fact that Trim's article is written in phonemic transcription with many words not separated by spaces makes it hard to read.

with the pitch of the syllables in the tail getting progressively higher. In such cases, the tonic syllable is the syllable on which the pitch movement of the tone begins, but that pitch movement is completed over the rest of the tone-unit (i.e. the tail). If, in rising progressively higher, the pitch reaches the highest part of the speaker's normal pitch range, subsequent syllables will continue at that top level.

We find a similar situation with the falling tone. On a single syllable '\why', the pitch movement might be of this sort:

but if there are syllables following, the fall may not be completed on the tonic syllable: \why did you go

Again, if the speaker's lowest pitch is reached before the end of the tail, the pitch continues at the bottom level. In the case of a level tone, syllables following in the tail will, of course, continue at the same level; since level tone is to be treated as a rather unusual type of tone, we will not examine it in more detail at this stage. The situation is more complicated when we have a tail following a fall–rise or a rise–fall, and this is described in Chapter 17.

Notes on problems and further reading

Almost all British analyses use a unit similar or identical to what I call a tone-unit, and link intonation to higher-level grammatical units. Different writers use different names: 'tone-group', 'sense-group', 'intonation unit' and 'tone-unit' are all more or less synonymous. It is possible to represent intonation as a simple sequence of tonic and non-tonic stressed syllables, and pauses, with no higher-level organisation; an example of this is the transcription used in the Spoken English Corpus (Williams, 1996). An early attempt at defining intonation units was that of Jones (1975: Chapter 30), where stretches of speech between pauses were called 'breath-groups' and marked with double vertical lines (‖), and smaller stretches within

17 Intonation 3

In Chapter 16 the structure of the tone-unit was introduced and it was explained that when a tonic syllable is followed by a tail, that tail continues and completes the tone begun on the tonic syllable. Examples were given to show how this happens in the case of rising and falling tones. We now go on to consider the rather more difficult cases of fall–rise and rise–fall tones.

17.1 Fall–rise and rise–fall tones followed by a tail

AU17, Exs 1 & 2

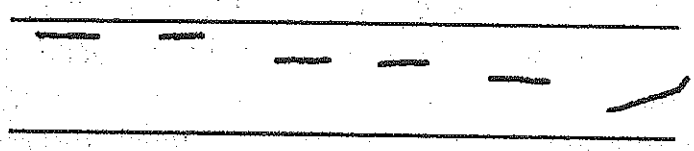
A rising or a falling tone is quite easy to identify, whether it falls on a single syllable or extends over more syllables in the case of a tonic syllable followed by a tail. Fall–rise and rise–fall tones, however, can be quite difficult to recognise when they are extended over tails, since their characteristic pitch movements are often broken up or distorted by the structure of the syllables they occur on. For example, the pitch movement on '˩some' will be something like this:

If we add a syllable, the "fall" part of the fall–rise is usually carried by the first syllable and the "rise" part by the second. The result may be a continuous pitch movement very similar to the one-syllable case, if there are no voiceless medial consonants to cause a break in the voicing. For example:

˩some ˈmen

3 The following sentences are given with intonation transcribed. Draw underneath them a diagram of the pitch movements, leaving a gap between each syllable. Example:

'would you 'like some 'more ˌmilk



a) 'Only when the ˌwind blows

b) ˌWhen did you say

c) 'What was the ˌname of the place

If the continuity of the voicing is broken, however, the pitch pattern might be more like this:

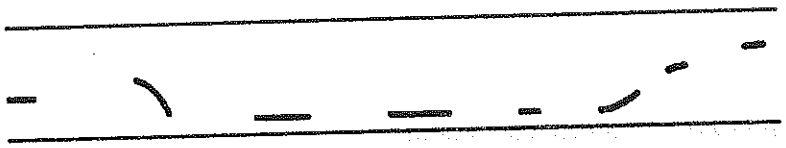
ˌsome ˈchairs

In this case it would be possible to say that there is a falling tone on 'some' and a rise on 'chairs'. However, most English speakers seem to feel that the pitch movement in this case is the same as that in the previous two examples; it can be said that there is a parallel with rhyming. Just as 'balloon' rhymes with 'moon', so we might say that 'ˌsome ˈchairs' has what could be called a **tonal rhyme** with 'ˌsome'.

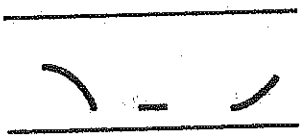
If there is a tail of two or more syllables, the normal pitch movement is for the pitch to fall on the tonic syllable and to remain low until the last stressed syllable in the tail. The pitch then rises from that point up to the end of the tone-unit. If there is *no* stressed syllable in the tail, the rise happens on the final syllable. Here are some examples:

i) I ˌmight ˈbuy it

I ˌmight have ˈthought of ˈbuying it

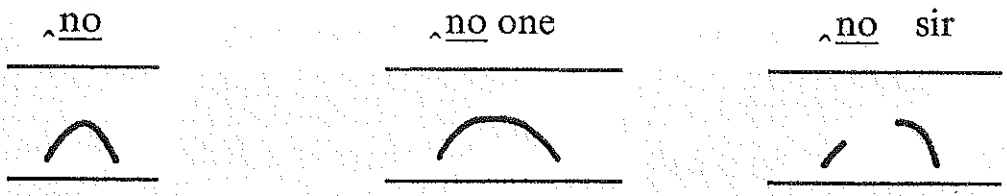


ii) ˌmost of them

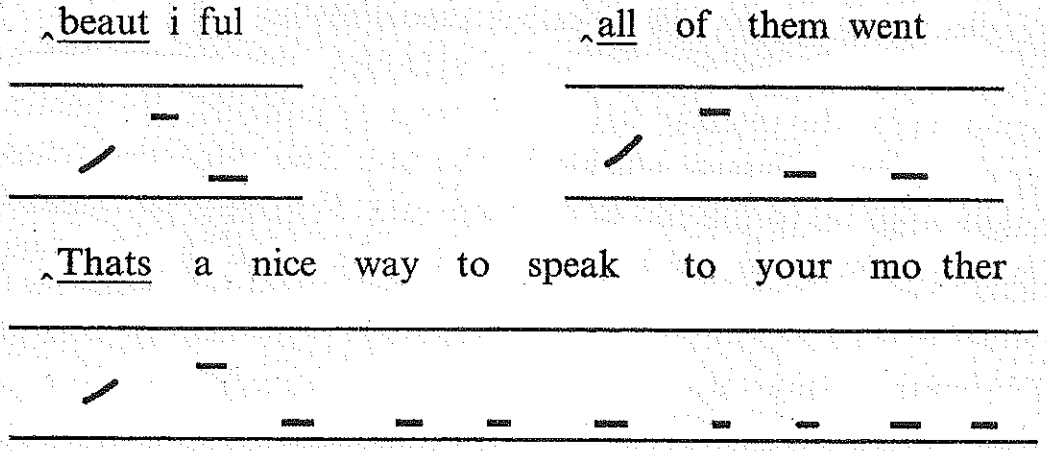


ˌmost of it was for them

With the rise–fall tone we find a similar situation: if the tonic syllable is followed by a single syllable in the tail, the "rise" part of the tone takes place on the first (tonic) syllable and the "fall" part is on the second. Thus:



