SECOND EDITION

PHILIP CARR

ENGLISH PHONETICS AND PHONOLOGY

AN INTRODUCTION

WILEY-BLACKWELL

Contents

English Phonetics and Phonology

An Introduction

Second Edition

Philip Carr

WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

This second edition first published 2013

Edition History: Blackwell Publishing Ltd (1e, 1999)

Blackwell Publishing was acquired by John Wiley & Sons in February 2007. Blackwell's publishing program has been merged with Wiley's global Scientific, Technical, and Medical business to form Wiley-Blackwell.

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

350 Main Street, Malden, MA 02148-5020, USA
9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, for customer services, and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Carr, Philip, 1953–
English phonetics and phonology : an introduction / Philip Carr. — Second edition.
pages cm
Includes bibliographical references and index.
ISBN 978-1-4051-3454-5
1. English language—Phonology. 2. English language—Phonetics.
PE1133.C34 2013
421′.5—dc23
2012005351

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

Cover image: Patrick Henry Bruce, , c.

1925–6. Corcoran Gallery of Art, Washington, D.C. Photo © SuperStock.
Cover design by Simon Levy.
1 2013

Sound Recordings

These sound files accompany exercises, the treatment of intonation, and the description of some of the varieties of English given here. They are marked in the margins with a headphones symbol (as shown), and are available at: www.wiley.com/go/carrphonetics

Listen to sound files online

Track 1.1: Exercise 4
Track 1.2: Exercise 5
Track 1.3: Exercise 6

Track 2.1: Exercise 1
Track 2.2: Exercise 2
Track 2.3: Exercise 3

Track 3.1: Exercise 3

Track 4.1: Exercise 3

Track 5.1: Exercise 4

Track 6.1: (vowel neutralization in GA)
Track 6.2: Exercise 3

Track 13.10: Exercise 7 (RP)
Track 13.11: Exercise 7 (GA)

Prefaces to the First Edition

Preface for Teachers

Each year in the Department of English at Newcastle University, I am given eleven 50-minute lecture slots in which to introduce English phonetics and phonology to around a hundred students in the first semester of their first year on a variety of different undergraduate degree programmes, including English language and literature, linguistics, English language, modern languages, music, history and many others. Also included in the student body are European exchange undergraduates and students taking applied linguistics postgraduate degrees in media technology and in linguistics for teachers of English as a second language.

Given the range of degree types, this is a daunting task, made even more difficult by the fact that a substantial minority of the students do not have English as their first language. In a typical year, the student cohort will include speakers of Arabic, French, Spanish, German, Greek, Japanese, Korean, Mandarin or Cantonese Chinese, and Thai. Many of the non-native speakers will have been taught RP; others will have been taught General American. Amongst the native speakers of English, very few of the students will be speakers of RP, so that the non-native speakers are more likely to speak RP than the native speakers.

The vast majority of the student body will take their study of English phonetics and phonology no further, and the one factor which the majority of this diverse band of students shares is that they have no previous knowledge of phonetics or phonology; the course must therefore be .

One faces a dilemma in teaching such a course: on the one hand, one wants to cater to the small minority who will go on to study phonology at a more advanced level. On the other hand, one wants to introduce the subject without overwhelming the students with a mass of bewildering descriptive detail and an avalanche of seemingly arcane theoretical constructs. It is a moot point whether this dilemma can be resolved. However, this textbook was written as an attempt at a solution.

It is arguable that textbooks are harder to write than monographs, and that the more elementary the textbook, the harder it is to write: one can barely write a line without being aware of one's often questionable assumptions, and one has always to resist the temptation to question them in the body of the text. One continually has the sense of one's peers looking over one's shoulder and guffawing at the absurd oversimplifications which one is knowingly committing to print. But it has to be done: students have to learn to walk before they can learn to run; they also have to learn to crawl before they can learn to walk.

Writing and using textbooks is an empirical matter: it is very often immediately apparent when an exercise, chapter or book is simply not working, for a given body of students. Almost all of the textbooks which I have used on the first-year Newcastle course described here have proved to be unsuitable for this type of student cohort in one way or another; mostly, they have contained far too much detail. I have therefore set out to write a very short, very simple coursebook which deliberately ignores a great many descriptive/theoretical complexities.

My aim has not been to introduce students to phonological theory; rather, I have sought to introduce some of the bare essentials of English phonetics and phonology in a manner that is as theory-neutral as possible. This is fundamentally problematic, of course, since there is no such thing as theory-neutral description. I have therefore decided to adopt various theoretical/descriptive views, such as the tongue-arch/cardinal vowel approach to articulatory description, the phonemic approach to segmental phonology, the trochaic approach to English foot structure, and so on, on the purely pragmatic basis of what I have found to be easiest to convey to the students.

I have ignored acoustic phonetics for the very simple reason that our department lacks a phonetics lab, and I have not included distinctive features, since the mere sight of arrays of features marked with '+' and '−' symbols seems to render large numbers of my first-year students dizzy (particularly those majoring in English literature). I have also excluded feature geometry, the mora, under-specification and a great many other theoretical/descriptive notions, in an attempt to pare the subject down to a bare minimum of these.

The first four chapters are deliberately very short indeed, and contain only the most elementary introduction to articulatory phonetics. My aim there is to offer the student a gentle introduction to the course. I have spread the introduction of the phonemic principle over two chapters, since, in my experience, students find their first encounter with these ideas something of a quantum leap. The chapters on word stress, rhythm, connected speech phenomena and accent variation contain a very stripped-down, minimal, account of those subjects; I hope that there is enough there to act as a foundation for those students who wish to study these matters in more depth. In the chapter on syllable structure, I have been a little more ambitious in introducing analytical complexity, on the assumption that syllable structure is something that beginning students seem to be able to get the hang of more easily than, say, rhythm or intonation.

I believe that one of the most important duties of a university teacher is to induce in the student a sense of critical awareness, a grasp of argumentation and the role of evidence. On the other hand, one has to be very wary of introducing students at the most elementary stage to the idea of competing analyses: they find it difficult enough to get the hang of one sort of analysis, without being asked to assess the merits and demerits of competing analyses (even at the post-elementary stage, most undergraduates are very resistant to the idea of critically comparing different analyses). I have tried to overcome this dilemma by introducing competing analyses and assumptions at one or two points, while consciously ignoring them elsewhere.

The exercises are meant to be discussed at weekly seminar/tutorial meetings; my experience is that, if phonetics/phonology students are not made to do exercises, they easily come to believe that they have grasped the subject when in fact they have not. It is my hope that students who have completed this course would find it possible to tackle more advanced textbook treatments of these topics, such as those given by Giegerich (1992) and Spencer (1996). Whether that hope is fulfilled is, of course, very much an empirical matter.

Preface for Students

This is an elementary introduction to English phonetics and phonology, designed for those who have no previous knowledge whatsoever of the subject. It begins with a very elementary introduction to articulatory phonetics, and then proceeds to introduce the student to a very simplified account of some of the main aspects of the phonological structure of present-day English.

It is arguable that there are two main questions one might ask in studying the English language: what is it about English that makes it a language (as opposed to, say, a non-human communication system), and what is it about English that makes it English (as opposed to, say, French or Korean)? This book attempts to provide the beginnings of an answer to both of those questions, with respect to one aspect of English: its phonology.

Thus, although the subject matter of this book is English, there is reference to the phonology of other languages at several points, often in contrastive exercises which are designed to bring out one or more differences between English and another language. These contrastive exercises are included because native speakers of English, who often have little or no detailed knowledge of other languages, tend to assume that the phonology of English is the way it is as a matter of natural fact, a matter of necessity. For many such speakers, it will seem somehow natural, for instance, that the presence of the sound [f] as opposed to [v] functions to signal a difference in meaning (as in vs). To the English speaker, [f] and [v] will therefore seem easily distinguishable, and that too will appear to be a natural fact. But the fact that these sounds have that function in English is a conventional, not a necessary or natural fact: English need not have been that way, and may not always be that way. Just as one can gain a new perspective on one's own culture by learning about other cultures, so one can gain a fresh perspective on one's native language by learning a little about other languages. One can also, in learning about other cultures, gain some sense of what human cultures are like. Similarly, one can begin to get a sense of what human language phonologies are like by learning in what respects they resemble each other. Those points of resemblance concern general organizational properties of human language phonologies, such as the phonemic principle and the principles of syllable structure.

Reading a textbook on linguistic analysis is not like reading a novel. It is vital that the student complete the exercises at the end of each chapter before proceeding to the next chapter: they are designed to get the student to apply the ideas introduced in the chapter. The reader will not have properly grasped the ideas contained in this, or any other, textbook on phonology by simply sitting back in an armchair and reading the text, even if the student is under the impression of having understood the ideas. Vast numbers of students who have attempted to master linguistic analysis without actually doing it have ended up with disastrous exam results: no one ever became any good at linguistic analysis without actually doing it.

Like most linguistics textbooks, this book is cumulative in nature: what has been introduced in earlier chapters is presupposed in later chapters. It is fatal, therefore, to let several weeks go by without doing the reading and the exercises, in the hope of catching up later: the result is very likely to be that you will simply find yourself out of your depth, even though this is an elementary textbook. It is simply not possible to dip in and out of a linguistic analysis textbook, no matter how basic, in the way that one might dip in and out of a dictionary or an encyclopedia.

This book is designed to cater for students who, in all probability, will not pursue their studies in English phonetics and phonology any further. However, students who will be proceeding to a more advanced level should be able to tackle more advanced textbook treatments of these topics, such as those given by Giegerich and by Spencer (see Suggested Further Reading at the end of the book). Those students should also find it easier to tackle one of the many introductions to general phonological theory which are not focused on English (again, see Suggested Further Reading). In order to prepare such students for more advanced study, I have introduced, at some points, an indication of some of the difficulties with some of the assumptions made in this textbook, or a brief discussion of competing analyses. Although this textbook merely scratches the surface of the subject matter, I hope that there is enough here to make the subject of phonology seem intriguing to the student who intends to pursue his or her studies.

It is my hope that this book will be of some use to teachers of English as a foreign language, although it is not designed specifically for such readers. I am always surprised to discover how little in the way of knowledge of English phonetics and phonology such teachers often have. I have no experience of such teaching, and while I make no suggestions as to how the notions introduced in this book might be put to use in the TEFL classroom, I find it hard to believe that a knowledge of the basics of English phonetics and phonology could fail to be useful to the TEFL teacher in some way, even if only as background knowledge which extends the teacher's knowledge of English. I also hope that some of the contrastive exercises might help suggest ways in which one's native language phonology can interfere with one's attempt to acquire English as a second language.

Newcastle, February 1999

Preface to the Second Edition

The first edition of this book was written while I was teaching in an English university. Since then, I have moved to the English department at Montpellier University, in France. While I always had non-native speakers of English in my classes at Newcastle University, most of my students were native speakers of English; now, the vast majority of my students are not native speakers of English. Most are French, but there are also Spanish, Portuguese, Greek, German, Dutch, Polish, Russian and Bulgarian students, among others. The book has changed as a result: it is more orientated towards learners of English as a foreign language, but it is still useful for native speakers, I believe.

The main changes to the text concern the later chapters: chapters 8, 9 and 10 have been entirely rewritten, and there is a new chapter (chapter 11) on the relationship between spelling and pronunciation, known as grapho-phonemics. Teachers whose students are native speakers of English may choose to skip this chapter, but it could prove useful for students who wish to go on to teach English as a foreign language. I have expanded the appendix (renamed as chapter 13) to cover additional varieties of English. There are now sound files which accompany exercises, the treatment of intonation, and the description of some of the varieties of English given here: these are marked in the margins with a headphones symbol.

I have insisted on retaining practice at phonetic transcription, for two reasons. Firstly, I believe that it reinforces the distinction between phonetic transcription, based on listening to speech sounds, and phonological analysis, in which phonemes (as conceived of here) are not speech sounds, and cannot be heard. Secondly, I hope that some readers of this book will go on to engage in the empirical study of varieties of English, which typically involves both listening carefully to, and phonetically transcribing, recordings of speakers of various accents, and also engaging with theoretical issues in the analysis of those accents. The phonetic transcription exercises are now based on audio recordings.

The book is not intended as an introduction to phonological theory; some books of that sort are listed in the Suggested Further Reading. Inevitably, I have had to draw on notions proposed in various theoretical frameworks. Any proposed distinction between theory and description is fraught with difficulties: there can be no description without theoretical assumptions, as the philosopher of science Karl Popper pointed out. However, in my view, some kind of distinction between theory and description must be upheld. My aims here are primarily descriptive.

Any queries and/or corrections can be sent to: philip.carr@univ-montp3.fr

Montpellier, December 2011

Acknowledgements

Acknowledgements, First Edition

I am grateful to several cohorts of students in the Department of English Literary and Linguistic Studies at Newcastle University whose feedback has been valuable. I am also indebted to Patrick Honeybone, Maria Maza, Irenie Rowley and Charles Prescott, who acted as tutors to my students, and whose comments on several drafts have proved most helpful. Thanks too to my colleague Karen Corrigan, who commented on an early draft. Many years ago, in a small mud house on an island on the White Nile, I introduced James Dickins to the elements of phonetic description. He has kindly reciprocated by supplying me with the Arabic data on p. 40, for which I thank him. I also owe a debt of gratitude to the seven anonymous reviewers of my original proposal for their comments, and to Steve Smith, Mary Riso and Beth Remmes at Blackwell for their patience and encouragement. They may well be surprised that it took so long to write such a short book; I can, however, console myself with the fact that I do not have to offer the (perhaps apocryphal) apology: 'I am sorry this book is so long; I did not have the time to write a shorter one.' Finally, many thanks to Blackwell's reader, Andrew Spencer, whose intelligent, informed and insightful comments on the pre-final draft were immensely helpful. He is not, of course, responsible for any remaining inadequacies in the text.

Acknowledgements, Second Edition

I am grateful to Jacques Durand and Anne Przewozny, fellow co-directors of the PAC project (La Phonologie de l'anglais contemporain), and to Raphaël Dommange for permission to use some of their PAC recordings.

Thanks to my ex-Masters student Cécile Montforte for allowing me to use some of her PAC data from her dissertation 'Accent change and language attitudes in Scotland: a socio-linguistic investigation of Glasgow middle-class speakers' (Université Montpellier II, 2011).

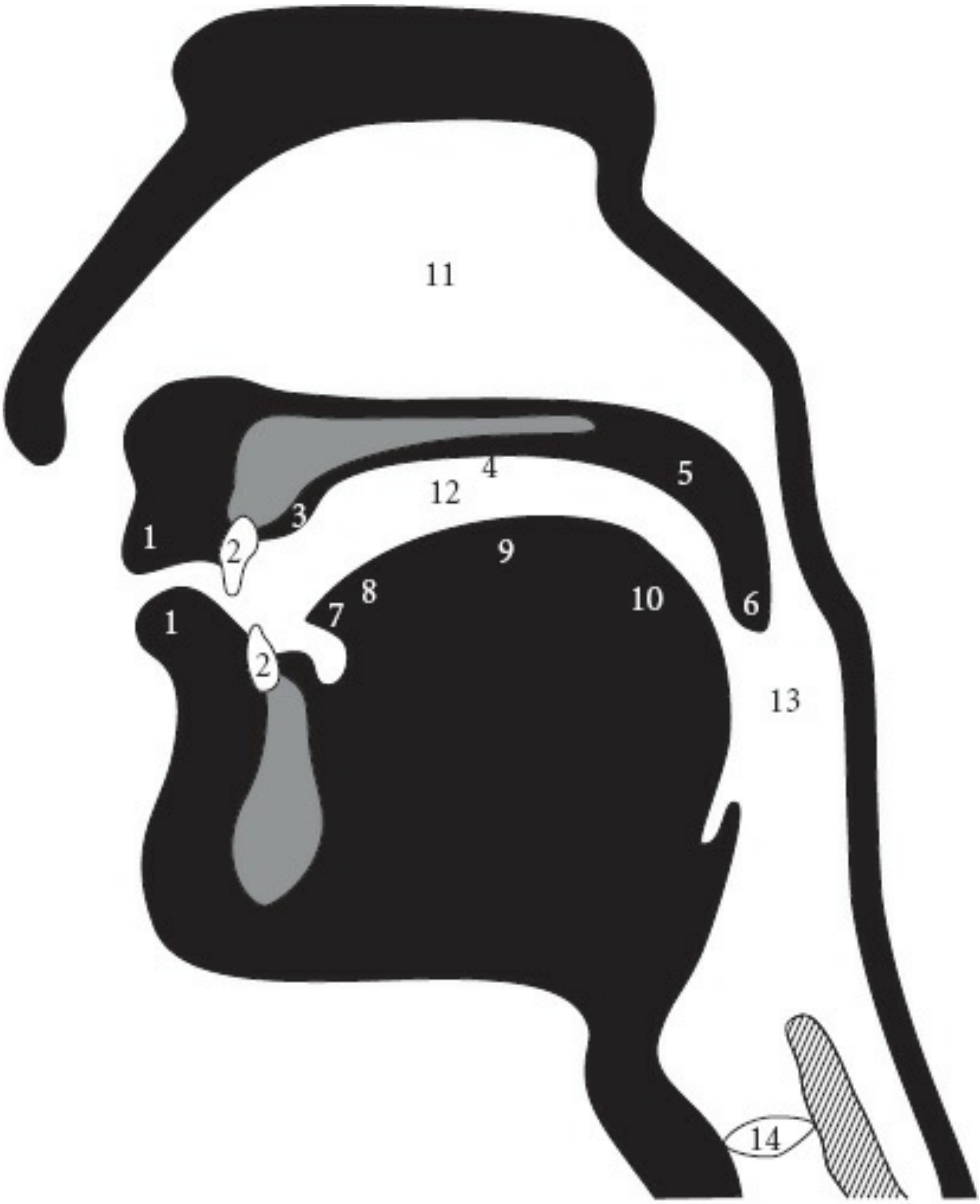
I am grateful to Francis Nolan, professor of phonetics at Cambridge University and principal investigator of the IViE project (Intonational Variation in English), for permission to use, for educational purposes, two sound files from the IViE corpus, collected as part of the ESRC project R000237145 'English Intonation in the British Isles'.

Many thanks to my friend and colleague Nick Myers, of the Université Paul Valéry here in Montpellier, to Inès Brulard and to my Masters student Alison Gilbreath for taking the time and trouble to record exercise material for me.

I thank two anonymous readers of the draft chapters for their detailed comments.

I thank Danielle Descoteaux and Julia Kirk at Wiley-Blackwell for their patience and encouragement. I also thank Colleen Fitzgerald for permission to use the written passage which she constructed for the recordings of Texan English, my copy-editor Fiona Sewell for her help with the draft text and Leah Morin for her help with the proofs.

Figure 1 The organs of speech



1 Lips	2 Teeth
3 Alveolar ridge	4 Hard palate
5 Soft palate (velum)	6 Uvula
7 Tip of the tongue	8 Blade of the tongue
9 Front of the tongue	10 Back of the tongue
11 Nasal cavity	12 Oral cavity
13 Pharynx	14 Larynx

Figure 2 The International Phonetic Alphabet (Department of Theoretical and Applied Linguistics, School of English, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece)

CONSONANTS (PULMONIC)

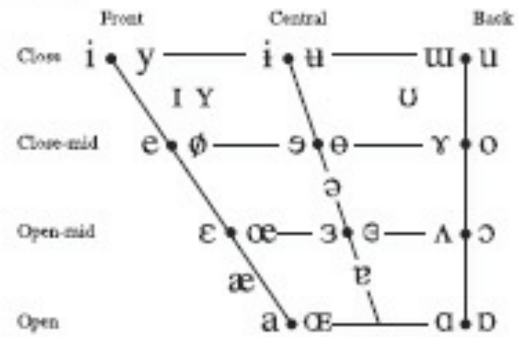
	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

CONSONANTS (NON-PULMONIC)

Clicks		Voiced implosives		Ejectives	
ʘ	Bilabial	ɓ	Bilabial	ʼ	Examples:
ǀ	Dental	ɗ	Dental/alveolar	pʼ	Bilabial
ǃ	(Post)alveolar	ʄ	Palatal	tʼ	Dental/alveolar
ǂ	Palatoalveolar	ɠ	Velar	kʼ	Velar
ǁ	Alveolar lateral	ʛ	Uvular	sʼ	Alveolar fricative

VOWELS



Where symbols appear in pairs, the one to the right represents a rounded vowel.

OTHER SYMBOLS

ʍ Voiceless labial-velar fricative

w Voiced labial-velar approximant

ɥ Voiced labial-palatal approximant

ʜ Voiceless epiglottal fricative

ʢ Voiced epiglottal fricative

ʡ Epiglottal plosive

ɕ ʑ Alveolo-palatal fricatives

ɺ Voiced alveolar lateral flap

ɧ Simultaneous ʃ and x

Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary. k͡p t͡s

SUPRASEGMENTALS

ˈ Primary stress

ˌ Secondary stress ˌfoʊnəˈtɪʃən

ː Long eː

ˑ Half-long eˑ

˘ Extra-short ĕ

| Minor (foot) group

‖ Major (intonation) group

. Syllable break ɹi.ækt

‿ Linking (absence of a break)

DIACRITICS Diacritics may be placed above a symbol with a descender, e.g. ŋ̊

Voiceless	n̥ d̥	Breathy voiced	b̤ a̤	Dental	t̪ d̪
Voiced	s̬ t̬	Creaky voiced	b̰ a̰	Apical	t̺ d̺
ʰ Aspirated	tʰ dʰ	Linguolabial	t̼ d̼	Laminal	t̻ d̻
More rounded	ɔ̹	ʷ Labialized	tʷ dʷ	Nasalized	ẽ
Less rounded	ɔ̜	ʲ Palatalized	tʲ dʲ	ⁿ Nasal release	dⁿ
Advanced	u̟	ˠ Velarized	tˠ dˠ	ˡ Lateral release	dˡ
Retracted	e̠	ˤ Pharyngealized	tˤ dˤ	No audible release	d̚
Centralized	ë	Velarized or pharyngealized	ɫ		
Mid-centralized	e̽	Raised	e̝ (ɹ̝ = voiced alveolar fricative)		
Syllabic	n̩	Lowered	e̞ (β̞ = voiced bilabial approximant)		
Non-syllabic	e̯	Advanced Tongue Root	e̘		
Rhoticity	ɚ a˞	Retracted Tongue Root	e̙		

TONES AND WORD ACCENTS

LEVEL			CONTOUR		
e̋ or	˥	Extra high	ě or	˩˥	Rising
é	˦	High	ê	˥˩	Falling
ē	˧	Mid	e᷄	˧˥	High rising
è	˨	Low	e᷅	˩˨	Low rising
ȅ	˩	Extra low	e᷈	˧˦˧	Rising-falling
↓		Downstep	↗		Global rise
↑		Upstep	↘		Global fall

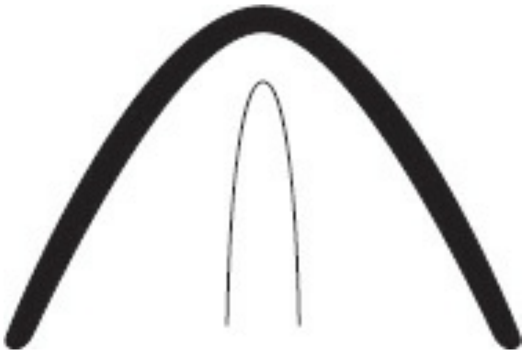
1

English Phonetics: Consonants (i)

1.1 Airstream and Articulation

Speech sounds are made by modifying an airstream. The airstream we will be concerned with in this book involves the passage of air from the lungs out through the **oral** and **nasal cavities** (see figure 1). There are many at which that stream of air can be modified, and several in which it can be modified (i.e. constricted in some way). The first point at which the flow of air can be modified, as it passes from the lungs, is in the **larynx** (you can feel the front of this, the Adam's apple, protruding slightly at the front of your throat; see figure 1), in which are located the **vocal folds** (or vocal **cords**). The vocal folds may lie open, in which case the airstream passes through them unimpeded. Viewed from above, the vocal folds, when they lie open, look like this:

Open vocal folds



The vocal folds may be brought together so that they are closed, and no air may flow through them from the lungs:

Closed vocal folds

One way in which the outgoing stream of air may be modified is by applying a certain level of constant muscular pressure sufficient to close the vocal folds along their length, but only just; the build-up of air pressure underneath this closure is sufficient, given the degree of muscular pressure, to force that closure open, but the air pressure then drops, and the muscular pressure causes the folds to close again. The sequence is then repeated, very rapidly, and results in what is called vocal fold vibration. You should be able to feel this vibration if you put your fingers to your larynx and produce the sound which is written as <z> in the word (although you will probably also feel vibration elsewhere in your head). Sounds which are produced with this vocal fold vibration are said to be **voiced** sounds, whereas sounds produced without such vibration are said to be **voiceless**.

To transcribe speech sounds, phoneticians use the International Phonetic Alphabet (the IPA: see figure 2); the IPA symbol for the sound written <z> in is [z]. You should be able to feel the presence of vibration in [z] if you put your fingers to your larynx and produce [z], then [s] (as in), then [z] again: [z] is voiced, whereas [s] is voiceless. This distinction will constitute the first of three descriptive parameters by means of which we will describe a given consonantal speech sound: we will say, for any given consonant, whether it is voiced or voiceless.

1.2 Place of Articulation

We will refer to the points at which the flow of air can be modified as places of articulation. We have just identified the vocal folds as a place of articulation; since the space between the vocal cords is referred to as the glottis, we will refer to sounds produced at this place of articulation as **glottal** sounds. There are many other places of articulation; we will identify a further seven.

Firstly, sounds in which the airflow is modified by forming a constriction between the lower lip and the upper lip are referred to as **bilabial** sounds. An example is the first sound in .

A bilabial sound: the first sound in

Secondly, sounds in which there is a constriction between the lower lip and the upper teeth are referred to as **labio-dental** sounds. An example is the first sound in .

A labio-dental sound: the first sound in

Thirdly, sounds in which there is a constriction between the tip of the tongue and the upper teeth are referred to as **dental** sounds. An example is the first sound in .

A dental sound: the first sound in

For the remaining places of articulation, let us distinguish between the tip, the blade of the tongue, the front of the tongue and the back of the tongue (as in figure 1). Let us also distinguish various points along the upper part of the mouth. We will identify four different areas: the **alveolar ridge** (the hard, bony ridge behind the teeth; see figure 1), the **hard palate** (the hard, bony part of the roof of the mouth; see figure 1), the **palato-alveolar** (or **post-alveolar**) **region**[1] (the area in between the alveolar ridge and the hard palate), and the **velum** (the soft part at the back of the roof of the mouth, also known as the **soft palate**; see figure 1).

Sounds in which there is a constriction between the blade or tip of the tongue and the alveolar ridge are called **alveolar** sounds. An example is the first sound in .

An alveolar sound: the first sound in

Sounds in which there is a constriction between the blade of the tongue and the palato-alveolar (or post-alveolar) region are called **palato-alveolar** sounds. An example is the first sound in

.

A palato-alveolar sound: the first sound in

Sounds in which there is a constriction between the front of the tongue and the hard palate are called **palatal** sounds. An example is the first sound in (although this may be less obvious to you; we will return to this sound below).

A palatal sound: the first sound in

Sounds in which there is a constriction between the back of the tongue and the velum are called **velar** sounds. An example is the first sound in .

A velar sound: the first sound in

1.3 Manner of Articulation: Stops, Fricatives and Approximants

We have now identified eight places of articulation: glottal, bilabial, labio-dental, dental, alveolar, palato-alveolar, palatal and velar. For any given sound we will say whether it is voiced or voiceless, and what its place of articulation is. But to distinguish between the full range of speech sounds, we will require a third descriptive parameter: **manner of articulation**. To identify the manner in which a sound is articulated, we will identify three different (complete closure, close approximation and open approximation), and thus three different categories of consonant: stops, fricatives and approximants.

1.3.1 Stops

The articulators in question may form a stricture of **complete closure**; this is what happens when one produces the first sound in . Here the lower and upper lips completely block the flow of air from the lungs; that closure may then be released, as it is in , and may then produce a sudden outflow of air. Sounds which are produced with complete closure are referred to as **stops** (or **plosives**).

We may describe the first sound in as a voiceless bilabial stop (transcribed as [p]) and we will henceforth identify all consonants with three-term labels of this sort. The consonant in is also a bilabial stop, but differs from that in : it is voiced. This consonant (transcribed as [b]) is a voiced bilabial stop.

The first sound in is a voiceless alveolar stop; it is transcribed as [t]. Its voiced counterpart is the consonant in . This sound, the voiced alveolar stop, is transcribed as [d].

The first sound in is a voiceless velar stop; it is transcribed as [k]. Its voiced counterpart, the voiced velar stop, is transcribed as [g]; an example is the consonant in .

We have now identified bilabial, alveolar and velar stops; stops may be made at many other places of articulation, but we will ignore those, as they are not relevant to the study of English. There is one further stop which we must mention, however, as it is very common in the speech of most speakers of English. This is the **glottal stop** (transcribed as [ʔ]). It is made by forming a constriction of complete closure between the vocal folds. This is the sound made instead of [t] in many Scottish and Cockney pronunciations of, for example, the word . We will see that it is present in the speech of almost every speaker of English, no matter what the accent. There is no question of describing the glottal stop as voiced or voiceless, since it is articulated in the glottis itself.

1.3.2 Fricatives

Let us now distinguish between complete closure and another, less extreme, degree of constriction: **close approximation**. Sounds which are produced with this kind of constriction entail a bringing together of the two articulators to the point where the airflow is not quite fully blocked: enough of a gap remains for air to escape, but the articulators are so close together that friction is created as the air escapes. Sounds of this sort are referred to as **fricatives**.

The first sound in is created by bringing the lower lip close to the upper teeth in a constriction of close approximation. This sound is a voiceless labio-dental fricative (transcribed as [f]). Its voiced counterpart (the voiced labio-dental fricative, transcribed as [v]) is the consonant in .

The first sound in is created by bringing the tip of the tongue into a constriction of close approximation with the upper teeth. This sound is a voiceless dental fricative, transcribed as [θ]. Its voiced counterpart, the voiced dental fricative (transcribed as [ð]) is, for some speakers, the first sound in the word .[2]

The first sound in is created by bringing the tip or blade of the tongue into a constriction of close approximation with the alveolar ridge. This sound, transcribed as [s], is a voiceless alveolar fricative. Its voiced counterpart, the voiced alveolar fricative (transcribed as [z]) is the consonant in .

The first sound in is created by bringing the blade of the tongue into a constriction of close approximation with the palato-alveolar region. This sound, transcribed as [ʃ], is a voiceless palato-alveolar fricative. Its voiced counterpart, transcribed as [ʒ], is the second consonant in .

Fricatives may be articulated at any point of articulation, but many of those sounds are irrelevant to the study of English. However, we will mention three.

One is the voiceless velar fricative [x], found in the speech of many Scots, in words such as . Another is the voiceless fricative [ʍ], again found in the speech of many Scots, as in words like (as opposed to) and (as opposed to); its place of articulation is labial-velar (explained in 1.3.3).

A third is the glottal fricative [h], as in the first sound in . This sound is produced by bringing the vocal cords into a constriction of close approximation, so that friction is produced. As the vocal cords are not vibrating, we will take it that this is a voiceless sound.

1.3.3 Approximants

The least radical degree of constriction occurs when the articulators come fairly close together, but not sufficiently close together to create friction. This kind of stricture is called **open approximation**. Consonants produced in this way are called **approximants**.

The first sound in is an approximant. It is produced by bringing the front of the tongue close to the hard palate. Although the sides of the tongue are in a constriction of complete closure with the upper gums, the air escapes along a central groove in which the front of the tongue is not close enough to the hard palate to create friction. This sound, transcribed as [j], is a voiced palatal approximant. Approximants are normally voiced, so we will not discuss any voiceless counterparts for these sounds.

The first sound in many English speakers' pronunciation of , etc. is an approximant. It is produced by bringing the blade of the tongue into a constriction of open approximation with the alveolar ridge. This approximant, transcribed as [ɹ], is referred to as an alveolar approximant. As with [j], the sides of the tongue form a constriction of complete closure with the gums at the sides of the mouth, but the air escapes along a central groove without creating friction. For most speakers (and in varying degrees, depending on the accent), the tongue body is somewhat retracted when [ɹ] is uttered; it is therefore often referred to as a -alveolar approximant, but 'alveolar approximant' will suffice for our purposes.[3]

We will be looking at more English approximants in chapter 2. For the moment, let us identify one further such sound, the sound at the beginning of . In producing this sound, the lips form a constriction of open approximation: there is no friction produced. But its articulation is more complicated than that of [j], the palatal approximant, since it also involves another articulation, between the back of the tongue and the velum (i.e. a velar articulation). We will therefore refer to it as a voiced approximant; it is transcribed as [w].

Notes

1 Many phonologists and phoneticians use the term 'palato-alveolar', but the chart of symbols used by the International Phonetics Association uses the term 'post-alveolar'. It will suffice for our purposes if the student takes the two terms to be interchangeable. There are no rigid physiological divisions between the alveolar ridge and the hard palate; the transition from one to the other is a continuum. And the range of articulations which can be made in between the two is

relatively varied, leading some phoneticians to distinguish alveo-palatal from palato-alveolar articulations. We will simplify by ignoring these details.

2 Many speakers of English do not have a voiced dental fricative; rather, the sound lacks friction: it is a voiced dental approximant.

3 The articulation of an [ɹ] kind of articulation in some American and West Country accents is also referred to by some as approximant. The term 'retroflex' means that the blade and tip of the tongue are curled upwards and backwards to some extent, so that the underside of a part of the tongue forms the relevant articulation. Somewhat inaccurately, we will use [ɹ] for these sounds.

Exercises

1 Give the appropriate three-term description for each of the following sounds (e.g. [k]: voiceless velar stop):

[θ] [b] [f] [ʃ] [j] [t]

2 Give the appropriate phonetic symbol for each of the following sounds:

(a) a voiced palato-alveolar fricative

(b) a voiced alveolar stop

(c) a voiced velar stop

(d) a voiced dental fricative

(e) a voiced labio-dental fricative

3 What phonetic property distinguishes each of the following pairs of sounds (e.g. [p] and [b]: voicing; [s] and [ʃ]: place of articulation; [t] and [s]: manner of

articulation)?

(a) [k] and [g]
(b) [b] and [d]
(c) [d] and [z]
(d) [z] and [ʒ]
(e) [ʃ] and [ʒ]
(f) [d] and [g]


Listen to sound files online

4 Listen to Track 1.1 at www.wiley.com/go/carrphonetics. Which of the words on the recording begin with a fricative? The words are listed below.

5 Listen to Track 1.2. Which of the words on the recording end with a fricative? The words are listed below.

6 Listen to Track 1.3. Which of the words on the recording begin with a stop? The words are listed below.

7 Describe the position and action of the articulators during the production of the following sounds (e.g. [d]: the blade of the tongue forms a constriction of complete closure with the alveolar ridge; the vocal cords are vibrating):

[v] [θ] [k] [b]

2

English Phonetics: Consonants (ii)

2.1 Central vs Lateral

In discussing the alveolar approximant [ɹ], we said that the air escapes along a central groove (of the tongue, in this case; the same kind of groove can be formed by the lips). This is true for all of the fricatives and approximants described in chapter 1: they are all fricatives and approximants. However, it is possible to produce fricatives and approximants in which this is not the case. For instance, in the first sound in , the centre of the blade of the tongue forms a stricture of complete closure with part of the alveolar ridge, but the articulation which 'counts' is that between the of the tongue and the alveolar ridge. Since the sides of the tongue form a constriction of open articulation with the alveolar ridge, and no friction is created, we refer to this sound (transcribed as [l]) as a voiced alveolar lateral approximant. Since English fricatives and approximants are typically central, we will use the term 'lateral' for laterals, and omit the term 'central' in describing central fricatives and approximants in English speech. The sounds [l] and [ɹ] are, clearly, quite similar: both are approximants, both are voiced, both are alveolar. The principal difference is that the former is lateral and the latter central.[1]

2.2 Taps and Trills

We have said that, for a great many speakers of English, the sound at the beginning of words such as , etc. is a post-alveolar approximant: [ɹ]. The same is true of the sound which occurs after stops in words such as , etc. However, some speakers utter, not an approximant, but a sound which is very like a voiced alveolar stop of very short duration. Many Scots utter this sound, rather than [ɹ], after stops, as in the words just cited. During the articulation of this sound, the blade of the tongue comes into a momentary constriction of complete closure with the alveolar ridge. This sound, transcribed as [ɾ], is referred to as a voiced alveolar **tap** (or flap). This is also the sound that many American speakers have instead of [t] or [d] in words such as , etc.