When there are two or more syllables in the tail, the syllable immediately following the tonic syllable is always higher and any following syllables are low. For example:

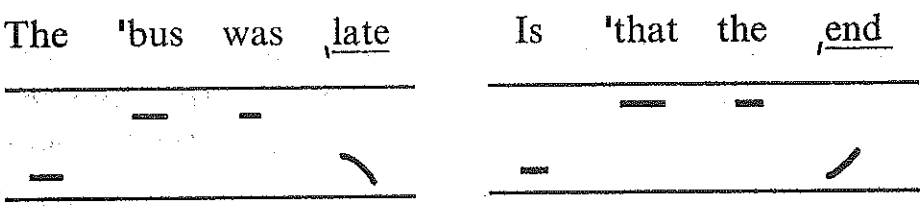


It should be clear by now that the speaker does not have a choice in the matter of the pitch of the syllables in the tail. This is completely determined by the choice of tone for the tonic syllable.

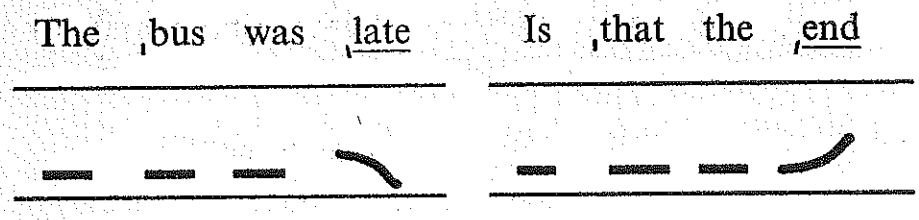
17.2 High and low heads

AU17, Ex 3

The head was defined in Chapter 16 as "all that part of a tone-unit that extends from the first stressed syllable up to, but not including, the tonic syllable". In our description of intonation up to this point, the only pitch contrasts found in the tone-unit are the different possible choices of tone for the tonic syllable. However, we can identify different pitch possibilities in the head, although these are limited to two which we will call **high head** and **low head**. In the case of the high head, the stressed syllable which begins the head is high in pitch; usually it is higher than the beginning pitch of the tone on the tonic syllable. For example:



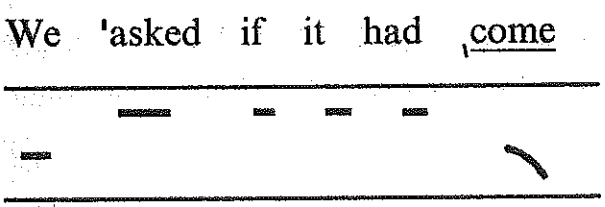
In the low head the stressed syllable which begins the head is low in pitch; usually it is lower than the beginning pitch of the tone on the tonic syllable. To mark this stressed syllable in the low head we will use a different symbol, ˌ as in 'ˌlow'. As an example, the heads of the above sentences will be changed from high to low:



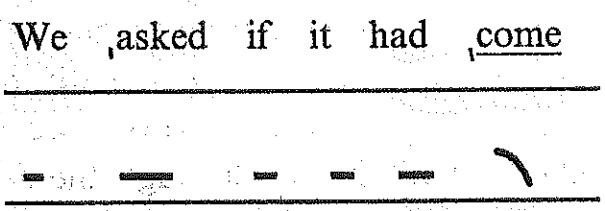
The two different versions (high and low head) will usually sound slightly different to English listeners, although it is not easy to say just what the difference *is*, as will be made clear in Chapter 18.

It is usual for unstressed syllables to continue the pitch of the stressed syllable that precedes them. In the following example, the three unstressed syllables 'if it had' continue at the same pitch as the stressed syllable 'asked'.

i) with high head



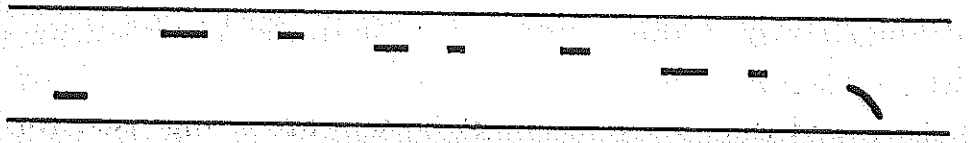
ii) with low head



When there is more than one stressed syllable in the head there is usually a slight change in pitch from the level of one stressed syllable to that of the next, the change being in the direction of the beginning

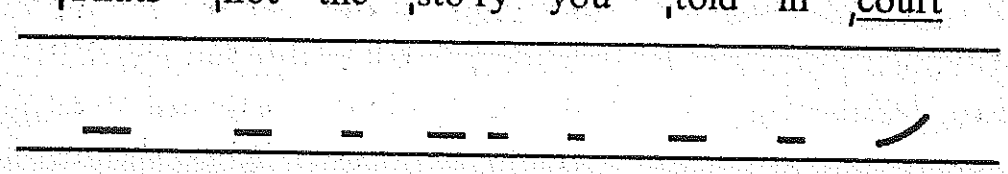
pitch of the tone on the tonic syllable. We will use some long examples to illustrate this, although heads of this length are not very frequently found in natural speech. In the first example the stressed syllables in the high head step downwards progressively to approach the beginning of the tone:

The 'rain was 'com ing 'down 'fair ly ˌhard



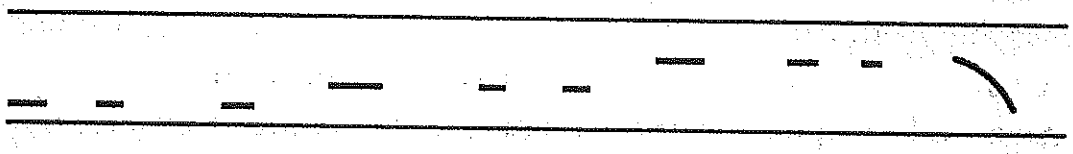
In the next example the head is low; since the tone also starts low, being a rise, there is no upward movement in the head:

ˌThats ˌnot the ˌsto ry you ˌtold in ˌcourt



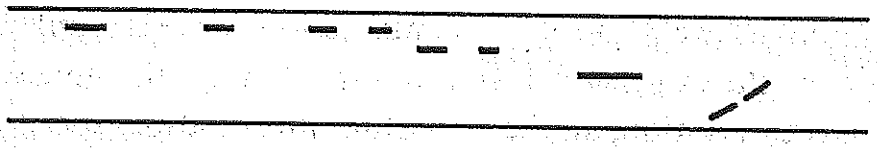
When there is a low head followed by a falling tone, successive stressed syllables in the head tend to move upwards towards the beginning pitch of the tone:

ˌI could have ˌbought it for ˌless than a ˌpound



When a high head is followed by a rise the stressed syllables tend to move downwards, as one would expect, towards the beginning pitch of the tone:

'Will there be a'nother 'train ˌlater



Of course, when we examine the intonation of polysyllabic heads we find much greater variety than these simple examples suggest. However, the division into high and low heads as general types is probably the most basic that can be made, and it would be pointless to set up a more elaborate system to represent differences if these

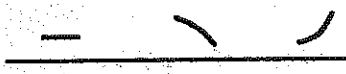
in many British studies of English intonation. There are certain difficulties that all of these studies have had to confront, and it is useful to give a brief summary of what the major difficulties are.