Speakers of certain accents of English may utter neither an [ɾ] nor an [ɹ] in words such as and , but a sound referred to as a voiced alveolar **trill**. Trills are produced by holding one articulator (e.g. the blade of the tongue) next to the other (e.g. the alveolar ridge) in a constriction of complete closure, but without the same muscular pressure as one finds in stops. The result is that air pressure builds up behind the closure and forces it open; the air pressure then reduces, and the muscular pressure again creates a constriction of complete closure. This sequence may be repeated in quick succession, producing, in the case of an alveolar trill, a series of taps of the tongue against the alveolar ridge. The alveolar trill is transcribed as [r], but is relatively rare. Scots are often said to produce this sound; however, most speakers of Scottish varieties of English typically produce, not an alveolar trill, but an alveolar tap.

2.3 Secondary Articulation

We have said that the lateral approximant [l] is alveolar. However, laterals may also be produced with an additional articulation, such as one formed between the back of the tongue and the velum, i.e. a velar articulation. When this happens, we may distinguish between the alveolar articulation as the **primary articulation** and the velar one as the **secondary articulation**. Where a secondary articulation is velar, this process is referred to as **velarization**: we say that the lateral is velarized. A velarized lateral approximant is transcribed using the velarization diacritic, thus: [ɫ]. This sound is often referred to as **'dark l'**.[2] Where a secondary articulation is palatal (formed between the front of the tongue and the hard palate), this process is referred to as **palatalization**; we say that the lateral is palatalized. A palatalized lateral is transcribed using the palatalization diacritic, thus: [lʲ]. The term **'clear l'** is often used to refer to [lʲ], or to [l] (neither palatalized nor 'dark'). In subsequent chapters, we will consider the status of 'dark l' and 'clear l' in different accents of English.

2.4 Affricates

We have, thus far, distinguished three classes of consonant according to degree of constriction: stops, fricatives and approximants. Consider the first sound in : it is like a stop in that there is complete closure between the blade of the tongue and the palato-alveolar region. However, it is like a fricative in that it clearly involves friction. That friction occurs during the of the closure, which we referred to in 1.3.1. Sounds produced with a constriction of complete closure followed by a release phase in which friction occurs are called **affricates**. We might say that one of the main differences (place of articulation apart) between the first sound in and the first sound in is that, during the release phase of the [t] in , there is no friction of the sort one finds during the release phase of the first sound in . We might therefore think of affricates as stops with a slow, fricative, release phase. The affricate in is a voiceless palato-alveolar affricate, transcribed as [tʃ]. Its voiced counterpart is [dʒ], the first sound in , etc.[3]

These two affricates occur in the speech of most speakers of English. In later chapters, we will examine some other affricates which occur in the speech of speakers of certain accents of English.

2.5 Aspiration

The first stop in , we said, is a voiceless bilabial stop. So too is the first stop in . But the bilabial stop in differs phonetically from the bilabial stop in : if you hold the palm of your hand up close to your mouth when uttering , you will feel a stronger puff of air on releasing the bilabial stop than you will when you utter . That ‘stronger puff of air’ phenomenon is called **aspiration**: we say that the bilabial stop in is an voiceless stop, whereas the stop in is . Aspirated voiceless stops are transcribed with the aspiration diacritic ([ʰ]), so that the bilabial stop in is transcribed as [pʰ]. Unaspirated stops are transcribed without that diacritic, so that the bilabial stop in is transcribed as [p].

2.6 Nasal Stops

We have been making an assumption in our discussion thus far, concerning the position of the velum in the production of the speech sounds we have described. We have assumed that, in all of these sounds, the air from the lungs is escaping only through the mouth (the oral cavity). This is true if the velum is in the position, such that it prevents the flow of air out through the nasal cavity (see figure 1). In all of the sounds discussed thus far, the velum is indeed raised: we describe all such sounds as **oral** sounds. But the velum may be lowered, to allow escape of air through the nasal cavity (see figure 1). Sounds produced with the velum lowered, and with air escaping through the nasal cavity alone, are referred to as **nasal stops**.[4] These may occur at most places of articulation;

let us consider those which are relevant for the study of English.

While nasal stops may be either voiced or voiceless, they are typically voiced in most human languages; we will therefore ignore voiceless nasal stops and use the term 'nasal stop' to imply ' nasal stop'.

Bilabial nasal stops (transcribed [m]) entail, as one would expect, complete closure between the lips, voicing, and escape of the air through the nasal cavity. An example is the first consonant in .

Labio-dental nasal stops (transcribed [ɱ]) entail complete closure between the lower lip and the upper teeth, voicing, and escape of the air through the nasal cavity. An example is the second consonant in . In English, they occur before labio-dental sounds, as in this case. The nasal stop articulation in cases such as these reflects a process of **assimilation**. Assimilation processes are processes in which one sound becomes similar to an adjacent sound. In this case, the nasal is assimilated to the following fricative, in the sense that it 'takes on' the place of articulation of the fricative. Such processes involve a principle of **ease of articulation**. In this case, if the nasal in is articulated at the same place as the following fricative, this saves the speaker the articulatory effort of moving from a bilabial to a labio-dental articulation. We will return to such processes in chapter 6.

Dental nasal stops (transcribed as [n̪]) entail complete closure between the tip of the tongue and the upper teeth, voicing, and escape of the air through the nasal cavity. An example is the second consonant in . As in this case, they occur before other dental sounds, and this too is a matter of assimilation involving place of articulation.

Alveolar nasal stops (transcribed as [n]) entail complete closure between the blade of the tongue and the alveolar ridge, voicing, and escape of the air through the nasal cavity. An example is the first sound in .

Velar nasal stops (transcribed as [ŋ]) entail complete closure between the back of the tongue and the velum, voicing, and escape of the air through the nasal cavity. An example is the last sound in or the nasal stop as it is often articulated (especially in faster or more casual speech styles) in the word Once again, the latter case involves assimilation.

Notes

1 The central approximant [ɹ] also differs from [l] in having tongue body retraction and lip rounding. We will see shortly that alveolar laterals may be produced with retraction too.

2 The term 'dark l' can also be used to refer to lateral approximants in which the body/back of the tongue is retracted and/or lowered. Accents of English vary with respect to the exact articulatory nature of their 'dark l's: some are velarized, while others have no velar articulation, but have, instead, retraction and/or lowering of the back/body of the tongue. Such retraction can lead to loss of alveolar contact, and thus to [l]-vocalization, in which the articulation becomes vowel-like.

3 Some authors transcribe [tʃ] as [č] and [dʒ] as [ǰ]. We should, if we were to stick strictly to the conventions of the International Phonetics Association, transcribe both affricates

with a ‘tie bar’ above the two symbols; we depart here from the conventions of the IPA chart, which does not contain an ‘affricate’ category.

4 The term ‘nasalized’, as opposed to ‘nasal’, is used to describe sounds in which air escapes through cavities, the oral the nasal. The term ‘nasal’ is used to describe sounds in which the air escapes through the nasal cavity alone.

Exercises

Listen to sound files online

1 Listen to Track 2.1 at www.wiley.com/go/carrphonetics. For each of the words on the recording, identify (a) any oral stops, (b) any fricatives, (c) any approximants, (d) any affricates and (e) any nasals. For each sound that you identify, say whether it is voiced or voiceless and what its place of articulation is (e.g. the word : voiceless alveolar stop [t] and voiceless bilabial stop [p]; voiceless alveolar fricative [s]; no approximants, affricates or nasals). The words are:

2 Listen to Track 2.2. Which of the words on the recording begins with an affricate, and which (if any) with a stop?

Many speakers of English typically utter words like and with an affricate at the beginning of the word. This means that and are typically indistinguishable. None the less, when asked in a phonetics class whether they utter words such as with an affricate, such speakers often deny that they do. These speakers typically have a more careful pronunciation of words such as and , in which there a [j].

Notice, however, that there is no such more careful pronunciation of words like and : one never hears these pronounced with [tj] and [dj]. In order to explain the difference between and , we need to say that the speaker in some sense intends to utter [dj] in , but that ease of articulation results in a palato-alveolar affricated release of the stop closure, rather than a transition from an alveolar closure to a stricture of open approximation between the front of the tongue and the hard palate. In the case of and , the intended articulation is a palato-alveolar affricate.

If you are a speaker of General American, you may well never utter a [j] in words like and , in which case you will utter a stop followed by a vowel. However, you may well also have been told at school that the ‘correct’ pronunciation of such words has a [j] after the stop. Your speech may well vary with respect to the presence or absence of the [j]. If your speech does vary in this way, how do you pronounce ?

3 Listen to Track 2.3. Give a phonetic transcription of each of the words on the recording, using a ‘V’ for the vowels. The words are:

You may well have noticed that the nasal stop in is velar, rather than alveolar. It requires considerable conscious effort to utter that nasal stop as alveolar, and when one does so, the resulting pronunciation sounds quite unnatural. This appears to be the result of a process of anticipatory assimilation: the tongue adopts the articulatory position for the velar stop [k] during the pronunciation of the nasal.

But what about the nasal-plus-velar stop sequence in ? Many speakers of English find it easier to utter an alveolar, rather than a velar, nasal there, despite the fact that cases like also contain a sequence of a nasal stop followed by a velar stop. Do you have any hunches as to why the two cases should be different?

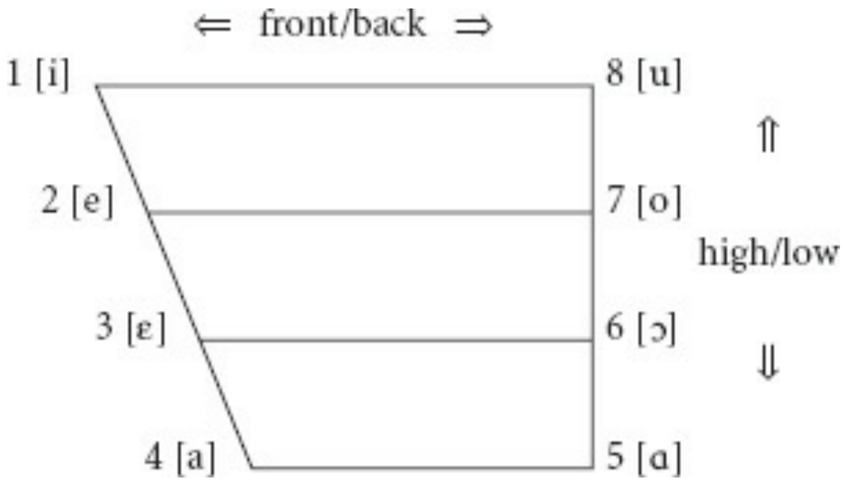
3

English Phonetics: Vowels (i)

3.1 The Primary Cardinal Vowels

Let us begin by assuming that all vowels are voiced and are articulated with a constriction of open approximation. We will also assume, for the moment, that all vowels are sounds (i.e. that the velum is raised during their production). The range of positions which the tongue can occupy within the oral cavity while remaining in a constriction of open approximation is quite large. Let us call the entire available space for such articulations the **vowel space**. We will require a means of plotting the point at which a given vowel is articulated in the vowel space. In order to do this, we will appeal to an idealized chart of that space, as follows (this chart is repeated in the IPA chart in figure 2):

(1) The vowel space and the primary cardinal vowels



In this diagram, we represent the vowel space along two dimensions. The first is the **high/low** dimension (also referred to as the close/open dimension), depicting the height of the body of the tongue during the articulation of a vowel (i.e. depicting). This is represented as the vertical axis in the diagram. The second is the **front/back** dimension, depicting the extent to which the body of the tongue lies towards the front of the vowel space. This is represented as the horizontal axis in the diagram. We may identify three arbitrary points along this dimension: **front**, **central** and **back**. In using these two dimensions, we can say, for any given vowel, how high in the vowel space it is articulated, and whether it is a front, central or back vowel. To these two descriptive parameters, we will add a third, which refers to lip position: we will say, for a given vowel, whether, during its articulation, the lips are rounded or not. We will refer to the former sort of vowel as a **rounded** vowel and the latter as an **unrounded** vowel.

It is convenient to identify several points along the perimeter of the vowel space. Once we have done this, we can plot the location of any given vowel in relation to those points. Vowels articulated at those points are called **the cardinal vowels**. We will now identify eight of them.

Let us begin with the vowel which is produced when the lips are unrounded and the tongue is located as high as possible and as front as possible, without causing friction, in the vowel space. This is **cardinal vowel no. 1**, depicted at the top left-hand corner of the diagram in (1) above. That vowel is transcribed as [i]. Using our three descriptive parameters, we will refer to this as a high front unrounded vowel. We will not seek to exemplify cardinal vowels with words from English, or any other language, since, typically, speakers do not utter vowel sounds which are quite as peripheral in the vowel space as the cardinal vowels. Rather, we will plot the place of articulation of English vowels the cardinal vowels, using the vowel space diagram as a map of the vowel space. The vowel in many English speakers' pronunciation of the word , for instance, is quite close to cardinal vowel no. 1: it too is a **high front rounded vowel**, but it is not quite as as cardinal vowel no. 1: it is typically slightly less high and slightly less front in its articulation.

Let us now identify the cardinal vowel which lies at the 'opposite end' of the vowel space: the vowel which is produced when the lips are unrounded and the body of the tongue is as low as possible and as far back as possible, without causing friction. This is **cardinal vowel no. 5**. Its location is depicted at the bottom right-hand corner of the diagram in (1) above. Transcribed as [ɑ], it is a **low back unrounded vowel**.

We have now identified two 'anchor' points in the vowel space; we may now proceed to identify further cardinal vowels in relation to these. If the lips remain unrounded and the body of the tongue remains as low as possible in the vowel space (as for cardinal vowel no. 5), but the tongue is moved as far to the front of that space as is possible without causing friction, then **cardinal vowel no. 4** is produced. It is transcribed as [a].

We have now identified two vowel heights: high and low. You should be able to feel this difference in tongue height if you utter cardinal vowel no. 1 followed by cardinal vowel no. 4: the jaw opens considerably and the body of the tongue lowers considerably as one moves from the former to the latter. There is a continuum of vowel heights between these two heights; we will identify two arbitrary points along this continuum: **high-mid** and **low-mid**. If the lips remain unrounded and the body of the tongue remains as far front as is possible, but the tongue height is lowered somewhat from the cardinal vowel no. 1 position, one arrives at the front, **high-mid unrounded vowel** known as **cardinal vowel no. 2**. This is transcribed as [e].

In retaining the same lip position and the same degree of frontness, one may lower the body of the tongue further still to the low-mid position, and arrive at the **front low-mid unrounded vowel** known as **cardinal vowel no. 3**. This is transcribed as [ɛ].

If you articulate cardinal vowel no. 1, then cardinal vowels nos. 2, 3 and 4, you should feel the body of the tongue lowering progressively. These are all front unrounded vowels: the difference between them lies in the height of the tongue.

Let us now consider the back cardinal vowels. If the body of the tongue is as high as possible and as far back as possible without causing friction, and the lips are, this time, rounded, then **cardinal vowel no. 8** is produced. This **high back rounded vowel** is transcribed as [u].

If the lips remain rounded and the tongue remains as far back as possible, but the tongue height is lowered to the high-mid position, **cardinal vowel no. 7** is produced. This **high-mid back rounded vowel** is transcribed as [o].

In retaining the same degree of backness and the same lip position, one may lower the height of the tongue still further, to the low-mid position, and thus produce the **low-mid back rounded vowel** known as **cardinal vowel no. 6**. This is transcribed as [ɔ].

You should be able to feel the tongue lowering progressively as you make the transition from cardinal vowel no. 8, through cardinal vowel no. 7, to cardinal vowel no. 6; the tongue goes through the same lowering process as it does for the transition from cardinal vowel no. 1, through no. 2, to no. 3.

We have now identified the eight primary cardinal vowels. With these reference points established, we may describe the articulation of specific English vowels in relation to them. Let us begin by looking at those referred to as the English short vowels.

3.2 RP and GA Short Vowels

There is considerable variation in the vowel sounds uttered by speakers of different accents of English, and we will be considering that variation in later chapters. For the moment, we will begin with two particular accents; we will later describe others. We will, somewhat arbitrarily, begin with the accents known as **Received Pronunciation (RP)** and **General American (GA)**. RP is the accent often referred to as the 'prestige' accent in British society and associated with the speech of the graduates of the English public schools. It is thus defined largely in terms of the social class of its speakers. We do not select it as one of our starting points for that reason; rather, we select it as it tends to be the accent which foreign learners of British English are taught, and has thus been widely described. GA tends to be defined in terms of the geographical location, rather than the social class, of its speakers. The term 'GA' is an idealization over a group of accents whose speakers inhabit a vast proportion of the United States: it excludes Eastern accents such as the New York City accent, and Southern accents (such as that spoken in Texas).