Identifying the tonic syllable

It is often said that the tonic syllable can be identified because it is the only syllable in the tone-unit that carries a movement in pitch; this is in fact not always true. We have seen how when the tonic syllable is followed by a tail the tone is carried by the tonic plus tail together in such a way that in some cases practically no pitch movement is detectable on the tonic syllable itself. In addition it has been claimed that one of the tones is the *level* tone, which by definition may not have any pitch movement. It is therefore necessary to say in this particular case that the tonic syllable is identified simply as the most prominent syllable.

In addition, it sometimes seems as if some tone-units (though only a small number) contain not one but two tonic syllables, almost always with the first syllable having a fall on it and the other a rise. An example is:

Ive ˌseen ˌhim

i) 

In this example there seems to be equal prominence on 'seen' and 'him'. Of course, it could be claimed that this is the same thing as:

Ive ˬseen him

ii) 

It has, however, been pointed out that the two versions are different in several ways. Since 'him' has greater prominence in (i), it cannot occur in its weak form **ɪm**, but must be pronounced **hɪm**, whereas in (ii) the pronunciation is likely to be **aɪv ˬsɪn ɪm**. The two versions are said to convey different meanings, too. Version (i) might be said in conversation on hearing someone's name, as in this example:

differences were not recognised by most English speakers. Some writers on intonation claim that the intonation pattern starting at a fairly high pitch, with a gradual dropping down of pitch during the utterance, is the most basic, normal, "unmarked" intonation pattern; this movement is often called **declination**. The claim that declination is universally unmarked in English, or even in all languages, is a strong one. As far as English is concerned, it would be good to see more evidence from the full range of regional and national varieties in support of the claim.

It should be noted that the two marks ˈ and ˌ are being used for two different purposes in this course, as they are in many phonetics books. When stress is being discussed, the ˈ mark indicates primary stress and ˌ indicates secondary stress. For the purposes of marking intonation, however, the mark ˈ indicates a stressed syllable in a high head and the mark ˌ indicates a stressed syllable in a low head. In practice this is not usually found confusing as long as one is aware of whether one is marking stress levels or intonation. When the high and low marks ˈ and ˌ are being used to indicate intonation it is no longer possible to mark two different levels of stress within the word. However, when looking at speech at the level of the tone-unit we are not usually interested in this; a much more important difference here is that between tonic stress (marked by underlining the tonic syllable and placing before it one of the five tone-marks) and non-tonic stressed syllables (marked ˈ or ˌ in the head or · in the tail).

It needs to be emphasised that in marking intonation, only stressed syllables are marked; this implies that intonation is carried entirely by the stressed syllables of a tone-unit and that the pitch of unstressed syllables is either predictable from that of stressed syllables or is of so little importance that it is not worth marking. Remember that the additional information given in the examples above by drawing pitch levels and movements between lines is only included here to make the examples clearer and is not normally given with our system of transcription; all the important information about intonation must, therefore, be given by the marks placed in the text.

17.3 Problems in analysing the form of intonation

The analysis of intonational form presented in this chapter and in Chapters 15 and 16 is similar in most respects to the approaches used

A: John Cleese is a very funny actor.
B: 'Oh ˌyes | Ive ˌseen ˌhim

In version (ii), on the other hand, the word 'seen' is given the greatest prominence, and it is likely to sound as though the speaker has some reservation, or has something further to say:

A: Have you seen my father yet?
B: Ive ˬseen him | but I 'havent had 'time to ˌtalk to him

The same is found with 'her', as in

Ive ˌseen ˌher
aɪv ˌsiːn ˌhɜː

compared with

Ive ˬseen her
aɪv ˬsiːn ə

This is a difficult problem, since it weakens the general claim made earlier that each tone-unit contains only one tonic syllable.

Identifying tone-unit boundaries

It is a generally accepted principle in the study of grammar that utterances may contain one or more sentences, and that one can identify on grammatical grounds the places where one sentence ends and another begins. In a similar way, in suprasegmental phonology it is claimed that utterances may be divided up into tone-units, and that one can identify on phonetic or phonological grounds the places where one tone-unit ends and another tone-unit begins. However, giving rules for determining where the boundaries are placed is not easy, except in cases where a clear pause separates tone-units. Two principles are usually mentioned: one is that it is possible in most cases to detect some sudden change from the pitch level at the end of one tone-unit to the pitch level that starts the following tone-unit, and recognition of the start of the following tone-unit is made easier by the fact that speakers tend to "return home" to a particular pitch level at the beginning of a tone-unit. The second principle used in tone-unit boundary identification is a rhythmical one: it is claimed that within the tone-unit, speech has a regular rhythm, but that rhythm is broken or interrupted at the tone-unit boundary. Both the

above principles are useful guides, but one regularly finds, in analysing natural speech, cases where it remains difficult or impossible to make a clear decision; the principles may well be factually correct, but it should be emphasised that at present there is no conclusive evidence from instrumental study in the laboratory that they are.

Anomalous tone-units

However comprehensive one's descriptive framework may be (and the one given in this course is very limited), there will inevitably be cases which do not fit within it. For example other tones such as fall–rise–fall or rise–fall–rise are occasionally found. In the head, we sometimes find cases where the stressed syllables are not all high or all low, as in the following example:

ˌAfter ˌone of the 'worst 'days of my ˬlife

It can also happen that a speaker is interrupted and leaves a tone-unit incomplete – for example, lacking a tonic syllable. To return to the analogy with grammar, in natural speech one often finds sentences which are grammatically anomalous or incomplete, but this does not deter the grammarian from describing "normal" sentence structure. Similarly, although there are inevitably problems and exceptions, we continue to treat the tone-unit as something that can be described, defined and recognised.

17.4 Autosegmental treatment of intonation

In recent years a rather different way of analysing intonation, sometimes referred to as **autosegmental**, has become quite widely used, especially in American work. In this approach, all intonational phenomena can be reduced to just two basic phonological elements: H (high tone) and L (low tone). A movement of pitch from high to low (a fall) is treated as the sequence HL. Individual stressed ("accented") syllables must all be marked as H or L, or with a combination marking a pitch movement. In addition to this process, H and L tones are associated with boundaries. A major tone-unit boundary (equivalent to what we have been marking with ||) is given the symbol %, but it must also be given a H or a L tone. Let us take an utterance like 'It's time to leave', which might be pronounced

its 'time to ˌleave (using our usual transcription)

In a simple version the alternative transcription will look like this:

H H L%
its time to leave

Instead of marking a falling tone on the word 'leave', the high-pitched part of the word is shown by the H and the low part by the L associated with the boundary %. There is another boundary (corresponding to the minor tone-unit boundary |) which is marked with –, and again this must be marked with either a H or a L. There must always be one of these boundaries marked before a % boundary. So, the following utterance would be transcribed like this in the system used in this book:

we ˌlooked at the ˌsky | and 'saw the ˌclouds

and in this way using autosegmental transcription:

L L H– H H L– L%
we looked at the sky and saw the clouds

How would this approach deal with complex tones spread over several syllables?