It has often been pointed out that terms such as 'RP' and 'GA' entail a great deal of idealization, in that they are used to cover a variety of somewhat different, if converging, accents. We accept this as inevitable: it will be true of any term we use to describe an accent (e.g. 'New York City', 'Cockney', 'Scouse', 'Geordie', 'South African', etc.) and indeed it is true of the term 'accent' itself. But we need some way of expressing valid generalizations about the speech sounds which members of different speech communities utter. For instance, it is generally true that, while RP speakers pronounce and differently, many speakers with accents found in the North of England do not. To refuse to speak of different accents would be to throw the baby out with the bathwater, and to deny ourselves the opportunity of expressing statements which are informative, if subject to certain caveats.

We have said nothing, as yet, about the length of vowels. For speakers of RP and GA, the vowels in and differ in several respects, one of which is vowel length. If you are an RP or a GA speaker, and you utter the two words, you will probably agree that the vowel in the former is longer than that in the latter. We will, accordingly, refer to the former as a long vowel and the latter as a short vowel. Vowel length is a relative matter: when we say that the vowel in is a short vowel, we are not referring to its duration in milliseconds; rather, we are saying that it is short in relation to other vowels, such as that in . The vowel in is typically articulated with the body of the tongue fairly front and fairly high, and with the lips unrounded. We will transcribe that vowel as [ɪ]. While it is a high front unrounded vowel, it is less high and less front than the vowel in . Its location is depicted in (2) below.

Now consider the vowel in RP and GA speakers' pronunciation of the vowel in the word . This is, for many speakers, a high back rounded vowel, articulated in the region near to cardinal vowel no. 8. It is similar to the vowel in , but less high and less back. It is also shorter than that in . We will transcribe this short vowel as [ʊ]; its location is depicted in (2) below.

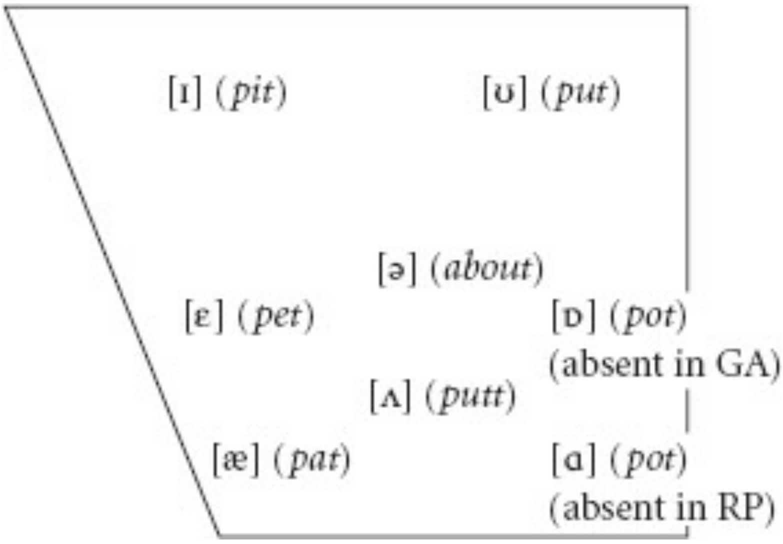
For RP and GA speakers, there is a distinction between the vowel in and that in . Both are short vowels, but they differ in several respects. Firstly, the latter vowel is unrounded. Secondly, the vowel in is articulated with a fairly low tongue height: typically, it is just below the low-mid position. Thirdly, the vowel in is located at around the half-way point on the front/back axis. We will refer to vowels located in this region as **central** vowels. We will transcribe this vowel as [ʌ]; its location is depicted in (2) below.

In both RP and GA, the vowels in and differ. Both vowels are unrounded, but the vowel in is shorter than that in , and the vowel in is a low front vowel, whereas that in is a low back vowel. The low front unrounded vowel in is articulated higher and less front than cardinal vowel no. 4. We will transcribe this as [æ]; its location is depicted in (2) below (although the GA vowel is higher than the RP vowel, and sounds rather [ɛ]-like to British speakers).

The short vowel in RP and GA speakers' pronunciation of the word is a front unrounded vowel, whose height is somewhere between cardinal vowels nos. 2 and 3. For most RP and GA speakers, it is closer to cardinal vowel no. 3 than to cardinal vowel no. 2 in height; it is also somewhat more centralized than cardinal vowel no. 3. For convenience' sake, we will transcribe it as [ɛ]; its location is depicted in (2) below.

The short vowel in the RP speaker's pronunciation of the word is a back rounded vowel which is articulated with a tongue height somewhere between low and lowmid (i.e. between cardinal vowels nos. 5 and 6). It is transcribed as [ɒ]; its location is given in (2) below. This vowel is absent from the GA system: GA speakers have the vowel [ɑ] in words such as . [ɑ] is a short back rounded low vowel.

(2) RP and GA short vowels



We have used the words , , , , and to

illustrate these vowels, since these words differ in pronunciation only with respect to the vowel. In discussing vowels, we will also adopt the lexical sets adopted by Wells (1982; see Suggested Further Reading). These are key words selected by Wells to bring out the similarities and differences between RP and GA. We will therefore, at times, refer to the vowel in words such as as the KIT vowel. The vowel in words such as we will call the DRESS vowel; words such as have the TRAP vowel; words such as have the LOT vowel; words such as have the FOOT vowel, and words such as have the STRUT vowel.

There is one further vowel sound, indicated above, which we must consider at this stage. It is the first vowel sound which occurs in most speakers' pronunciation of the word . This vowel is referred to as **schwa**; it is produced without lip rounding, and with the body of the tongue lying in the most central part of the vowel space, between high-mid and low-mid, and between back and front. Schwa is transcribed as [ə]. This vowel is typically even shorter than the short vowels we have just described, and it differs from those in that it may never occur in a stressed syllable (in , it occurs in the unstressed first syllable; in , it occurs in the unstressed second syllable; in , it occurs in the unstressed initial and final syllables). This vowel occurs in the speech of almost every speaker of English; in later chapters, we will consider its relation to English stressed vowels in more detail.

Exercises

1 Describe the position of the body of the tongue and the lips in the production of the following vowels:

[i] (cardinal vowel no. 1)

[u] (cardinal vowel no. 8)

[ɑ] (cardinal vowel no. 5)

2 Give an appropriate vowel symbol for the vowel in each of the following words, as you would utter them. Say (a) whether the vowel is rounded or not, (b) how back or front it is, and (c) how low or high it is (do this in relation to the cardinal vowels):

If you are discussing these exercises in a tutorial group, you may well already have begun to notice differences in the speech of the members of the group, depending on the accents they speak. Clearly, there is little point, if one has, say, a West Yorkshire or a New York City accent, in transcribing these words one were an RP or a GA speaker. What you should do is to try to work out (preferably with the help of a tutor) what the quality of each vowel is, and to adopt an appropriate phonetic symbol for that vowel, which you can then use consistently in your transcriptions. In due course, we will be examining accent variation in more detail.

Listen to sound files online

3 Listen to Track 3.1 at www.wiley.com/go/carrphonetics). Give a phonetic

transcription, with as much phonetic detail as possible, for each of the words you hear:

4

English Phonetics: Vowels (ii)

4.1 RP and GA Long Vowels

We noted that the RP/GA vowel in ([ʊ]) is shorter than that in ; we also said that it is less back and less high than that in . We will transcribe the vowel in as [u:], where the ‘ : ’ diacritic denotes vowel length. This is a high back rounded vowel, articulated closer to cardinal vowel 8 than [ʊ].

The RP/GA short vowel [ɪ], as in , which we described in chapter 3, is a fairly high, fairly front, unrounded vowel. It differs from the RP/GA vowel in , which is longer, more front and higher. We will transcribe this as [i:]; it is a high front unrounded vowel which is closer to cardinal vowel 1 than [ɪ].

It is worth noting that in RP and GA, when words such as and are uttered in isolation, they contain, respectively, the vowels [u:] and [i:], so that is pronounced in the same way as and . But ‘function’ words like and (which are not nouns, adjectives or verbs) are often uttered without stress, in which case they may be uttered with a schwa ([ə]), or in a shortened form, as in (pronounced either as [təi:t] or as [tuiːt]) and (pronounced either [ʃəwɔ] or [ʃiwɔː]). The shortened form of [iː] is also found in various suffixes, as in the suffix in : [wɪti] and in the suffix in : [kʰwɪkli]. It occurs too in the unstressed syllable of words such as : [pʰɹɪti].[1]

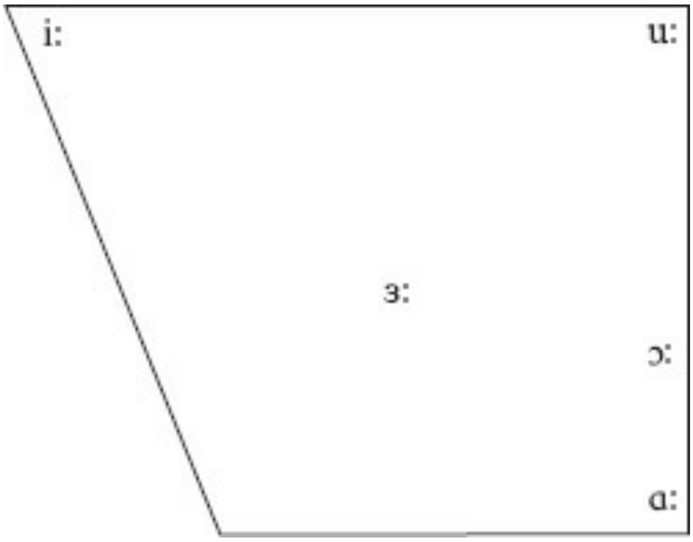
The RP vowel in and is longer than that in and ; it is a low-mid back rounded vowel, articulated closer to cardinal vowel 6 than is the [ɒ] in RP and . We will transcribe it as [ɔː]. This is also the vowel which GA speakers utter in words like (although the GA vowel is somewhat shorter than the RP vowel). Thus, although both GA and RP speakers distinguish between pairs such as and , GA has [ɑ] in whereas RP speakers have [ɒ]. In GA, words such as and , with an [ɹ] after the vowel, are typically uttered as [hɔɹs] and (see below on /oʊ/ in RP and GA).

The RP and GA short vowel [æ], as in , is, as we have seen, a fairly low, rather front, unrounded vowel. It differs from that in , which is a low back unrounded vowel, articulated in the region of cardinal vowel 5. The RP/GA vowel in is also longer than the RP/GA vowel in . We will transcribe it as [ɑː]. Thus, whereas RP has a three-way distinction between [ɒ], [ɑː] and [ɔː], GA has only a two-way distinction between [ɑ] and [ɔː]. We will return to this difference between the accents below.

RP and GA speakers utter a long vowel in words like , although GA speakers utter an [ɹ] in words such as these, while RP speakers do not. The articulation for this vowel is pretty much the same as that for schwa: it is central on both the high/low and front/back dimensions, and is unrounded. Unlike schwa, it appears in stressed syllables. We will transcribe it as [ɜ:].

We may depict the approximate areas of articulation of these vowels in the vowel space as follows:

(1) RP and GA long vowels



We will, following Wells (1982; see Suggested Further Reading), refer to [i:] as the FLEECE vowel, [u:] as the GOOSE vowel and [ɜ:] as the NURSE vowel. Wells uses three key words for the [ɔ:] vowel: THOUGHT, FORCE and NORTH; we will see why at a later stage. Similarly, Wells uses three key words for the [ɑ:] vowel: START, BATH and PALM. One of the reasons for this is that in words of the set BATH, GA has [æ], whereas RP has [ɑ:], whereas in words of the sets START and PALM, both GA and RP have [ɑ:].

4.2 RP and GA Diphthongs

In all of the RP and GA vowel sounds we have considered thus far, the articulators remain more or less in the same position throughout the articulation of the vowel. This means that the vowel quality (the acoustic effect created during the articulation of the vowel) remains more or less constant. That kind of vowel is a **monophthong**. However, there are vowel sounds in which this is not the case. This kind of vowel sound, called a **diphthong**, entails some kind of change of position of the articulators during its production, and thus a change in the vowel quality produced. A diphthong is a vowel whose quality changes . A diphthong is not simply a sequence of two vowels. For instance, in both the RP and the GA pronunciations of the word ([si:ɪŋ]), the vowel [i:] is followed by the vowel [ɪ], but the resulting sequence is not a diphthong, because the [i:] and the [ɪ] are not in the same syllable: has two syllables, the first of which ends in [i:] and the second of which begins with [ɪ].

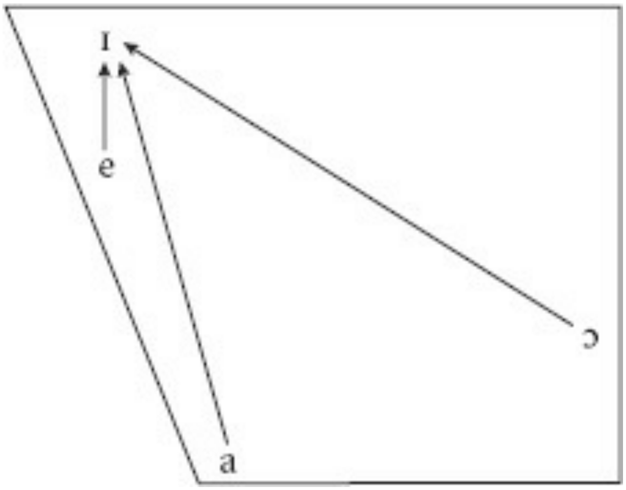
Let us begin with diphthongs which end in an [ɪ]-like quality. In the RP and GA pronunciations of words such as , etc., the vowel begins with an [a]-like quality (in the region of cardinal vowel 4) and ends in an [ɪ]-like quality. We will transcribe this as [aɪ].

In the RP and GA pronunciations of , etc., the vowel begins with an [e]-like quality (in the region of cardinal vowel 2) and ends in an [ɪ]-like quality. We will transcribe this as [eɪ]. In words such as , the GA pronunciation is a monophthongal [e], followed by an [ɹ].

In the RP and GA pronunciations of , etc., the vowel begins with an [ɔ]-like quality (in the region of cardinal vowel 6) and ends in an [ɪ]-like quality. We will transcribe this as [ɔɪ].

We may represent these diphthongs in the vowel space diagram as follows:

(2) RP and GA diphthongs ending in [ɪ]



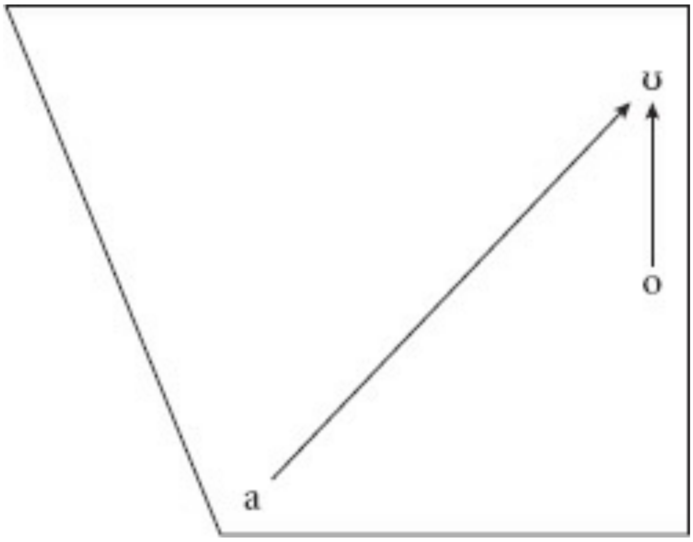
Wells uses the key word FACE for the [eɪ] diphthong, CHOICE for the [ɔɪ] vowel and PRICE for the [aɪ] vowel; we will follow this practice when it proves useful.

There are two diphthongs in RP and GA which end in an [ʊ]-like quality. The first of these begins with a low, rather back, unrounded quality. It is found in the RP and GA pronunciations of words such as . We will transcribe this diphthong as [aʊ].

The second of these diphthongs begins, among GA speakers, and among more conservative RP speakers, with an [o]-like quality. It occurs in words such as . We will transcribe this as [oʊ]. Among more modern RP speakers, words such as these are pronounced with an [əʊ]-like quality.[2] Words such as , uttered with the long vowel [ɔ:] but no [ɹ] in RP, are uttered with an [o] followed by an [ɹ] in GA.

These two diphthongs may be represented within the vowel space as follows:

(3) RP and GA diphthongs ending in [ʊ]



For these diphthongs, Wells uses the key word MOUTH for the [aʊ] vowel and GOAT for the [oʊ] diphthong.

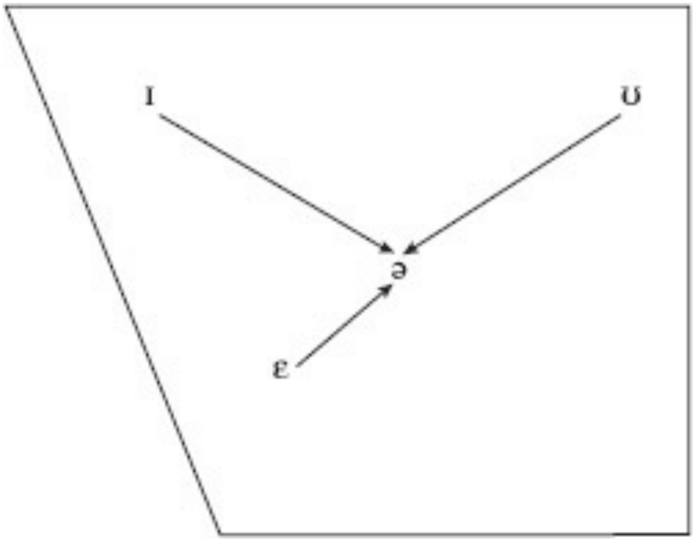
Many RP speakers utter a series of diphthongs which end in an [ə]-like quality, i.e. schwa. Since schwa is pronounced in the centre of the vowel space, these are often called **centring diphthongs**. The first of these diphthongs begins with an [ɪ]-like quality. It occurs in words such as and . We will transcribe this as [ɪə].

Another diphthong of this sort begins with an [ɛ]-like quality (in the region of cardinal vowel 3). This occurs in the RP pronunciation of words such as and . We will transcribe this as [ɛə]. Some RP speakers pronounce words of this sort with [ɛ:], a long vowel which is not a diphthong at all, but is more like a long version of a vowel in the region of cardinal vowel 3.