H L– H%
ˌmost of them could be transcribed most of them

Although this type of analysis has some attractions, especially in the way it fits with contemporary phonological theory, it seems unlikely that it would be more useful to learners of English than the traditional analysis presented here.

Notes on problems and further reading

The main concern of this chapter is to complete the description of intonational form, including analysis of perhaps the most difficult aspect: that of recognising fall–rise and rise–fall tones when they are extended over a number of syllables. This is necessary since no complete analysis of intonation can be done without having studied these "extended tones".

Cruttenden (1997: Chapters 3 and 4) gives a good introduction to the problems of analysing tones both within the traditional British framework and in autosegmental terms. A very detailed discussion of the difference between fall–rise and the "compound" fall-plus-rise is

given in Sharp (1958), although this is not easy reading and some of the examples are difficult to follow. On tone-unit boundaries, there is a clear explanation of the problems in Cruttenden (1997: Section 3.2), and in more detail in Crystal (1969: 204–7). The study of Scottish English by Brown *et al.* (1980) gives ample evidence that tone-units in real life are not as easy to identify as tone-units in textbooks. There has recently been a growth of interest in the comparative study of intonation in different languages and dialects: see Cruttenden (1997: Chapter 5); Hirst and di Cristo (1998); Ladd (1996: Chapter 4).

On declination, see Cruttenden (1997: 121–3).

For reading on autosegmental analysis (which is often also given the name **ToBI**), a good introduction is Cruttenden (1997: 56–67). A fuller and more critical analysis can be read in Ladd (1996: Chapters 2 and 3); see also Roca and Johnson (1999: Chapter 14). A short account of the problems found in trying to compare this approach with the traditional British analysis is given in Roach (1994).

Note for teachers

I would like to emphasise how valuable an exercise it is for students and teachers to attempt to analyse some recorded speech for themselves. For beginners it is best to start on slow, careful speech – such as that of newsreaders – before attempting conversational speech. One can learn more about intonation in an hour of this work than in days of reading textbooks on the subject, and one's interest in and understanding of theoretical problems becomes much more profound.

Written exercises

1 The following sentences are given with intonation marks. Sketch the pitch within the lines below, leaving a gap between each syllable.

a) 'Which was the ˌcheap one did you say

b) I 'only 'want to ˬtaste it

c) ˌShe would have ˌthought it was ˏobvious

d) There 'wasnt 'even a 'piece of ˌbread in the ·house

e) ˌNow will you be ·lieve me

2 This exercise is similar, but here you are given polysyllabic words and a tone. You must draw an appropriate pitch movement between the lines.

a) (rise) opportunity

b) (fall–rise) actually

c) (fall) confidently

d) (rise–fall) magnificent

e) (rise) relationship

f) (fall–rise) afternoon

18 Functions of intonation 1

The form of intonation has now been described in some detail, and we will move on to look more closely at its functions. Perhaps the best way to start is to ask ourselves what would be lost if we were to speak *without* intonation: you should try to imagine speech in which every syllable was said on the same level pitch, with no pauses and no changes in speed or loudness. This is the sort of speech that would be produced by a "mechanical speech" device (as described at the beginning of Chapter 14) that made sentences by putting together recordings of isolated words. To put it in the broadest possible terms, we can see that intonation makes it easier for a listener to understand what a speaker is trying to convey. The ways in which intonation does this are very complex, and many suggestions have been made for ways of isolating different functions. Among the most often proposed are the following:

i) Intonation enables us to express emotions and attitudes as we speak, and this adds a special kind of "meaning" to spoken language. This is often called the **attitudinal function** of intonation.

ii) Intonation helps to produce the effect of prominence on syllables that need to be perceived as stressed, and in particular the placing of tonic stress on a particular syllable marks out the word to which it belongs as the most important in the tone-unit. This has been called the **accentual function** of intonation.

iii) The listener is better able to recognise the grammar and syntactic structure of what is being said by using the information contained in the intonation; for example, such things as the placement of boundaries between phrases, clauses or sentences, the difference between questions and statements and the use of

grammatical subordination may be indicated. This has been called the **grammatical function** of intonation.

iv) Looking at the act of speaking in a broader way, we can see that intonation can signal to the listener what is to be taken as "new" information and what is already "given", can suggest when the speaker is indicating some sort of contrast or link with material in another tone-unit and, in conversation, can convey to the listener what kind of response is expected. Such functions are examples of intonation's **discourse function**.

The attitudinal function has been given so much importance in past work on intonation that it will be discussed separately in this chapter, although it should eventually become clear that it overlaps considerably with the discourse function. In the case of the other three functions, it will be argued that it is difficult to see how they could be treated as separate; for example, the placement of tonic stress is closely linked to the presentation of "new" information, while the question/statement distinction and the indication of contrast seem to be equally important in grammar and discourse. What seems to be common to accentual, grammatical and discourse functions is the indication, by means of intonation, of the relationship between some linguistic element and the context in which it occurs. The fact that they overlap with each other to a large degree is not so important if one does not insist on defining watertight boundaries between them.

The rest of this chapter is concerned with a critical examination of the attitudinal function.

18.1 The attitudinal function of intonation

Many writers have expressed the view that intonation is used to convey our feelings and attitudes; for example, the same sentence can be said in different ways, which might be labelled "angry", "happy", "grateful", "bored" and so on. It has also been widely observed that the form of intonation is different in different languages; for example, the intonation of languages such as Swedish, Italian or Hindi is instantly recognisable as being different from that of English. Not surprisingly, it has often been said that foreign learners of English need to learn English intonation. Some have gone further than this and claimed that, unless the foreign learner learns the

appropriate way to use intonation in a given situation, there is a risk that he or she may unintentionally give offence; for example, the learner might use an intonation suitable for expressing boredom or discontent when what is needed is an expression of gratitude or affection. This misleading view of intonation must have caused unnecessary anxiety to many learners of the language.