A third such diphthong begins with an [ʊ]-like quality, and occurs in words such as and . We will transcribe this as [ʊə]. Some RP speakers pronounce some of these words (e.g.) as a long monophthong in the region of cardinal vowel 6. If you encounter this, you may reasonably transcribe it as [ɔ:].

These three diphthongs may be represented as follows:

(4) RP diphthongs ending in [ə] (centring diphthongs)



These are called centring diphthongs since schwa is located at the centre of the vowel space. Wells uses the key words NEAR for the [ɪə] vowel, CURE for the [ʊə] vowel, and SQUARE for the [ɛə] vowel. These diphthongs are all absent in GA. Their presence in RP results from the loss of [ɹ] after vowels in the historical development of RP: the schwa is, as it were, the only remaining trace of the [ɹ] which once existed in the accents from which RP evolved, in the pronunciation of words such as and , which are pronounced [hi:ɹ], [heɹ] and [pʰu:ɹ] in GA. In RP, it is common to find a monophthongal variant. For the SQUARE vowel in contemporary RP, it is common to find a long monophthong: [ɛ:]. For the CURE vowel, many words of that lexical set are now pronounced by RP speakers with the long monophthong [ɔ:], as in the word , pronounced [ʃɔ:].

Notes

1 Accents of English vary with respect to the final vowel in words such as and . Some have an [i]-type vowel, while others have an [ɪ]-type vowel. The former vowel is sometimes said to be ‘more tense’ than the latter, and accents with [i] in these words are sometimes described as having ‘ tensing’. The latter term is due to Wells (1982; see Suggested Further Reading).

2 Among younger speakers of RP, these diphthongs are frequently uttered with a fronted unrounded second element; we could transcribe these pronunciations as [əɨ] and [aɨ], where the second symbol denotes a relatively high central unrounded vowel, as in : [kʰəɨk] and : [daɨn]. The effect is to make sound rather like , and rather like . There are many other such pairs; the principal point about them is that the pairs are still distinct, but less markedly so than in the past.

Exercises

1 Transcribe phonetically the vowel which you utter in each of the following words:

(See the note under chapter 3, exercise 2.)

2 For native speakers of English: the vowels in the following words are normally diphthongs in RP. For each word, phonetically transcribe the vowel as you would

normally say it, with the appropriate symbol. If it is a diphthong in your speech, describe the initial and final lip and tongue configurations. If it is not a diphthong, say how back/front it is, how high/low and whether it is rounded.

(Again, see the note under chapter 3, exercise 2.)

Listen to sound files online

3 Listen to Track 4.1 at www.wiley.com/go/carrphonetics. Transcribe, with as much phonetic detail as possible, each of the words you hear:

5

The Phonemic Principle

5.1 Introduction: Linguistic Knowledge

We have been dealing, thus far, with phonetics, that is (as we have defined it), with the study of human speech sounds (although we have dealt exclusively with phonetics, and in particular, exclusively phonetics, ignoring important facts about the acoustic properties of the speech sounds we have been discussing). We will, henceforth, be dealing with , as well as phonetics. Phonology, we will claim, is to do with something more than properties of human speech sounds . Phonology is the study of certain sorts of mental organization. In particular, it is the study of certain types of mental category, mentally stored representations, and generalizations concerning those categories and representations. On this view, phonology is the study of human speech sounds , although phonetics and phonology are inextricably intertwined. The point of this chapter is to demonstrate what the difference between the two is, and to begin to introduce the reader to the phonology of

English. Let us begin by considering some general questions concerning what it is to know a language.

Let us assume that when we say that someone knows a language, in the sense of being a native speaker of that language, he or she is in a certain mental state, or possesses a certain sort of linguistic knowledge. Knowledge of a native language is, apparently, largely unconscious knowledge. It appears to contain semantic knowledge (to do with the meanings of words, phrases and sentences) and syntactic knowledge (to do with the syntactic categories of words, with the structure of phrases and sentences and with the syntactic relations between words, phrases and clauses). We know that this is so, since speakers are able to make syntactic and semantic judgements, based on that knowledge. For instance, a native speaker of English can judge that

is an English sentence, and that

is not. The speaker knows, again intuitively, that the difference between the two amounts to more than the difference between the mere presence of the word as opposed to the presence of the word . He or she also knows intuitively (not necessarily fully consciously) in what sense is ambiguous, and in what sense the two interpretations of that sequence of words differs in structure and meaning from , over and above the superficial fact that one sequence contains and the other . That knowledge is clearly unconscious knowledge, since we require no instruction to be able to make such judgements, and we can make them in the absence of any conscious knowledge whatsoever of the syntax and semantics of English (one could make such judgements even if one had not the faintest idea of what a noun or a verb might be, or what the syntactic categories of and might be).

We will take the view in this book that a speaker's (largely) unconscious knowledge of his or her native language(s) must also contain phonological knowledge. One of the reasons many linguists take this view is that speakers can make judgements which, it is claimed, are in some sense parallel to those made with respect to syntactic states of affairs. For instance, a native speaker of English can tell how many syllables there are in a word without having the faintest idea, consciously, as to what a syllable is. This shows that the native speaker has the ability to recognize syllables, even if the recognition of syllables lies below the level of consciousness. In a similar fashion, it is claimed, a native speaker of English can tell that the sequence of segments [blʌg], considered as an utterance of a word, is an English sequence, whereas the sequence of segments [t^{h}lʌg] is not, despite the fact that she or he may well never have heard either sequence in her or his life. Let us postulate that, in making such judgements, the native speaker of English gains access to a kind of unconscious knowledge which constitutes 'the phonology of English'.

Our task, in this book, will be to begin to consider, in an elementary way, what form that knowledge takes. The discipline of phonology, under this view, differs from that of phonetics, since it is the study, not of speech sounds , but of mental abilities and largely unconscious mental states. Clearly, the phonologist must pay close attention to speech sounds and their properties; they will constitute much of the evidence the phonologist brings to bear on his or her hypotheses about speakers' unconscious phonological knowledge, but they do not constitute his or her object of inquiry as such.

5.2 Contrast vs Predictability: The Phoneme

Let us begin by considering voiceless unaspirated and voiceless aspirated stops in English and Korean. Speakers of most accents of English habitually utter both aspirated and unaspirated voiceless stops. The following English data exhibit both of these.[1]

(1) Aspirated and unaspirated voiceless stops in English
(a)
[ˈpʰuːɫ]
pool
(b)
[əˈpʰɪə]
appear
(c)
[ˈspɜːt]
spurt
(d)
[dəˈspaɪt]
despite
(e)
[ˈtʰɒp]
top
(f)

[əˈtʰæk]
attack
(g)
[ˈstɒp]
stop
(h)
[dəˈstɹɔɪ]
destroy
(i)
[ˈkʰɪlɪŋ]
killing
(j)
[əˈkʰɹu:]
accrue
(k)
[ˈskoʊɫd]
scold
(l)
[dɪˈskʌvə]
discover

The diacritic which precedes certain symbols in these data (the one which precedes the ‘p’ symbol in [ˈpʰu:ɫ] indicates the beginning of a stressed syllable.

From these data, it appears that voiceless stops are aspirated when they are at the beginning of a stressed syllable, as in and , but unaspirated when preceded by a voiceless alveolar fricative, as in . That is, in these data, wherever the unaspirated voiceless stops appear, the aspirated ones do not, and vice versa. Compare the English data with the following data from Korean:

(2) Aspirated and unaspirated voiceless stops in Korean

(a)
[pʰul]
'grass'

(b)
[pul]
'fire'

(c)
[tʰal]
'mask'

(d)
[tal]
'moon'

(e)
[kʰɛda]
'dig'

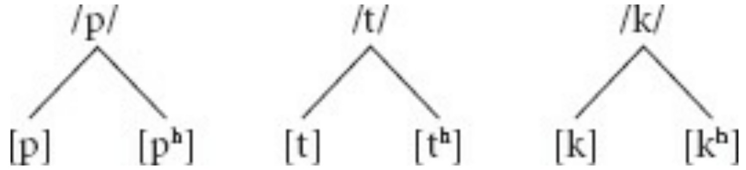
(f)
[kɛda]
'fold'

In these Korean data, aspirated and unaspirated voiceless stops may occur in the same place (at the beginning of a word). The range of places within a word which a given sound may occur in is called its **distribution**. In the English data we have looked at, the distribution of unaspirated and aspirated stops is : where you get one kind of stop, you never get the other. This is called **complementary distribution**.

Furthermore, if we take, say, the stops [t] and [tʰ] in the English data, it is clear that they are **phonetically similar**: both are stops, both are voiceless, both are alveolar. And yet, for most speakers of English, the alveolar stops in, say, and sound the same, despite the fact that the former is unaspirated and the latter aspirated. For the English speaker, these two phonetically distinct sounds ‘count as the same thing’. We cannot say, without contradiction, that they are simultaneously ‘the same sound’ and ‘not the same sound’. What we will say is that, while they are ,

they are . That is, the two types of stop correspond to, are interpreted as belonging to, a single mental category. We will refer to such a category as a **phoneme**. The English speaker interprets the six phonetic segments [p], [p^h], [t], [t^h], [k] and [k^h] in terms of only three phonemes: /p/, /t/ and /k/. We may depict this as follows:

(3) English voiceless stop phonemes



The top line here represents the three voiceless stop phonemes (mental categories) in terms of which the six types of phonetic segment are perceived. The relationship between phonemes and their associated phonetic segments is one of **realization**, so that the phoneme /p/, for instance, is realized as [p] after a voiceless alveolar fricative, and as [p^h] elsewhere. The most important point is that, on the data we have seen thus far, aspiration or the lack of it is entirely predictable in English: there is a generalization, expressible as a general rule, as to the contexts in which voiceless stops will and will not be aspirated. For most accents of English, this generalization is one that is internalized by children when they acquire English as their native language. The generalization forms part of what native speakers know in knowing their native language, even if that knowledge is largely knowledge. Realizations of a phoneme which are entirely predictable from context are called its **allophones**. We therefore say that [p] and [p^h] are allophones of the /p/ phoneme in most accents of English. We are claiming that native speakers of English possess phonemes (which are mental categories) and phonological generalizations or rules as part of their (largely unconscious) knowledge of their native language, and that native speakers perceive the allophones they hear in terms of those categories and generalizations.

Compare the English situation with the Korean one. It is clear that the distribution of aspirated and unaspirated voiceless stops in Korean is : there is at least one place (at the beginning of words) in which either type of sound may occur. This kind of distribution is referred to as **parallel distribution**, where ‘parallel’ means ‘overlapping to some degree’.

Furthermore, the distinction between aspirated and unaspirated voiceless stops can make a crucial difference in Korean: when the Korean speaker says [p^{h}ul], it does not mean the same thing as [pul]. The difference between the two sounds is said to be semantically **contrastive**. Pairs of words which differ with respect to only one sound are called **minimal pairs**. Their existence is important, since they demonstrate that the two sounds in question are both in parallel distribution and semantically contrastive.

We therefore want to say that, unlike the English speaker, the Korean perceives the six aspirated and unaspirated voiceless stops [p], [p^h], [t], [t^h], [k] and [k^h] in terms of six different mental categories. That is, [p], for instance, is a realization of the /p/ phoneme, whereas [p^h] is a realization of a distinct $/p^h/$ phoneme. We may depict (part of)[2] the Korean system thus:

(4) Some Korean voiceless stop phonemes

$/p/$	$/p^h/$	$/t/$	$/t^h/$	$/k/$	$/k^h/$
|	|	|	|	|	|
$[p]$	$[p^h]$	$[t]$	$[t^h]$	$[k]$	$[k^h]$

The distinction between aspirated and unaspirated voiceless stops is in Korean but in English. Both English and Korean speakers habitually utter both aspirated and unaspirated voiceless stops. On the phonetic level, the two languages are therefore equivalent as far as bilabial, alveolar and velar voiceless stops are concerned. But at the phonemic level (the mental level), the two languages are quite distinct: the Korean speaker has six mental categories where the English speaker has only three. As far as voiceless stops are concerned, Korean speakers have twice as many phonemic contrasts as English speakers. The difficulty which the English speaker encounters in learning to pronounce and perceive Korean voiceless stops is therefore a one; it is a difficulty, not a purely articulatory one.

This is not to deny that there can be purely articulatory difficulties in learning to speak another language (difficulties in articulating new types of sound which one is not in the habit of articulating). For instance, most speakers of Japanese who are learning to speak English will have to learn to pronounce the sound [l], which they are not in the habit of pronouncing. When learners of a foreign language face this task, they often utter a sound from their native language which is similar to the target sound: in this case, the tap [ɾ] which, like [l], is voiced and alveolar. Similarly, a speaker of French who is trying to master the English sound [ð] will often utter the voiced alveolar fricative [z] or the voiced dental stop [d̪], which she or he is used to uttering in her or his native language. The former is similar to the target sound in being a voiced fricative, while the latter is similar in being a voiced dental sound. Such problems with the pronunciation of foreign languages are widespread. But they are distinct in kind from the kind of problem we have just discussed.

We need not deny either that there may be difficulties in the pronunciation of a foreign language which involve purely articulatory phonological difficulties. For instance, the English speaker who is learning Korean must learn to articulate a third kind of stop which is distinct from voiced stops, aspirated voiceless stops and unaspirated voiceless stops. These are the voiceless stops of Korean which are articulated with 'glottal tension': during their production, the vocal cords do not vibrate, but nor are the vocal cords spread apart, as they are for the voiceless aspirated stops; rather, the vocal cords are constricted.

The English speaker must also learn to (in a sense) the distinction between all three sorts of stop in Korean; since the glottally constricted voiceless stops are a new category of sound, they may seem to the English speaker to sound like stops he or she is more used to hearing (voiced stops, for instance). And that is a phonological difficulty, added to the purely articulatory one which the English speaker also has. However, it is clear from the data we have looked at here that there is a type of difficulty which is exclusively phonological, and it is that kind of difficulty which justifies our making a distinction between the kind of articulatory phonetics discussed in the preceding chapters, which constitutes the study of the articulation of speech sounds in and of themselves, and phonology, the study of the system of mental categories in terms of which we interpret those speech sounds.

In examining the phonological differences between Korean and English voiceless stops, we have adopted what is known as **the phonemic principle**, which consists of two sets of two criteria, as follows:

(5) The phonemic principle

Two or more sounds are realizations of phoneme if:

(a) they are in complementary distribution

and

(b) they are phonetically similar.

Two or more sounds are realizations of phonemes if:

(a) they are in parallel (overlapping) distribution

and

(b) they serve to signal a semantic contrast.

It is on the basis of the phonemic principle that we say that phonetic differences involving aspiration are allophonic in English but phonemic in Korean.

We have just seen a case where the Korean speaker has more phonemic contrasts than the English speaker. Let us now look at another set of data where the converse is the case. Native speakers of some varieties of Scottish English habitually utter the speech sounds we have represented as '[ɾ]' and '[l]', i.e. the voiced alveolar tap and the voiced lateral alveolar approximant (as in and). So do speakers of Korean. Here are some examples of Scottish English and Korean words which contain those sounds:

(6) [ɾ] and [l] in Scottish English and Korean

	Scottish English			Korean	
(a)	[læm]	lamb	(b)	[mul]	'water'
(c)	[ɾæm]	ram	(d)	[mulkama]	'place for water'
(e)	[lɪp]	lip	(f)	[muɾe]	'at the water'
(g)	[ɾɪp]	rip	(h)	[mal]	'horse'
(i)	[bɛɾi]	berry	(j)	[malkama]	'place for horse'
(k)	[bɛli]	belly	(l)	[maɾe]	'at the horse'

While speakers of Scottish English and Korean habitually utter both sounds, we can predict that many native speakers of Korean who are learning to speak this variety of Scottish English would find the distinction between [l] and [ɾ], when they speak Scottish English, rather difficult to get the hang of. On the face of it, this is puzzling because, as we have just said, Korean speakers have no difficulty in uttering the two sounds, and may well have uttered many thousands of them, long before beginning to learn Scottish English. So wherein does the problem reside? One possibility that can be immediately discounted is the suggestion that Korean speakers are encountering some kind of physical, articulatory difficulty: it is clearly the case, as we have seen, that either of the sounds is new to them.

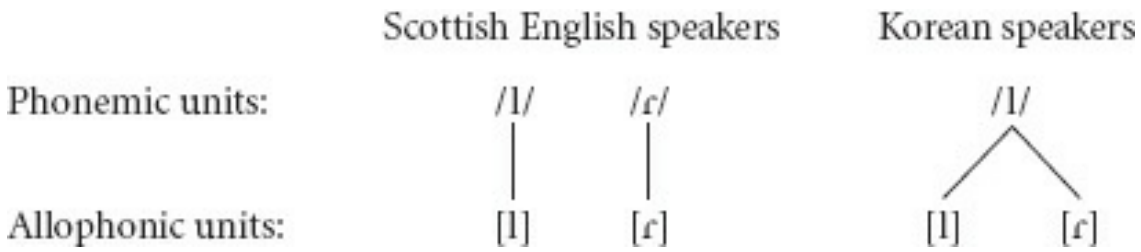
The difficulty is of a nature, and if one examines the table of data in (6) above, it is clear that, in Scottish English, the two sounds may occur in the same places within a word, e.g. at the beginning of words, or between vowels. Furthermore, two words may differ with respect to the segments [ɾ] and [l]: there are minimal pairs involving the two sounds ([ɾæm] vs [læm], for instance). In this variety of Scottish English, [ɾ] and [l] are in parallel distribution and can function to signal a semantic contrast. It is important to bear in mind that, when we say that a phonetic difference is contrastive, we refer to a contrast, and to a phonetic difference between the sounds.