Let us begin by considering how one might analyse the attitudinal function of intonation. One possibility would be for the analyst to invent a large number of sentences and to try saying them with different intonation patterns (i.e. different combinations of head and tone), noting what attitude was supposed to correspond to the intonation in each case; of course, the results are then very subjective, and based on an artificial performance that has little resemblance to conversational speech. Alternatively, the analyst could say these different sentences to a group of listeners and ask them all to write down what attitudes they thought were being expressed; however, we have a vast range of adjectives available for labelling attitudes and the members of the group would probably produce a very large number of such adjectives, leaving the analyst with the problem of deciding whether pairs such as 'pompous' and 'stuck-up', or 'obsequious' and 'sycophantic' were synonyms or represented different attitudes. To overcome this difficulty, one could ask the members of the group to choose among a small number of adjectives (or "labels") given by the analyst; the results would then inevitably be easier to quantify (that is, the job of counting the different responses would be simpler) but the results would no longer represent the listeners' free choices of label. An alternative procedure would be to ask a lot of speakers to say a list of sentences in different ways according to labels provided by the analyst, and see what intonational features are found in common (for example, one might count how many speakers used a low head in saying something in a "hostile" way). The results of such experiments are usually very variable and difficult to interpret, not least because the range of acting talent in a randomly selected group is considerable.

A much more useful and realistic approach is to study recordings of different speakers' natural, spontaneous speech and try to make generalisations about attitudes and intonation on this basis. Many problems remain, however. In the method described previously, the

analyst tries to select sentences (or passages of some other size) whose meaning is fairly "neutral" from the emotional point of view, and will tend to avoid material such as 'Why don't you leave me alone?' or 'How can I ever thank you enough?' because the lexical meaning of the words used already makes the speaker's attitude pretty clear, whereas sentences such as 'She's going to buy it tomorrow' or 'The paper has fallen under the table' are less likely to prejudice the listener. The choice of material is much less free for someone studying natural speech. Nevertheless, if we are ever to make new discoveries about intonation, it will be as a result of studying what people actually say rather than inventing examples of what they *might* say.

The notion of "expressing an emotion or attitude" is itself a more complex one than is generally realised. First, an emotion may be expressed involuntarily or voluntarily; if I say something in a "happy" way, this may be because I *feel* happy, or because I want to convey to you the *impression* that I am happy. Second, an attitude that is expressed could be an attitude towards the listener (e.g. if I say something in a "friendly" way), towards what is being said (e.g. if I say something in a "sceptical" or "dubious" way) or towards some external event or situation (e.g. "regretful" or "disapproving").

However, one point is much more important and fundamental than all the problems discussed above. To understand this point you should imagine (or even actually perform) your pronunciation of a sentence in a number of different ways for example, if the sentence was 'I want to buy a new car' and you were to say it in the following ways: "pleading", "angry", "sad", "happy", "proud", it is certain that at least some of your performances will be different from some others, but it is also certain that the technique for analysing and transcribing intonation introduced earlier in the course will be found inadequate to represent the different things you do. You will have used variations in loudness and speed, for example; almost certainly you will have used different voice qualities for different attitudes. You may have used your pitch range (which was introduced in Section 15.3) in different ways: your pitch movements may have taken place within quite a narrow range (**narrow pitch range**) or using the full range between high and low (**wide pitch range**); if you did not use wide pitch range, you may have used

different **keys**: **high key** (using the upper part of your pitch range), **mid key** (using the middle part of the range) or **low key** (the lower part). It is very likely that you will have used different facial expressions and even gestures and body movements. These factors are all of great importance in conveying attitudes and emotions, yet the traditional handbooks on English pronunciation have almost completely ignored them.

If we accept the importance of these factors it becomes necessary to consider how they are related to intonation, and what intonation itself consists of. We can isolate three distinct types of suprasegmental variable: sequential, prosodic and paralinguistic.

Sequential

These components of intonation are found as elements in sequences of other such elements occurring one after another (never simultaneously). These are:

i) pre-heads, heads, tonic syllables and tails (with their pitch possibilities);
ii) pauses;
iii) tone-unit boundaries.

These have all been introduced in previous chapters.

Prosodic

These components are characteristics of speech which are constantly present and observable while speech is going on. The most important are:

i) width of pitch range;
ii) key;
iii) loudness;
iv) speed;
v) voice quality.

It is not possible to speak without one's speech having some degree or type of pitch range, loudness, speed and voice quality (with the possible exception that pitch factors are largely lost in whispered speech). Different speakers do, of course, have their own typical pitch range, loudness, voice quality, etc., and contrasts among

prosodic components should be seen as relative to these "background" speaker characteristics.

Each of these prosodic components needs a proper framework for categorisation, and this is an interesting area of current research. One example of the prosodic component "width of pitch range" has already been mentioned in Section 15.3, when "extra pitch height" was introduced, and the "rhythmicality" discussed in Section 14.1 could be regarded as another prosodic component. Prosodic components should be regarded as part of intonation along with sequential components.

Paralinguistic

Mention was made above of facial expressions, gestures and body movements. People who study human behaviour often use the term **body language** for such activity. One could also mention certain vocal effects such as laughs and sobs. These paralinguistic effects are obviously relevant to the act of speaking but could not themselves properly be regarded as components of speech. Again, they need a proper descriptive and classificatory system, but this is not something that comes within the scope of this course, nor in my opinion should they be regarded as components of intonation.

What advice, then, can be given to the foreign learner of English who wants to learn "correct intonation"? It is certainly true that a few generalisations can be made about the attitudinal functions of some components of intonation. Within tone, for example, most books agree on some basic meanings; here are some examples:

1. Fall
   Finality, definiteness: That is the end of the ˌnews
   I'm absolutely ˌcertain
   Stop ˌtalking

2. Rise
   Most of the functions attributed to rises are nearer to grammatical than attitudinal, as in the first three examples given below; they are included here mainly to give a fuller picture of intonational function.
   General questions: Can you ˌhelp me
   Is it ˌover

Listing: ˌRed ˌbrown ˌyellow or ˌblue
(fall is normal on the last item)
"More to follow": I phoned them right a ˌway
(and they agreed to come)
You must write it aˌgain
(and this time, get it right)
Encouraging: It wont ˌhurt

3. Fall–rise
Uncertainty, doubt: You ˬmay be right
Its ˬpossible
Requesting: Can I ˬbuy it
Will you ˬlend it to me

4. Rise–fall
Surprise, being impressed: You were ˆfirst
ˆAll of them

Generalisations such as these are, however, very broad, and foreign learners do not find it easy to learn to use intonation through studying them. Similarly, within the area of prosodic components most generalisations tend to be rather obvious: wider pitch range tends to be used in excited or enthusiastic speaking, slower speed is typical of the speech of someone who is tired or bored, and so on. Most of the generalisations one could make are probably true for a lot of other languages as well. In short, of the rules and generalisations that could be made about conveying attitudes through intonation, those which are not actually wrong are likely to be too trivial to be worth learning. I have witnessed many occasions when foreigners have unintentionally caused misunderstanding or even offence in speaking to a native English speaker, but can remember only a few occasions when this could be attributed to "using the wrong intonation"; most such cases have involved native speakers of different varieties of English, rather than learners of English. Sometimes an intonation mistake can cause a difference in apparent grammatical meaning (something that is dealt with in Chapter 19). It should not be concluded that intonation is not important for conveying attitudes. What is being claimed here is that, although it is of great importance, the complexity of the total set of sequential and prosodic components of intonation and of paralinguistic features

makes it a very difficult thing to teach or learn. One might compare the difficulty with that of trying to write rules for how one might indicate to someone that one finds him or her sexually attractive; while psychologists and biologists might make detailed observations and generalisations about how human beings of a particular culture behave in such a situation, most people would rightly feel that studying these generalisations would be no substitute for practical experience, and that relying on a textbook could lead to hilarious consequences. The attitudinal use of intonation is something that is best acquired through talking with and listening to English speakers, and this course aims simply to train learners to be more aware of and sensitive to the way English speakers use intonation.

Notes on problems and further reading

Perhaps the most controversial question concerning English intonation is what its function is; pedagogically speaking, this is a very important question, since one would not wish to devote time to teaching something without knowing what its value is likely to be. At the beginning of this chapter I list four commonly cited functions; it is possible to construct a longer list; Lee (1958), for example, proposed ten.

For general introductory reading on the functions of intonation, there is a good survey in Cruttenden (1997: Chapter 4). Critical views are expressed in Brazil *et al.* (1980: 98–103) and Crystal (1969: 282–308). There are many useful examples in Brazil (1994). Few people have carried out experiments on listeners' perception of attitudes through intonation, probably because it is extremely difficult to design properly controlled experiments.

Once one has recognised the importance of features other than pitch, it is necessary to devise a framework for categorising these features. There are many different views about the meaning of the term "paralinguistic". In the framework presented in Crystal and Quirk (1964), paralinguistic features of the "vocal effect" type are treated as part of intonation, and it is not made sufficiently clear how these are to be distinguished from prosodic features. Crystal (1969) defines paralinguistic features as: 'vocal effects which are primarily the result of physiological mechanisms other than the vocal cords, such as the direct results of the workings of the pharyngeal, oral or

nasal cavities' but this does not seem to me to fit the facts. In my view, "paralinguistic" implies 'outside the system of contrasts used in spoken language' – which does not, of course, necessarily mean 'non-vocal'. I would therefore treat prosodic variables as linguistic – and consequently part of intonation – while I would treat vocal effects like laughs or sobs as non-linguistic vocal effects to be classed with gestures and facial expressions. Brown (1990), on the other hand, uses "paralinguistic" to include what I call "prosodic", and appears to have no separate term for non-linguistic vocal effects. A recent paper on transcription of prosodie and paralinguistic features is Roach *et al.* (1998).

The term 'voice quality' needs comment, as it tends to be used with different meanings: sometimes the term is used to refer to the personal, "background" characteristics that make one person's voice recognisably different from another, mainly as a result of the complex interaction of laryngeal and supralaryngeal features (Crystal, 1969: 100–4; Laver, 1980; 1994); for some writers, however, "voice quality" is the auditory result of different types of vocal fold vibration. A better name for this is **phonation type** (Catford, 1964).

Note for teachers

Audio Unit 18 consists of extracts from a recording of spontaneous dialogue. Students usually feel that listening to these unfamiliar voices chopped up into small pieces is hard work, but generally the transcription exercise is not found nearly as difficult as expected.

Written exercises

In the following bits of conversation, you are supplied with an "opening line" and a response that you must imagine saying. You are given an indication in brackets of the feeling or attitude expressed, and you must mark on the text the intonation you think is appropriate (mark only the response). As usual in intonation work in this book, punctuation is left out, since it can cause confusion.

1	It looks nice for a ˌswim	Its rather cold (*doubtful*)
2	Why not get a ˌcar	Because I cant afford it (*impatient*)
3	Ive lost my ˌticket	Youre silly then (*stating the obvious*)

4	You cant have an ice ˌcream	Oh please (*pleading*)
5	What times are the ˌbuses	Seven o'clock seven thirty and eight (*listing*)
6	She got four ˌ'A' levels	Four (*impressed*)
7	How much ˌwork have you got to do	Ive got to do the shopping (*and more things after that*)
8	Will the ˬchildren go	Some of them might (*uncertain*)

19 Functions of intonation 2

In the previous chapter a distinction was made between the attitudinal function of intonation and several other functions that were given the collective name of syntagmatic functions. They include accentual, grammatical and discourse functions, and these are discussed below.

19.1 The accentual function of intonation

The term accentual is derived from "accent", a word used by some writers to refer to what in this course is called "stress". When writers say that intonation has accentual function they imply that the placement of stress is something that is determined by intonation. It is possible to argue against this view: in Chapters 10 and 11 word stress is presented as something quite independent of intonation, and subsequently (p. 176) it was said that 'intonation is carried entirely by the stressed syllables of a tone-unit'. This means that the presentation so far has implied that the placing of stress is independent of and prior to the choice of intonation. However, one particular aspect of stress *could* be regarded as part of intonation: this is the placement of the tonic stress within the tone-unit. It would be reasonable to suggest that while word stress is independent of intonation, the placement of tonic stress is a function (the accentual function) of intonation. Some older pronunciation handbooks refer to this area as "sentence stress", which is not an appropriate name: the sentence is a unit of grammar, while the location of tonic stress is a matter which concerns the tone unit, a unit of phonology.

The location of the tonic syllable is of considerable linguistic importance. The most common position for this is on the last lexical word (e.g. noun, adjective, verb, adverb as distinct from the

function words introduced in Chapter 12, pp. 112–13) of the tone-unit. For contrastive purposes, however, any word may become the bearer of the tonic syllable. It is frequently said that the placement of the tonic syllable indicates the **focus** of the information. In the following pairs of examples, (i) represents normal placement and (ii) contrastive:

i) | I ˌwant to ˌknow ˌwhere hes ˌtravelling to |
(The word 'to', being a preposition and not a lexical word, is not stressed.)
ii) (I 'dont want to 'know where hes 'travelling ˎfrom)
| I ˌwant to ˌknow ˌ where hes ˌ travelling ˌto |

i) | She was 'wearing a 'red ˌdress |
ii) (She 'wasnt 'wearing a ˎgreen ·dress) | She was 'wearing a ˌred ·dress |

Similarly, for the purpose of emphasis we may place the tonic stress in other positions; in these examples, (i) is non-emphatic and (ii) is emphatic:

i) | It was 'very ˌboring |
ii) | It was ˌvery ·boring |

i) | You 'mustnt 'talk so ˌloudly |
ii) | You ˌmustnt ·talk so ·loudly |

However, it would be wrong to say that the only cases of departure from putting tonic stress on the last lexical word were cases of contrast or emphasis. There are quite a few situations where it is normal for the tonic syllable to come earlier in the tone-unit. A well-known example is the sentence 'I have plans to leave'; this is ambiguous:

i) | I have 'plans to ˌleave |
(i.e. I am planning to leave.)
ii) | I have ˌplans to ·leave |
(i.e. I have some plans/diagrams/drawings that I have to leave.)