In Korean, the distinction between [ɾ] and [l] can be contrastive, since [ɾ] and [l] may never occur in the same place. They are in complementary distribution: where one occurs, the other never does, and vice versa. Specifically, [ɾ] in Korean occurs between vowels but nowhere else, whereas [l] occurs between vowels, but may occur elsewhere. Because of this, it is impossible to find minimal pairs involving these two sounds in Korean. The two sounds are also phonetically similar: both are voiced and both entail a closure made between the centre of the tongue blade and the alveolar ridge. Therefore the two sounds are realizations of the same phoneme in Korean.

In this variety of Scottish English, there is a phonemic /ɾ/ vs /l/ contrast. In Korean, on the other hand, there is no such phonemic contrast: whereas this variety of Scottish English has /ɾ/ vs /l/, Korean has only one phoneme: /l/, which has two allophones, [ɾ] and [l]. Put another way, the difference between the sounds [ɾ] and [l] is in Scottish English, whereas the difference between [ɾ] and [l] is in Korean. Speakers of this variety of English perceive [ɾ] and [l] in terms of two mental categories, whereas Korean speakers perceive them in terms of mental category. In Korean, the phoneme /l/ is realized as [ɾ] between vowels, and is realized as [l] elsewhere.

We may depict this phonological difference between this variety of Scottish English and Korean as follows:

(7) The phonemic status of [ɾ] and [l] in Scottish English and Korean



We have now shown where the Korean speakers' difficulty resides: at the level of their (largely) unconscious knowledge of their language. As far as these segments are concerned, Korean and this variety of Scottish English do not differ at the allophonic level: both have [ɾ] and [l]. But they differ at the level: the Scottish English speaker has a mental distinction which the Korean speaker lacks; the Korean speakers' problem is thus (specifically, perceptual) in nature, not articulatory.

We have said that it is entirely predictable which allophone of the Korean /l/ phoneme will occur in a given context. We may say that there is a **phonological generalization** governing the occurrence of the allophones, which the native speakers of Korean have unconsciously grasped, and which forms part of their linguistic knowledge. We may express that generalization in terms of a **phonological rule**, as follows:

(8) /l/ realization in Korean

/l/ is realized as [ɾ] between vowels.

As we will see, the linguistic knowledge of native speakers contains many such generalizations. As far as [ɾ] and [l] are concerned, the phonological knowledge of the Korean speaker and that of the Scottish English speaker differ in two respects:

(a) the Scottish English speaker has a phonological distinction which the Korean speaker lacks, and (b) the Korean speaker possesses a phonological generalization which the Scottish English speaker lacks. Phonological knowledge therefore consists of, among other things, phonological categories and phonological generalizations.

In several varieties of English, the /l/ phoneme also has allophones: 'clear l' ([l]) and 'dark l' ([ɫ]).[3] The following data show the typical distribution of these two sounds in those varieties:

(9) English 'clear l' and 'dark l'
(a)
[kʰlɛvə]
clever
(b)
[bɛɫz]
bells
(c)
[pʰleɪn]
plain
(d)
[tɹeɪɫ]
trail
(e)
[lʊk]
look
(f)
[pʰʊɫ]
pull
(g)
[lɔ:]

law
(h)
[bɔːɫz]
balls
(i)
[laɪ]
lie
(j)
[pʰaɪɫ]
pile

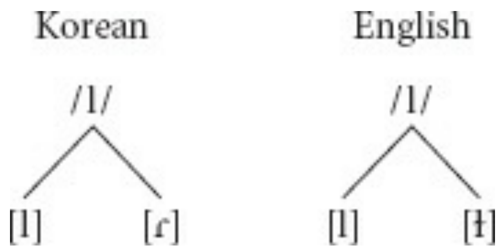
One way of stating the distribution of the allophones is to say that ‘clear l’ occurs immediately before vowels, whereas ‘dark l’ occurs immediately after vowels. We may state the relationship between the /l/ phoneme and its clear and dark allophones in terms of the following rule (which we will later express in terms of syllable structure):

(10) /l/ realization in English

/l/ is realized as [ɫ] immediately after a vowel.

We may depict the realizations of Korean /l/ and /l/ in certain varieties of English as follows:

(11) /l/ realizations in Korean and English



5.3 Phonemes, Allophones and Contexts

We have said that the allophones of a phoneme are predictable realizations of that phoneme. We can predict which allophone will occur, given a specific context. The sorts of context we have cited are, in some cases, rather general. For instance, in the Korean data we considered, we saw that aspirated and unaspirated voiceless stops may both occur at the beginning of a word. We also saw, in the Korean data that we looked at, that Korean /l/ is realized as [ɾ] between vowels. ‘At the beginning of a word’ and ‘between vowels’ are quite general contexts. So is ‘at the end of a word’, or ‘before a consonant’, or ‘after a vowel’.

In other cases, the contexts we need to refer to are more specific. For instance, in the English data we looked at, we saw that the unaspirated voiceless stops occurred after a voiceless alveolar fricative. In many cases, there appears to be some kind of phonetic connection between the context in which an allophone occurs and the nature of the allophone itself. Let us consider an example.

In many accents of English, the /ɹ/ phoneme has two realizations: [ɹ] and [ɹ̥] (in which the subscript diacritic denotes voicelessness). The following data exemplify this:[4]

(12) Voiced and voiceless allophones of /ɹ/ in English

(a)

[tʰɹ̥aɪ]

try

(b)

[əɹeɪ]

array

(c)

[pʰɹ̥u:v]

prove

(d)
[gɹoʊ]
grow
(e)
[kʰɹ̥eɪv]
crave
(f)
[bɹeɪk]
break
(g)
[fɹ̥iː]
free
(h)
[dɹɪŋk]
drink
(i)
[θɹ̥iː]
three
(j)
[bæɹoʊ]
barrow

It is clear that the voiced and voiceless alveolar approximants are in complementary distribution: the voiceless one appears only after voiceless consonants, and the voiced one appears elsewhere. The question is whether we should say that there is a voiced alveolar approximant phoneme which is realized as a voiceless allophone after voiceless consonants, or that there is a voiceless alveolar approximant phoneme which is realized as a voiced approximant after voiced consonants and between vowels. We choose the former claim, since it is more phonetically natural: approximants are normally voiced.

Additionally, we can make phonetic sense of the claim that a voiced phoneme has a voiceless realization when it follows voiceless consonants: the realization is assimilating to the preceding segment (it is becoming more like an adjacent segment).

Let us consider another case of this sort. In many accents of English, there are stops which are articulated in front of the velar place of articulation, close to the hard palate. The following data exemplify this ([c] and [ɟ] represent a voiceless and a voiced palatal stop, respectively):

(13) Velar and palatal stops in English

(a)
[kʰuːɫ]
cool
(b)
[cʰiːp]
keep
(c)
[kʰoʊɫ]
coal
(d)
[cʰiːn]
keen
(e)
[kʰɒp]
cop
(f)
[cʰɪt]
kit
(g)
[kʰɒːt]

cart
(h)
[scɪp]
skip
(i)
[guːɫ]
ghoul
(j)
[ɟɪə]
gear
(k)
[goʊɫ]
goal
(l)
[ɟiɫ]
gill

Once again, the two segment types are in complementary distribution: the advanced, palatal articulations occur before high front vowels, and the velar ones occur elsewhere. We postulate a /k/ phoneme which is 'fundamentally' velar in its place of articulation, but which has a fronted or advanced realization before high front vowels. This makes phonetic sense: high front vowels are palatal articulations (the articulators are the front of the tongue and the hard palate), so we can say that the velar phoneme is assimilating to the following vowel when it is a high front vowel.

We are adopting the view that phonemes often have a kind of 'default' or 'basic' phonetic realization, and that it is this realization which will occur in the absence of specifiable contexts which 'shift' the realization from its default one.

5.4 Summing Up

In this chapter, we have begun to distinguish between phonetics, defined as the study of speech sounds , and phonology, the study of the system of mental representation, categories and generalizations to which those sounds are related. Native speakers of a language tend to take its phonological system for granted. Speakers of English, for instance, think it perfectly obvious that [ɹ] and [l] are quite distinct, despite the fact that they are, phonetically, very similar. Equally, speakers of English cannot easily see that [p] and [pʰ] are different, despite the fact that they are. This chapter has sought to show that what underlies these perceptions is the phonological system of the native language, as distinct from, if intimately related to, the set of speech sounds uttered by native speakers of the language. What sounds one takes to be 'the same' or 'different' depend to a large extent on the system of mental categories which constitutes one's native language phonology. But it is clear that phonetics and phonology are intimately connected.

The extent to which our mentally stored system of language-specific phonological categories governs our perception of a stream of speech sounds was well expressed by the linguist Edward Sapir, who worked with North American Indian languages in the early twentieth century:

> the unschooled recorder of language, provided he has a good ear and a genuine instinct for language, is often at a great advantage as compared with the minute phonetician, who is apt to be swamped by his mass of observations. I have already employed my experience in teaching Indians to write their own language for its testing value in another

> connection. It yields equally valuable evidence here. I found that it was difficult or impossible to teach an Indian to make phonetic distinctions that did not correspond to ‘points in the pattern of his language’, however these differences might strike our objective ear, but that subtle, barely audible, phonetic differences, if only they hit the ‘points in the pattern’, were easily and voluntarily expressed in writing. In watching my Nootka interpreter write his language, I often had the curious feeling that he was transcribing an ideal flow of phonetic elements which he heard, inadequately from a purely objective standpoint, as the intention of the actual rumble of speech.[5]

One can begin to appreciate the extent to which one’s native language phonemic categories affect one’s perception when one considers that any normal 6-month-old child, no matter what language he or she is beginning to acquire, can distinguish aspirated and unaspirated voiceless stops. Clearly, then, the aspirated/unaspirated difference is one that could in principle act as the basis for a phonemic distinction, and it clearly does act that way in many human languages. But a child who acquires a language (such as most varieties of English) in which the aspirated/unaspirated distinction is allophonic rather than phonemic will come to ignore that distinction at a certain level of awareness. Acquiring the phonology of one’s native language can therefore result in a kind of loss of perceptual discrimination, but only at one level of awareness: when a speaker of, say, South African English utters unaspirated stops instead of aspirated stops, this will often be noticed by a speaker of, say, RP, even if the RP speaker notices only that there is something different about the speech of the South African English speaker. Indeed, such differences can be quite striking to the speaker of a language in which unaspirated stops never occur word-initially before a stressed vowel. Such speakers, on being suddenly confronted by English spoken with, say, a Greek accent (on arrival, say, at a Greek airport) will typically think that a word such as (one of the London airports) is being pronounced as [gadwig]. In cases such as this, the English speaker not only perceives the fact that the stops in question are unaspirated, but also assigns them to the category of English voiced stops, because voiced stops in English are unaspirated, and word-initial and word-final voiced stops in English are barely voiced at all.

Both the native speaker and the adult learner of English can begin to develop an awareness of her or his own phonological system, and of the immense influence this has on one's perception of speech sounds, by comparing and contrasting languages which are phonetically identical (or nearly identical), but phonologically distinct, with respect to some set of sounds. The examples given in this chapter are designed to begin to induce this kind of awareness, as are the exercises which follow.

Notes

1 These data do not show the full range of places in which aspirated and unaspirated voiceless stops occur in most English accents. What we will have to say about their phonological status is therefore very much oversimplified. But the data will suffice to illustrate a valid point.

2 Korean has a third phonemic category of stops, which we discuss below.

3 There are also devoiced allophones of /l/; we ignore these here.

4 We indicate these devoiced sounds here, but henceforth we will not transcribe them using the ‘voiceless’ diacritic in cases where the devoicing phenomenon is irrelevant to the point being made.

5 Edward Sapir (1921), , New York: Harcourt Brace, p. 56.

Exercises

1 [d] and [ð] in English and Spanish

(a) English

1.
[dɛn]
den

2.
[ðɛn]
then

3.
[doʊz]
doze

4.
[ðoʊz]
those

5.
[dɛə]
dare

6.
[ðɛə]

their

7.

[ʌdə]

udder

8.

[ʌðə]

other

9.

[aɪdə]

Eider

10.

[aɪðə]

either

In many (not all) accents of English, [d] and [ð] are realizations of phonemes, as these data show. They are in parallel distribution (both occur at the beginnings of words and between vowels). They also function contrastively: there are minimal pairs involving the two sounds. We are therefore justified in postulating a /d/ vs /ð/ phonemic distinction for most accents of English.

(b) Spanish

Now consider the following Spanish data. (The voiced stop in question is in fact dental in Spanish. We overlook this fact here.) Is the distinction between [d] and [ð] phonemic or allophonic in Spanish? Justify your answer with evidence and argumentation.

1.

[dɐr]

'to give'

2.

[nɐðɐ]
‘nothing’
3.
[dɛβɛr]
‘to have to’
4.
[boðeɣɐ]
‘wine cellar’
5.
[diɐs]
‘days’
6.
[ɐblɐðo]
‘spoken’
7.
[bɐndɐ]
‘ribbon’
8.
[pɾɐðo]
‘meadow’
9.
[ɐndɐr]
‘to go’
10.
[poðɛr]
‘to be able’

2 Voiced stops in English and Korean

In Korean, /p/, /t/ and /k/ have allophones which are unreleased at the end of a word (as can be seen in the data below) or before another consonant. The /p/, /t/ and /k/ phonemes also have voiced stop allophones: [b], [d] and

[g]. Unlike English, Korean does not have voiced stop phonemes: [b], [d] and [g] are allophones of /p/, /t/ and /k/ in Korean. Examine the following data and say what contexts the voiced stop allophones occur in:

(a)(i)	[pul]	'fire'	(a)(ii)	[ibul]	'this fire'
(b)(i)	[tal]	'moon'	(b)(ii)	[idal]	'this moon'
(c)(i)	[kan]	'liver'	(c)(ii)	[igan]	'this liver'
(d)(i)	[pap˺]	'cooked rice'	(d)(ii)	[pabi]	'cooked rice' (subject)
(e)(i)	[tat˺]	'close'	(e)(ii)	[tadaɾa]	'close it!'
(f)(i)	[tʃɛk˺]	'book'	(f)(ii)	[tʃɛgi]	'book' (subject)

3 Glottal stops in English and Standard Arabic

The segment [ʔ] (glottal stop) occurs in the speech of most speakers of English, but there is no glottal stop (/ʔ/) in English, since [ʔ] never functions contrastively with any other segment. For instance, [kʰɛʔəɫ] and [kʰɛtəɫ] ( ) are not pronunciations of different words, but different pronunciations of the word. Below are some data from Standard Arabic. Is there a glottal stop phoneme in this language? Explain the reasoning behind your answer:

(a)
[faʔl]
'good fortune'
(b)
[fatl]
'twisting/twining'
(c)
[faʔɾ]
'rats'

(d)
[faːɾ]
‘it boiled’
(e)
[baʔs]
‘strength’
(f)
[baːs]
‘he kissed’
(g)
[buʔs]
‘misery’
(h)
[buːs]
‘bus’

Listen to sound files online

4 Further phonetic transcription practice

Listen to Track 5.1 at www.wiley.com/go/carrphonetics. Transcribe, with as much phonetic detail as possible, the words you hear on the recording, paying attention to details such as presence vs absence of aspiration, clear vs dark l, and devoiced allophones of /l/ and /ɹ/.

6

English Phonemes

6.1 English Consonant Phonemes

We have distinguished phonemes from phonetic segments, and have begun to formulate hypotheses about which phonemes might exist as part of the native speaker's phonological knowledge. Specifically, we have said that many English speakers have the consonant phonemes /l/, /ɹ/, /p/, /t/ and /k/, among others. We will shortly postulate a full system of consonant phonemes which English speakers have. But we must first be a little more precise about what we mean by 'English speakers'. Clearly, there are different varieties of English, which we will be considering in more detail later, and we will need some means of differentiating between them. Let us begin by considering a distinction which is often appealed to by linguists: that between **accent** and **dialect**. It is often said that differences in accent concern solely phonetic and phonological variation, whereas dialect differences involve more than this: they also include differences in vocabulary and syntax. This is a rather simplistic way of putting the distinction, and it is a distinction which is fraught with difficulties, but it will suffice for the present discussion.

We may exemplify the difference between accent and dialect as follows. Perhaps the most widely spoken (and written) English dialect is the 'prestige' dialect known as Standard English, which has its origins in the South East of England; this dialect is used widely, in Britain, in national radio and television, in the press, and indeed in most printed publications. It is possible to speak Standard English a New Zealand accent, a Tyneside accent, a New York City accent, or indeed accent of English. When this happens, we may say (simplifying somewhat) that the speaker is using the vocabulary and syntax of Standard English, while retaining the phonetics and phonology which constitutes the native accent.

Let us exemplify the difference between dialect and accent in a little more detail, as follows. Take the Standard English sentence

Uttered by a speaker with a Standard Scottish English (SSE) accent, the outcome would be:

(1) [jəłnɒʔbiebłtəp^{h}ʉʔðətʃɪłdɹənənðəflo:ɹ]

Now compare this with the same Standard English sentence uttered by a speaker of RP. The RP speaker might well utter:

(2) [jəłnɒʔbieɪbłtəp^{h}ʊʔðətʃɪłdɹənənðəflɔ:]

Both speakers are speaking Standard English (the syntax is the same, as is the vocabulary, if one excludes the phonological form of the morphemes in that vocabulary), but their accents differ: the SSE speaker's vowel sounds are not always identical to those of the RP speaker, and the SSE speaker utters an [ɹ] in , which the RP speaker does not.

Now let us imagine that the same SSE speaker wants to convey the same proposition, but speaking, this time, in a dialect other than Standard English: that of Lowland Scots.

The result might be:

(3) [jʌɫno:kɪmpʰɪʔðəbe:ɹɪnzonðəfɫe:ʌɹ]

(This might be written as

)

In (3), the syntax and vocabulary differ from that of Standard English; we may say that these are differences, and distinct in kind from the differences in accent which we noted between (1) and (2) above.

We will return to the matter of accent variation in a later chapter; for the moment, let us look at the consonant phoneme system shared by most varieties of English, which typically looks like this:

(4) English consonant phonemes

/p/

as in

/b/

as in

/t/

as in

/d/

as in

/k/

as in

/g/

as in

[tʃ]

as in

[dʒ]

as in
/θ/
as in
/ð/
as in
/f/
as in
/v/
as in
/s/
as in
/z/
as in
/h/
as in
/ʃ/
as in
/ʒ/
as in
/w/
as in
/l/
as in
/ɹ/
as in
/ɛ/
as in
/m/

as in
/n/
as in
/ŋ/
as in

The evidence comes partly in the form of the sorts of minimal pair cited in (4), such as ([mɛʒə]/[mɛʃə]) and ([saɪ]/[ʃaɪ]), but we have by no means presented all of the evidence here. Let us look briefly at some of the evidence for the three nasal stop phonemes postulated here:

(5) Evidence for English nasal stop phonemes
(a)
[mi:t]
meat
(b)
[ni:t]
neat
(c)
[moʊɫ]
mole
(d)
[noʊɫ]
knoll
(e)
[sɪn]
sin
(f)
[sɪŋ]
sing

(g)
[dɪm]
dim
(h)
[dɪn]
din
(i)
[wɪn]
win
(j)
[wɪŋ]
wing

It is clear that [m] and [n] are in parallel distribution: each may occur word-initially or word-finally. The distinction is also contrastive: it forms the basis for minimal pairs such as . It is also clear that [n] and [ŋ] are in parallel distribution: while [ŋ] does not appear in word-initial position, both may occur in word-final position. The distinction is also contrastive, as is shown by the existence of minimal pairs such as . The distinction between [m] and [ŋ] is contrastive too, as pairs such as show. We therefore have clear evidence for a three-way phonemic distinction between /m/, /n/ and /ŋ/. We will consider this analysis in more detail below. The main point to be made at the moment is that we postulate the existence of phonemes on the basis of evidence and argumentation; if phonemes are mental categories, they cannot be directly observed.

6.2 The Phonological Form of Morphemes

We have said that, in knowing a language, a speaker possesses largely unconscious linguistic knowledge, which subsumes semantic, syntactic and phonological knowledge. And we have said that the phonological units or categories we have called phonemes are part of that phonological knowledge. As we progress in this book, we will investigate the question of what other sorts of phonological knowledge speakers possess, besides phonemes alone. Let us begin this investigation by considering the internal structure of words. You will agree that the English word may be broken down into two component parts. Let us call those component parts **morphemes**. Then we may say that this word consists of a **root** morpheme and a plural morpheme (which, in this case, is a **suffix**). Let us say that words of this sort are **morphologically complex**: they consist of more than one morpheme. Let us say that a morpheme takes the form of a triple: a syntax, a semantics and a phonology. Take the morpheme : it has a syntax (it is a noun), a semantics (it means ‘cat’) and a phonology, which takes the form /kæt/; we will refer to this as the **phonological form** of the morpheme. The phonological form of a morpheme may, clearly, consist of more than one phoneme. Just as phonemes are mental objects, so the phonological form of this morpheme is a mental object: /kæt/ is a in the mind of a speaker, whereas the sequence [k^{h}æt] is a sequence.

Let us now consider the adjectives , ,

, , , , and

All consist of at least a **prefix** morpheme and a root morpheme (some of these words have a prefix, a root a suffix). Many speakers have the following pronunciations of these words:

(6)

(a)

[ɪmpʰɒsɪbɫ]

impossible

(b)

[ɪmbælənst]

imbalanced

(c)

[iɱfəlɪsɪtəs]

infelicitous

(d)

[ɪntʰændʒɪbɫ]

intangible

(e)

[ɪndɹɛkt]

indirect

(f)

[ɪnseɪn]

insane

(g)

[ɪŋkəɹɛkt]

incorrect

(h)

[ɪŋglɔːɹiəs]

inglorious

It is part of the native speaker's unconscious linguistic knowledge of English that these words all have the prefix. That prefix is one of the morphemes of English, and, like all morphemes in the language, has a syntax (it is a prefix), a semantics (it has a specific meaning) and a phonology. But what the phonological form of that morpheme? We know from the data that the suffix may be realized as [ɪm], [ɪɱ], [ɪn] or [ɪŋ]. It is clear, then, that the first phoneme in the prefix is /ɪ/ and the second phoneme is a nasal, but nasal phoneme? We claimed above that English has three nasal phonemes: /m/, /n/ and /ŋ/. So the phonological form of this prefix might be [ɪm], [ɪn] or [ɪŋ]. Let us consider [ɪŋ]. We could say that the /ŋ/ phoneme is realized as [n] before /t/, /d/ and /s/, and as [m] before [p] and [b]. This seems to make sense: we can say that, when the prefix is added to a root, the place of articulation of the nasal becomes identical to that of the first consonant in the root. Thus, it is alveolar when followed by an alveolar consonant (such as /t/, /d/ and /s/), labio-dental when followed by a labio-dental consonant (such as /f/), and bilabial when followed by a bilabial consonant (such as /p/ or /b/). This is the process of assimilation we referred to in chapter 2, in which one segment becomes similar, in some respect, to another when the two are adjacent. Here, the assimilation is in place of articulation. Further evidence that nasals in English undergo place of articulation assimilation is not hard to come by. Consider the following data, which are representative of the speech of many speakers of English:

(7) Nasal assimilation in English

(a)

[ʌŋkʰlɪə]

unclear

(b)

[ʌŋgɒdli]
ungodly
(c)
[ʌɱfɛə]
unfair
(d)
[ʌɱvælju:d]
unvalued
(e)
[ʌntʰɹu:]
untrue
(f)
[ʌndʌn]
undone
(g)
[ʌmbɛəɹəbəɫ]
unbearable
(h)
[ʌmbaɪəst]
unbiased

While the [ɪŋ] solution is plausible, it faces a difficulty: we might equally say that the phonological form of the morpheme is [ɪn], or [ɪm], and that, in either case, the nasal assimilates to a following consonant. On the evidence presented thus far, there is no non-arbitrary way of choosing between the three alternatives: each is as plausible as the others. The following data, however, allow us to make a non-arbitrary choice:

(8)
(a)
[ɪæktɪv]
inactive

(b)
[ɪnɐpɹətɪv]
inoperative
(c)
[ɪnɛfəbɫ]
ineffable
(d)
[ɪnədvaɪzəbɫ]
inadvisable
(e)
[ɪnɔ:dɪbɫ]
inaudible
(f)
[ɪneɪliənəbɫ]
inalienable

In each case, there is no consonant at the beginning of the root to which the nasal assimilate: each root begins with a vowel. From the fact that vowel-initial roots are realized with the [ɪn] form, we can therefore conclude that the phonology of the prefix takes the form [ɪn], and that the nasal does not change its place of articulation if the root-initial segment is a vowel or an alveolar consonant. Note that this is generally true of alveolar nasals in English, as the following data, involving the prefix seen in (7), suggest:

(9)
(a)
[ʌneɪdəd]
unaided
(b)
[ʌnət^{h}ɹæktɪv]
unattractive

(c)
[ʌnivɛntfəɫ]
uneventful
(d)
[ʌnɔːθədɒks]
unorthodox

We might, of course, have said that the morpheme in question has : /ɪm/, /ɪɱ/, /ɪn/ and /ɪŋ/, and that words such as , , and are each stored mentally with the appropriate prefix. There are two problems with this approach. Firstly, there is no independent evidence that there is an /ɱ/ phoneme in English ([ɱ] never functions contrastively with any other nasal). Secondly, even if there were no [ɪɱ] forms, we would be committed to claiming, under the 'several phonological forms' approach, that it is that the /ɪm/ form is attached to roots beginning with a bilabial consonant, the /ɪŋ/ form only to roots beginning with a velar consonant, and the /ɪn/ form only to roots beginning with alveolar consonants and vowels. But that is surely an implausible claim. So, for this sort of case at least, the idea that we should postulate more than one phonological form for a morpheme is deeply unattractive and implausible. Given the data we have seen thus far, it appears much more plausible to say that any given morpheme has a single phonological form. And if that is the case, then it is the phonologist's task to hypothesize as to what that form might be. In doing so, she or he will be guided by : the facts of the matter, since they are mental in nature, and thus not directly observable, will not be available for immediate inspection via the five senses.

In adopting this ‘one phonological form per morpheme’ approach, we are allowing that, while any given morpheme has only one phonological form, that phonological form may ‘correspond’ in some sense to a variety of different phonetic forms. In the case we have just looked at, the prefix has the phonological form [ɪn], but that in turn corresponds to four different phonetic forms: [ɪm], [ɪɱ], [ɪn] and [ɪŋ]. Phonologists refer to such phonetic forms as **alternants**: we say that there is an **alternation** between the four forms. Which alternant of a given morpheme occurs in a given word is entirely ; there is a generalization which captures that predictability, and we are able to express it in the form of a phonological rule, just as we did with [ɾ] and [l] in Korean. In our English case, the rule in question concerns nasals in general; it might be put informally as:

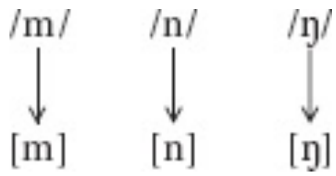
(10) The rule of nasal assimilation in English

> If the phonological form of a prefix ends in a nasal then that nasal will assimilate to the place of articulation of a following consonant.

We could have formulated this generalization as a formalized rule, or as some kind of constraint on the phonological form of morphemes in English. We will not go into the types of formalism required to express such generalizations, or inquire whether they are best expressed as rules or as constraints. The most important point is that native speakers appear to be in possession of generalizations of this sort, and that these appear to constitute a part of their largely unconscious phonological knowledge.

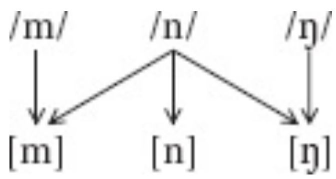
The data we have just considered also exemplify an important phenomenon: that of **phonemic overlapping**. On the basis of the data in (5), we postulated the following nasal stop phonemes, with the realizations shown:

(11)



However, we have allowed that the /n/ phoneme may also be realized as [m] if it precedes a bilabial consonant, or [ŋ] if it precedes a velar consonant. This means that a given occurrence of [m], for instance, may be either a realization of /m/, as in , or a realization of /n/, as in . That is, the /m/ and /n/ phonemes in their realizations. We may depict this as follows:

(12)



The question arises: how can the speaker of English tell whether a given [m] is a realization of /m/ or of /n/? The answer is that the phonological context allows the speaker to tell: an [m] which does precede a bilabial consonant will be a realization of the /m/ phoneme. The phonemic contrast between /m/ and /n/ is said to be **neutralized** before a bilabial consonant. Neutralization is the suspension of phonemic contrasts in one or more specifiable contexts.

6.3 English Vowel Phonemes

Accents of English vary considerably in their vowel phoneme systems and in the range of allophones that those phonemes have. We begin by depicting a set of postulated vowel phonemes for RP and GA. The RP and GA phonetic vowel qualities we presented and discussed in chapters 3 and 4 are typically contrastive for most speakers of those accents, and we may therefore postulate the following (stressed) vowel phonemes for RP and GA:

(13a)

RP vowel phonemes

Wells (1982; see Suggested Further Reading) lexical sets

/ʌ/

as in

STRUT

/ʊ/

as in

FOOT

/u:/

as in

GOOSE

/ɪ/

as in

KIT

/i:/

as in

FLEECE

/ɛ/

as in

DRESS

/eɪ/

as in

FACE

/ɒ/

as in

LOT

/oʊ/

as in

GOAT

/ɔː/

as in

NORTH, FORCE, THOUGHT

/æ/

as in

TRAP

/ɒː/

as in

START, BATH, PALM

/ɜː/

as in

NURSE

/ɔɪ/

as in

CHOICE

/aɪ/

as in

PRICE

/aʊ/

as in

MOUTH

/ɪə/

as in

NEAR

/ɛə/

as in

SQUARE

/ʊə/

as in

CURE

(13b)

GA vowel phonemes

Wells (1982) lexical sets

/ʌ/ as in

STRUT

/ʊ/ as in

FOOT, CURE

/uː/ as in

GOOSE

/ɪ/ as in

KIT

/iː/ as in

FLEECE, NEAR

/ɛ/ as in

DRESS, SQUARE

/eɪ/ as in

FACE

/oʊ/ as in

GOAT

/ɔ:/ as in
NORTH, FORCE, THOUGHT
/æ/ as in
TRAP, BATH
/ɒ/ as in
START, PALM, LOT
/ɜ/ as in
NURSE
/ɔɪ/ as in
CHOICE
/aɪ/ as in
PRICE
/aʊ/ as in
MOUTH

Again, what the set of RP or GA vowel phonemes might be is a matter for argumentation based on evidence and general theoretical considerations. For instance, we might have suggested that the second vowel in words like is a diphthong ([iu:]) and that, since forms a minimal pair with , , etc., then /iu:/ is an RP vowel phoneme. We will return, in due course, to this kind of question. We should also note that there is a further vowel phoneme which is not listed here: /ə/ (schwa), which differs from all the other phonemes listed above, since it does not occur in stressed position (as we noted in chapter 3). We will also return, in due course, to /ə/ and its relation to the phonetic segment [ə].

Like consonant phonemes, vowel phonemes may have allophones. For instance, speakers of many accents of English have two realizations of the vowel phoneme /i:/: [i:] and [i:ə].

The latter typically occurs before a velarized lateral (‘dark l’), as the following data show:

(14) Allophones of /i:/

(a)
[fi:t]
feet
(b)
[fi:əɫ]
feel
(c)
[di:p]
deep
(d)
[di:əɫ]
deal
(e)
[pʰi:k]
peak
(f)
[pʰi:əɫ]
peel
(g)
[si:m]
seem
(h)
[si:əɫ]
seal

We postulate /i:/ rather than /i:ə/ as the form of the phoneme, since we assume that the realization of the phoneme when it precedes a dark l is influenced by the dark l. The schwa articulation, which is retracted from the high front [i:] position, is a matter of the vowel articulation assimilating to the tongue body retraction in the dark l. In doing so, we appeal to the idea of **phonetic motivation**: our analysis is phonetically motivated in the sense that we can provide an articulatory reason for the /iː/ → [iːə] process, whereas we would be unable to provide any such motivation for a process in which /iːə/ → [iː] word-finally and before any consonant other than [ɫ].

We are also assuming, as we did in chapter 5, that the phoneme /l/ has two allophones, [l] and [ɫ]. We said there that there is an /l/ realization rule: /l/ is realized as [l] immediately vowels, and as [ɫ] immediately vowels.

These two claims appear to commit us to the idea that the rule governing the occurrence of [ɫ] must, in some sense, 'precede' the rule governing the phoneme /iː/, since we are claiming that [ɫ] only ever arises as a result of the application of the /l/ rule, and that [iːə] only ever arises when an [ɫ] follows. We may depict this claim about the interaction of the two rules as follows:

(15)

/fiːl/

/l/ rule
fiːɫ

/iː/ rule
fiːəɫ

This kind of depiction is referred to as a **derivation**: the phonetic realization of the phonological form /fiːl/ is derived from that phonological form by means of the of rules. This way of looking at the relationship between phonological forms and their phonetic realization therefore commits us to the idea of **rule ordering**, and thus to a rule-based, derivational view of phonological organization. While we will not pursue this conception of phonological organization in any depth, it is as well to acknowledge that we are implicitly assuming such a conception.

Another example of vowel allophones in English concerns vowel length. In many languages, vowel length is phonemic. For instance, in Limbu (spoken in Nepal), [sapma] means ‘to write’, but [sa:pma] means ‘to flatter’. Similarly, in Malayalam (spoken in Southern India), [ciri] means ‘smile’ but [ci:ri] means ‘shrieked’. But in other languages, vowel length is allophonic. In Scottish Standard English (SSE), for instance, some (not all) of the vowel phonemes have long and short allophones:

(16) Long and short vowels in Scottish Standard English

[hi:v]	heave	[bɹi:ð]	breathe	[bɹi:z]	breeze
[bi:ɹ]	beer	[bi:]	bee		
[bif]	beef	[hiθ]	heath	[fɫis]	fleece
[diɫ]	deal	[bit]	beat		
[mʉ:v]	move	[smʉ:ð]	smooth	[ɫʉ:z]	lose
[bʉ:ɹ]	boor	[blʉ:]	blue		
[hʉf]	hoof	[tʰʉθ]	tooth	[ɫʉs]	loose

The long allophones of the phonemes /i/, /u/ and /ai/ occur in the following contexts: at the end of a word or before one of the following voiced consonants: /v/, /ð/, /z/ and /ɹ/. Vowel length is therefore allophonic, rather than phonemic, in SSE.

Listen to sound files online

We considered the case of the neutralization of the contrast between the consonant phonemes /m/ and /n/ before bilabial consonants. Neutralization of contrast between vowel phonemes is also possible. An example comes from the contrast between /æ/, /ɛ/ and /eɪ/ in GA: these three phonemes are all realized as [ɛ] before the /ɹ/ phoneme. Listen to the sentence on Track 6.1 at www.wiley.com/go/carrphonetics. The GA speaker utters [ɛ] in each of the three words: they are homophones in GA. The other speakers are speaking SSE and RP respectively: each has distinct realizations for each of the three phonemes: the neutralization does not occur in SSE or RP.

Exercises

1 Phonemic contrasts in GA, SSE and RP

The following data sets are from General American (GA), Scottish Standard English (SSE) and RP (Received Pronunciation). On the basis of presence of minimal pairs in one variety vs absence of such pairs in another variety, identify any phonemic contrasts which are present in GA, but not in SSE, and vice versa. Then identify any phonemic contrasts which are present in RP, but not in GA, and vice versa. Finally, identify any phonemic contrasts which are present in SSE, but not RP, and vice versa. Assume that the data given here are fully representative of the varieties in question.