Version (ii) could not be described as contrastive or emphatic. There are many examples similar to (ii); perhaps the best rule to give is that the tonic syllable will *tend* to occur on the last lexical word in the

tone-unit, but may be placed earlier in the tone-unit if there is a word there with greater importance to what is being said. This can quite often happen as a result of the last part of the tone-unit being already "given" (i.e. something which has already been mentioned or is completely predictable); for example:

i) | 'Heres that ˌbook you ·asked me to ·bring |
(The fact that you asked me to bring it is not new.)
ii) | Ive ˌgot to ˌtake the ˌdog for a ·walk |
('For a walk' is by far the most probable thing to follow 'I've got to take the dog'; if the sentence ended with 'to the vet' the tonic syllable would probably be 'vet'.)

Placement of tonic stress is, therefore, important and is closely linked to intonation. A question that remains, however, is whether one can and should treat this matter as separate from the other functions described below.

19.2 The grammatical function of intonation

The word "grammatical" tends to be used in a very loose sense in this context. It is usual to illustrate the grammatical function by inventing sentences which when written are ambiguous, and whose ambiguity can only be removed by using differences of intonation. A typical example is the sentence 'Those who sold quickly made a profit'. This can be said in at least two different ways:

i) | 'Those who 'sold ˬquickly | ˌmade a ˌprofit
ii) | 'Those who ˬsold | ˌquickly ˌmade a ˌprofit |

The difference caused by the placement of the tone-unit boundary is seen to be equivalent to giving two different paraphrases of the sentences, as in:

i) A profit was made by those who sold quickly.
ii) A profit was quickly made by those who sold.

Let us look further at the role of tone-unit boundaries, and the link between the tone-unit and units of grammar. There is a strong tendency for tone-unit boundaries to occur at boundaries between grammatical units of higher order than words; it is extremely common to find a tone-unit boundary at a sentence boundary, as in:

| I 'wont have any ˌtea | I 'dont ˌlike it |

In sentences with a more complex structure, tone-unit boundaries are often found at phrase and clause boundaries as well, as in:

| In ˏFrance | where ˌfarms ˌtend to be ˏsmaller | the 'subsidies are 'more imˌportant

It is very unusual to find a tone-unit boundary at a place where the only grammatical boundary is a boundary between words. It would, for example, sound distinctly odd to have a tone-unit boundary between an article and a following noun, or between auxiliary and main verbs if they are adjacent (although we may, on occasions, hesitate or pause in such places within a tone-unit; it is interesting to note that some people who do a lot of arguing and debating, notably politicians and philosophers, develop the skill of pausing for breath in such intonationally unlikely places because they are less likely to be interrupted than if they pause at the end of a sentence). Tone-unit boundary placement can, then, indicate grammatical structure to the listener and we can find minimal pairs such as the following:

i) The Con'servatives who ˏlike the pro ˑposal | are ˌpleased
ii) The Con ˏservatives | who ˏlike the pro ˑposal | are ˌpleased

The intonation makes clear the difference between (i) "restrictive" and (ii) "non-restrictive" relative clauses; (i) implies that only *some* Conservatives like the proposal, while (ii) implies that *all* the Conservatives like it.

Another component of intonation that can be said to have grammatical significance is the choice of tone on the tonic syllable. One example that is very familiar is the use of a rising tone with questions. Many languages have the possibility of changing a statement into a question simply by changing the tone from falling to rising. This is, in fact, not used very much by itself in the variety of English being described here, where questions are usually grammatically marked. The sentence 'The price is going up' can be said as a statement like this:

| The ˌprice is going up |

(the tonic stress could equally well be on 'up'). It would be quite

acceptable in some dialects of English (e.g. many varieties of American English) to ask a question like this:

(Why do you want to buy it now?) | The ˌprice is going up |

But speakers in Britain would be more likely to ask the question like this:

(Why do you want to buy it now?) | 'Is the ˌprice going up |

It is by no means true that a rising tone is always used for questions in English; it is quite usual, for example, to use a falling tone with questions beginning with one of the "wh-question-words" like 'what', 'which', 'when', etc. Here are two examples with typical intonations, where (i) does not start with a "wh-word" and has a rising tone and (ii) begins with 'where' and has a falling tone.

i) | 'Did you 'park the ˌcar |
ii) | 'Where did you 'park the ˌcar |

However, the fall in (ii) is certainly not obligatory, and a rise is quite often heard in such a question. A fall is also possible in (i).

The intonation of **question-tags** (e.g. 'isn't it', 'can't he', 'should she', 'won't they', etc.) is often quoted as a case of a difference in meaning being due to the difference between falling and rising tone. In the following example, the question-tag is 'aren't they'; when it has a falling tone, as in (i), the implication is said to be that the speaker is comparatively certain that the information is correct, and simply expects the listener to provide confirmation, while the rising tone in (ii) is said to indicate a lesser degree of certainty, so that the question-tag functions more like a request for information.

i) | They 'are 'coming on ˌTuesday | ˌarent they |
ii) | They 'are 'coming on ˌTuesday | ˌarent they |

The difference illustrated here could reasonably be said to be as much attitudinal as grammatical. Certainly there is overlap between these two functions.

19.3 The discourse function of intonation

If we think of linguistic analysis as usually being linked to the sentence as the maximum unit of grammar, then the study of

discourse attempts to look at the larger contexts in which sentences occur. For example, consider the four sentences in the following:

A: Have you got any free time this morning?
B: I might have later on if that meeting's off.
A: They were talking about putting it later.
B: You can't be sure.

Each sentence could be studied in isolation and be analysed in terms of grammatical construction, lexical content and so on. But it is obvious that the sentences form part of some larger act of conversational interaction between two speakers; the sentences contain several references that presuppose shared knowledge (e.g. 'that meeting' implies that both speakers know which meeting is being spoken about), and in some cases the meaning of a sentence can only be correctly interpreted in the light of knowledge of what has preceded it in the conversation (e.g. 'You can't be sure').

If we consider how intonation may be studied in relation to discourse, we can identify two main areas: one of them is the use of intonation to focus the listener's attention on aspects of the message that are most important, and the other is concerned with the regulation of conversational behaviour. We will look at these in turn.

In the case of "attention focusing", the most obvious use has already been described: this is the placing of tonic stress on the appropriate syllable of one particular word in the tone-unit. In many cases it is easy to demonstrate that the tonic stress is placed on the word that is in some sense the "most important", as in:

| She 'went to ˌScotland |

Sometimes it seems more appropriate to describe tonic stress placement in terms of "information content": the more predictable a word's occurrence is in a given context, the lower its information content is. Tonic stress will tend to be placed on words with high information content, as suggested above when the term *focus* was introduced. This is the explanation that would be used in the case of the sentences suggested in Section 9.1:

i) | Ive ˌgot to ˌtake the ˌdog for a ·walk |
ii) | Ive ˌgot to ˌtake the ˌdog to the ˌvet |

The word 'vet' is less predictable (has a higher information content) than 'walk'. However, we still find many cases where it is difficult to explain tonic placement in terms of "importance" or "information". For example, in messages like:

Your coat's on fire	The wing's breaking up
The radio's gone wrong	Your uncle's died

probably the majority of English speakers would place the tonic stress on the subject noun, although it is difficult to see how this is more important than the last lexical word in each of the sentences. The placement of tonic stress is still to some extent an unsolved mystery; it is clear, however, that it is at least partly determined by the larger context (linguistic and non-linguistic) in which the tone-unit occurs.

We can see at least two other ways in which intonation can assist in focusing attention. The tone chosen can indicate whether the tone-unit in which it occurs is being used to present new information or to refer to information which is felt to be already possessed by speaker and hearer. For example, in the following sentence:

| 'Since the ˎlast time we ·met | when we had that 'huge ˎdinner |
Ive 'been on a ˌdiet |

the first two tone-units present information which is relevant to what the speaker is saying, but which is not something new and unknown to the listener. The final tone-unit, however, does present new information. Writers on discourse intonation have proposed that the falling tone indicates new information while rising (including falling–rising) tones indicate "shared" or "given" information.

Another use of intonation connected with the focusing of attention is **intonational subordination**; we can signal that a particular tone-unit is of comparatively low importance and as a result give correspondingly greater importance to adjacent tone-units. For example:

i) | As I ex ˌpect youve ˌheard | theyre 'only ad 'mitting eˌmergency ·cases |
ii) | The ˌJapa ˎnese | for ˌsome ˌreason or ˌother | 'drive on the ˌleft | like ˌus |

In a typical conversational pronunciation of these sentences, the first tone-unit of (i) and the second and fourth tone-units of (ii) might be treated as intonationally subordinate; the prosodic characteristics marking this are usually:

i) a drop to a lower part of the pitch range ("low key");
ii) increased speed;
iii) narrower range of pitch; and
iv) lower loudness, relative to the non-subordinate tone-unit(s).

The use of these components has the result that the subordinate tone-units are less easy to hear. Native speakers can usually still understand what is said, if necessary by guessing at inaudible or unrecognisable words on the basis of their knowledge of what the speaker is talking about. Foreign learners of English, on the other hand, having in general less "common ground" or shared knowledge with the speaker, often find that these subordinate tone-units – with their "throw-away", parenthetic style – cause serious difficulties in understanding.

We now turn to the second main area of intonational discourse function: the regulation of conversational behaviour. We have already seen how the study of sequences of tone-units in the speech of one speaker can reveal information carried by intonation which would not have been recognised if intonation were analysed only at the level of individual tone-units. Intonation is also important in the conversational interaction of two or more speakers. Most of the research on this has been on conversational interaction of a rather restricted kind – such as between doctor and patient, teacher and pupil or between the various speakers in court cases. In such material it is comparatively easy to identify what each speaker is actually doing in speaking – for example, questioning, challenging, advising, encouraging, disapproving, etc. It is likely that other forms of conversation can be analysed in the same way, although this is considerably more difficult. In a more general way, it can be seen that speakers use various prosodic components to indicate to others

that they have finished speaking, that another person is expected to speak, that a particular type of response is required and so on. A very familiar example is that quoted above (p. 197), where the difference between falling and rising intonation on question-tags is supposed to indicate to the listener what sort of response is expected. It seems that key (the part of the pitch range used) is important in signalling information about conversational interaction. We can observe many examples in non-linguistic behaviour of the use of signals to regulate turn-taking: in many sports, for example, it is necessary to do this – footballers can indicate that they are looking for someone to pass the ball to, or that they are ready to receive the ball, and doubles partners in tennis can indicate to each other who is to play a shot. Intonation, in conjunction with "body language" such as eye contact, facial expression, gestures and head-turning, is used for similar purposes in speech, as well as for establishing or confirming the status of the participants in a conversation.

19.4 Conclusions

It seems clear that studying intonation in relation to discourse makes it possible to explain much more comprehensively the uses that speakers make of intonation. Practically all the separate functions traditionally attributed to intonation (attitudinal, accentual and grammatical) could be seen as different aspects of discourse function. The risk, with such a broad approach, is that one might end up making generalisations that were too broad and had little power to predict with any accuracy the intonation that a speaker would use in a particular context. It is still too early to say how useful the discourse approach will be, but even if it achieves nothing else, it can at least be claimed to have shown the inadequacy of attempting to analyse the function of intonation on the basis of isolated sentences or tone-units, removed from their linguistic and situational context.

Notes on problems and further reading

Important work was done on the placement of tonic stress by Halliday (1967); his term for this is 'tonicity', and he adopts the widely-used linguistic term 'marked' for tonicity that deviates from

what I have called (for the sake of simplicity) 'normal'. Within generative phonology there has been much debate about whether one can determine the placing of tonic ('primary') stress without referring to the non-linguistic context in which the speaker says something. This debate was very active in the 1970s, well summarised and criticised in Schmerling (1976), but see Bolinger (1972). For more recent accounts, see Couper-Kuhlen (1986: Chapters 7 and 8) and Ladd (1996: 221–35).

One of the most interesting developments of recent years has been the emergence of a theory of discourse intonation. Readers unfamiliar with the study of discourse may find some initial difficulty in understanding the principles involved; the best introduction is Brazil *et al.* (1980), while the ideas set out there are given more practical expression in Brazil (1994). I have not been able to do more than suggest the rough outline of this approach.

The treatment of intonational subordination is based not on the work of Brazil but on Crystal and Quirk (1964: 52–6) and Crystal (1969: 235–52). The basic philosophy is the same, however, in that both views illustrate the fact that there is in intonation some organisation at a level higher than the isolated tone-unit; this was pointed out in the discussion of Trim (1959; see the notes on Chapter 16 above); see also Fox (1973). A parallel might be drawn with the relationship between the sentence and the paragraph in writing. It seems likely that a considerable amount of valuable new research on pronunciation will grow out of the study of discourse.

Note for teachers

The comment about Audio Unit 18 at the end of Chapter 18 applies also to Audio Unit 19: at first hearing it seems very difficult, but when worked on step by step it is far from impossible. In fact, although this passage sounds rapid and colloquial it is still easier to analyse than a full-speed conversational interchange.

Written exercises

1 In the following exercise, read the "opening line" and then decide the most suitable place for tonic stress placement (underline the syllable) in the response.

a)	I'd like you to ˌhelp me	(right) can I do the shopping for you
b)	I hear you're offering to do the ˌshopping for someone	(right) can I do the shopping for you
c)	What was the first thing that ˌhappened	first the professor explained her theory
d)	Was the theory explained by ˌstudents	no first the professor explained her theory
e)	Tell me how the ˌtheory was presented	first she explained her theory
f)	I think it starts at ten to ˌthree	no ten past three
g)	I think it starts at quarter past ˌthree	no ten past three
h)	I think it starts at ten past ˌfour	no ten past three

2 The following sentences are given without punctuation. Underline the appropriate tonic syllable places and mark tone-unit boundaries where you think they are appropriate.

a) (*he wrote the letter in a sad way*) he wrote the letter sadly
b) (*it's regrettable that he wrote the letter*) he wrote the letter sadly
c) four plus six divided by two equals five
d) four plus six divided by two equals seven
e) we broke one thing after another fell down
f) we broke one thing after another that night