	GA	SSE	RP	
(a)	[weɪlz]	[ʍeɫz]	[weɪlz]	whales
(b)	[weɪlz]	[weɫz]	[weɪlz]	Wales
(c)	[lɑk]	[ɫɔx]	[lɒk]	loch
(d)	[lɑk]	[ɫɔk]	[lɒk]	lock
(e)	[wɪn]	[ʍɪn]	[wɪn]	whin
(f)	[wɪn]	[wɪn]	[wɪn]	win
(g)	[pʰu:ɫ]	[pʰʉɫ]	[pʰu:ɫ]	pool
(h)	[pʰʊɫ]	[pʰʉɫ]	[pʰʊɫ]	pull
(i)	[sæm]	[sɐm]	[sæm]	Sam
(j)	[sɑm]	[sɐm]	[sɑ:m]	psalm
(k)	[hɔɹs]	[hɔɹs]	[hɔ:s]	horse
(l)	[hoɹs]	[hoɹs]	[hɔ:s]	hoarse
(m)	[kʰɔ:t]	[kʰɔt]	[kʰɔ:t]	caught
(n)	[kʰɑt]	[kʰɔt]	[kʰɒt]	cot

2 Nasal stops in English and Spanish

Recall that data such as the following led to our decision to postulate three nasal stop phonemes in English (/m/, /n/ and /ŋ/):

English

(a)

[mi:t]

meat

(b)

[ni:t]

neat

(c)

[moʊɫ]

mole

(d)

[noʊɫ]

knoll

(e)

[sɪn]
sin
(f)
[sɪŋ]
sing
(g)
[dɪm]
dim
(h)
[dɪn]
din
(i)
[wɪn]
win
(j)
[wɪŋ]
wing

Now examine the same three nasal stops ([m], [n], [ŋ]) in the following data from Spanish. (We ignore the fact that Spanish has a dental, rather than an alveolar, nasal stop. These data are representative of Spanish, the prestige accent spoken in Spain: many varieties of Spanish lack the /θ/ and /ð/ phonemes. This fact does not affect the point made in the exercise.) Assume that the data are fully representative:

Spanish
(a)
[muðo]
'mute'
(b)
[nuðo]

‘knot’
(c)
[metɐ]
‘goal’
(d)
[netɐ]
‘pure’ (feminine)
(e)
[ɔmbɾe]
‘man’
(f)
[sɛmblɐɾ]
‘to seem’
(g)
[ɐndɐɾ]
‘to go’
(h)
[ɐnte]
‘in the face of’
(i)
[ɐŋgulo]
‘angle’
(j)
[lɛŋgwɐ]
‘language’

. [ŋ] does not appear word-initially, word-finally or between vowels.

(i)
What evidence is there for postulating distinct /m/ and /n/ phonemes in Spanish?
(ii)

Are there any grounds for postulating an /ŋ/ phoneme in Spanish? What evidence is there for your answer?
Now look at the following Spanish data:
(k)
[feliθ]
'happy'
(l)
[lɱfeliθ]
'unhappy'
(m)
[posible]
'possible'
(n)
[imposible]
'impossible'
(o)
[dispɛnsable]
'dispensible'
(p)
[indispɛnsable]
'indispensible'
(q)
[konθebible]
'conceivable'
(r)
[iŋkonθebible]
'inconceivable'
(s)
[ɐkɔstɔmbɾɐðo]
'accustomed'
(t)
[inɐkɔstɔmbɾɐðo]

‘unaccustomed’

(iii)

Is there any phonemic overlapping of the Spanish nasal stop phonemes? Explain.

3 Further phonetic transcription practice

Listen to Track 6.2. Transcribe, with as much phonetic detail as possible, the words you hear on the recording. The words are:

7

English Syllable Structure

7.1 Introduction

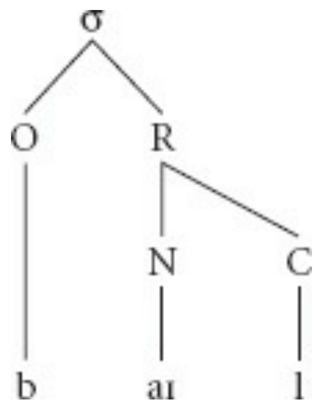
We have said morphemes are a kind of mental representation which have three properties: a syntactic category, a meaning and a phonological form. We have allowed, thus far, that the phonological form of a morpheme is present in the speaker's mentally constituted grammar, and that this phonological form consists in either a single phonological segment or a sequence of such segments. But this is only part of the story: there is more to the phonological form of a morpheme than that. There is evidence that those segments are organized into **phonological constituents**, rather in the way that words are organized into constituents (such as phrases and sentences). One of those constituents is **the syllable**. The evidence for the existence of the syllable comes largely in the form of phonological generalizations which cannot be adequately expressed without reference to the notion 'syllable'. The aim of this chapter is to examine the structure of the syllable in English, and exemplify some of the sorts of phonological generalization which are best expressed in terms of that structure.

7.2 Constituency in Syllable Structure

The two main constituents within a syllable are the **onset** and the **rhyme**. In the word , for instance, the first segment, /b/, constitutes the onset of the syllable and the last two segments, /aɪ/ and /l/, taken together, constitute the rhyme. The onset is defined as any and all consonants occurring before the vowel. What evidence is there for this division between onset and rhyme? The device of alliteration depends on identity of onsets, independently of the content of the rhyme, as in and , and , and so on. This constitutes evidence for the onset/rhyme division, and thus evidence that the rhyme is a well-founded syllabic constituent. Since that is so, then the onset as a constituent is equally well-founded (since the two are defined in contradistinction to each other). Slips of the tongue also show that the onset is a real unit in speech production. One type of slip of the tongue is the spoonerism, named after an academic called Spooner, who is said to have uttered sentences such as ‘You have missed my history lecture’ as ‘You have hissed my mystery lecture’, with an inversion of the onsets of ‘missed’ and ‘history’.

The rhyme may be further subdivided into the constituents **nucleus** and **coda**. Thus, in the word , the diphthong /aɪ/ constitutes the nucleus, and the consonant /l/ constitutes the coda. We may represent the constituency of the single-syllable morpheme as follows, where Greek ‘σ’ (sigma) stands for ‘syllable’, ‘O’ stands for ‘onset’, ‘R’ stands for ‘rhyme’, ‘N’ stands for ‘nucleus’, and ‘C’ stands for ‘coda’:

(1) :



A syllable such as this, which contains one or more consonants in coda position, is called a **closed syllable**, whereas a syllable which does not contain any consonants in coda position is referred to as an **open syllable**; as in the word :

(2) :

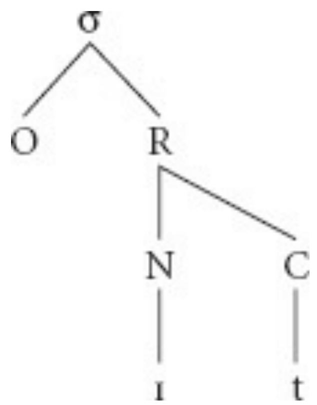
While a syllable have a nucleus, it is possible to have a well-formed syllable which does not contain any element other than a nucleus. The segment occupying the nucleus of the syllable is normally a vowel. An example of a word in English consisting of only one syllable, which in turn contains only a nucleus, is : /aɪ/. But the nucleus of a syllable in English may be preceded or followed by other segments, as we have seen, and those segments are typically consonants. In the word

, for instance, the nucleus is followed by a consonant in coda position: /aɪl/. In the word , the nucleus is preceded by a consonant in onset position: /baɪ/, and in the word , the nucleus is both preceded and followed by consonants: /baɪl/.

Morphemes like , which contain only one syllable, are said to be **monosyllabic**. In some languages, morphemes are monosyllabic. But in English, morphemes may contain more than one syllable: they may be **polysyllabic**. Examples are (which are bisyllabic), (which are trisyllabic), (which have four syllables) and so on.

In some languages, all syllables contain an onset consonant but, as we have seen, in English (and this is true of many other languages), this is not the case. For reasons to be explained later (connected with the notion of 'resyllabification'), we will represent such syllables with an **empty onset**, as follows:

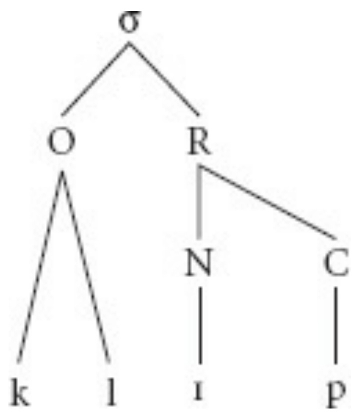
(3) :



In many languages, such as Hawaiian, onsets may contain a single consonant only, but in many others, English included, onsets may contain two segments (as in , etc.);

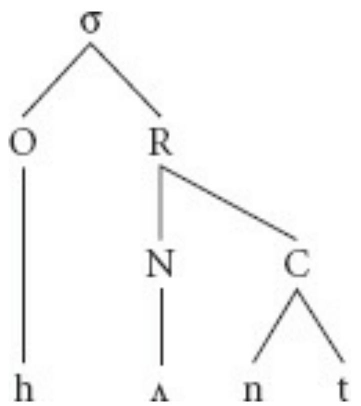
we will refer to these as **branching onsets**, and represent them as follows:

(4) :



Just as onsets may be branching, so codas may branch, as in the word :

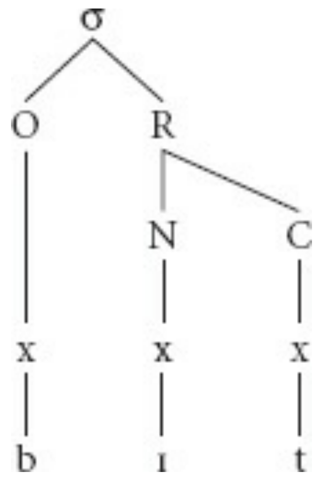
(5) :



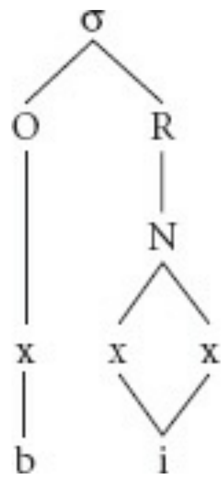
The distinction between, on the one hand, short vowels and, on the other, long vowels and diphthongs can be represented by taking the latter class of vowels to occupy a **branching nucleus**, with the former class occupying a non-branching nucleus. To represent the fact that long vowels and diphthongs are longer than short vowels, we say that segments are attached to a series of , referred to as the **skeletal tier**. The idea is that one can represent the difference between short vowels on the one hand, and long vowels (including

English diphthongs) on the other, by taking the former to be connected to a single skeletal slot and the latter to be connected to two skeletal slots. Thus, has a short vowel as its nucleus and is therefore represented with a non-branching nucleus, whereas and have branching nuclei:

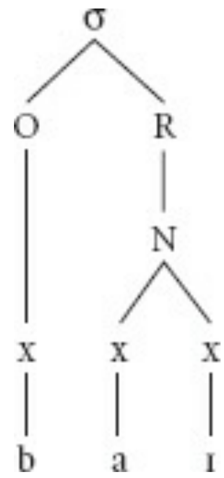
(6) :



(7) :



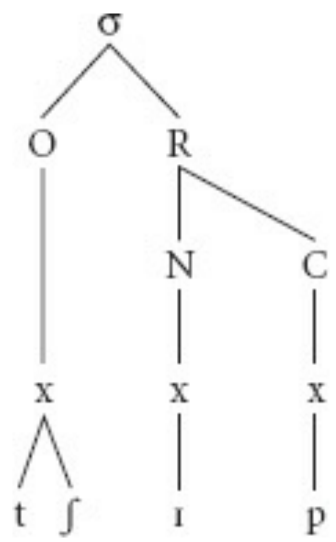
(8) :



What is intended by the representations in (7) and (8) is that long vowels are constituted as a single vowel quality which is attached to two skeletal slots, whereas long diphthongs, as in , have two different vowel qualities. The point is that nuclei with long vowels and with diphthongs are parallel with respect to the number of timing slots within the nucleus. We will henceforth adopt the skeletal tier in our representations of syllabic structure.

The skeletal tier enables us to say that affricates, which, as we have seen, have a closure element and a fricative release element, as in [tʃ] and [dʒ], are **complex segments**, since they behave like single segments (they occupy a single unit of timing) while having an internal structure which resembles two segments:

(9) :



7.3 The Sonority Hierarchy, Maximal Onset and Syllable Weight

You will agree, if you are a native speaker of English, that /bliŋk/ is a well-formed syllable (and happens to constitute the phonological form of an English word), as are the following: /blʊŋk/, /blɛŋk/, /blɐŋk/ and /blæŋk/ (most of which happen not to constitute the phonological form of English words). That is, your native speaker knowledge of English allows you to judge that these are syllabically well-formed, even though there are no words in English which have those phonological forms. That unconscious knowledge also allows you to judge that the following are ill-formed: /ibɪŋk/, /ibɪkŋ/, /tlɪnk/ and /blaɪmp/. The question is: what form does this unconscious knowledge take? What is it that we know, unconsciously, which allows us to make these judgements? Let us now seek to answer that question.

It is widely believed that there are both **universal** and **language-specific** constraints on the form that syllables may take, that is, constraints on the **syllabification** of sequences of segments. Among the universal constraints, we may mention two. Firstly, it is claimed that sequences of segments are syllabified in accordance with a **sonority scale**, which takes the following form:

(10) Sonority scale

Low vowels
High vowels
Approximants
Nasals
Voiced fricatives
Voiceless fricatives
Voiced stops
Voiceless stops

The idea is that, as one proceeds from the bottom to the top of the scale, the class of segments becomes more sonorous, or more vowel-like. Sonority is an acoustic effect: the more sonorous a sound, the more it resonates. Vowels have greater resonance than consonants, and voiced consonants have greater resonance than voiceless ones. If you listen to a singer holding a note for any length of time, the sound in question will most probably be a vowel. There are two articulatory reasons why it is easier to hold a vowel sound for longer than a consonant sound, and both are relevant to the production of sonority. The first is , as discussed in chapter 1, whereby stops are said to involve a greater degree of stricture than fricatives, which in turn involve greater constriction than approximants and vowels. Similarly, the more open a vowel articulation, the less stricture there is in the oral cavity. The acoustic effect of these sorts of articulation is that the lesser the degree of constriction, the greater the degree of sonority. The second articulatory factor is : voiceless segments are less vowel-like, less sonorant, than voiced segments: vowels are typically voiced, and voicing creates greater sonority.

Applied to syllable structure, the idea is that the most sonorous element in a syllable will be located within the nucleus, and that the further one gets from the nucleus, the less sonorous are the segments. Thus, in , the /b/ is less sonorous than the /l/, which is, in turn, less sonorous than the vowel: as one approaches the nucleus, so sonority increases. As one leaves the nucleus, we may note that the /ŋ/ is less sonorous than the vowel, and the /k/ less sonorous in turn than the preceding /ŋ/.

The 'degree of sonority' idea is very convincing, even if it runs into some difficulties. For instance, [s] + consonant onset clusters in English undermine the predictions made by the sonority hierarchy, since, in cases such as , the sonority scale principle makes the right predictions except with respect to the initial [s]. However, this merely serves further to underline the peculiarity of English sC (s + consonant) onset clusters: only [s]-initial onsets violate the sonority hierarchy, and the only three-way branching onsets in English are those which begin with an [s].

Another universal principle of syllabification concerns the syllabification of polysyllabic words, and is referred to as the principle of **Maximal Onset**. We have considered only monosyllabic words thus far; let us therefore consider the syllabification of the English word , whose segmental form is, let us say, /əpɹeɪz/. It is clear that the word is bisyllabic; the question is where the boundary between the syllables lies. We know that /p/ may occur in coda position in English, as in , etc. We also know that /pɹ/ is a well-formed onset, as in , etc., and we know that /ɹ/ may occur alone in onset position, as in , etc. Furthermore, we know that /pɹ/ is not a well-formed coda cluster: it violates the predictions of the sonority hierarchy. Thus, /u:pɹ/, /sɪpɹ/ etc. are ill-formed. We must therefore decide whether the syllabification of is /ə.pɹeɪz/ or /əp.ɹeɪz/ (where the full stop indicates the syllable boundary). The principle of Maximal Onset says that, in cases like this, where the language-specific phonotactics will allow for two or more syllabifications across a syllable boundary, it is the syllabification which maximizes the material in the following onset which is preferred. In this case, that is the former syllabification.

The principle of Maximal Onset is intimately connected with a universal fact about syllable structure: that syllables with an onset consonant are in some sense more basic than those without, and that presence of onset consonants is in some sense more basic than presence of coda consonants. It appears that the most 'basic' syllable structure in human languages is CV syllable structure, with a single onset consonant followed by a vowel. There are several types of evidence for this claim.

Firstly, CV-type syllables appear to be the syllable types that human children first utter when they begin to speak (e.g. [ba], [ma]) regardless of what language their parents speak. At that stage in the development of the child's syllable structure, syllables in the adult language with branching onsets will be uttered as CV structures. So too will syllables with coda consonants: the coda consonants will simply be absent at that stage. This strongly suggests that onset consonants are in some way more basic, in articulatory, and perhaps perceptual, terms than coda consonants.

Secondly, in many cases of aphasia, where post-stroke patients have suffered damage to their speech, CV syllable structures also appear to be the sort that first begins to appear as the patient recovers his or her speech, even if his or her native language has branching onsets and coda consonants.

Thirdly, languages which have both onset and coda consonants typically allow for a wider range of consonants to occur in onset position than in coda position.

Fourthly, coda consonants are much more likely to undergo loss of articulation in the course of the historical development of languages than onset consonants. This is what has happened with /l/ in coda position in some varieties of English, where the realization of /l/ has become vocalized ([w], which is vowel-like, rather than consonantal) in coda position, but not in onset position, so that [l] occurs in words like and , where the /l/ is in an onset, but [w] occurs in words like and , where the /l/ occurs in coda position (except in cases where words such as are followed by a word or suffix beginning with an empty onset, in which case the /l/ occupies that position and is realized as [l]; see 7.7 below on resyllabification). This kind of weakening of articulation can lead to complete **elision** (non-pronunciation) of a consonant. This is what has happened with [ɹ] in coda position in many accents of English. In those accents, words like and have lost the coda [ɹ], while retaining it in onsets, as in words like and .

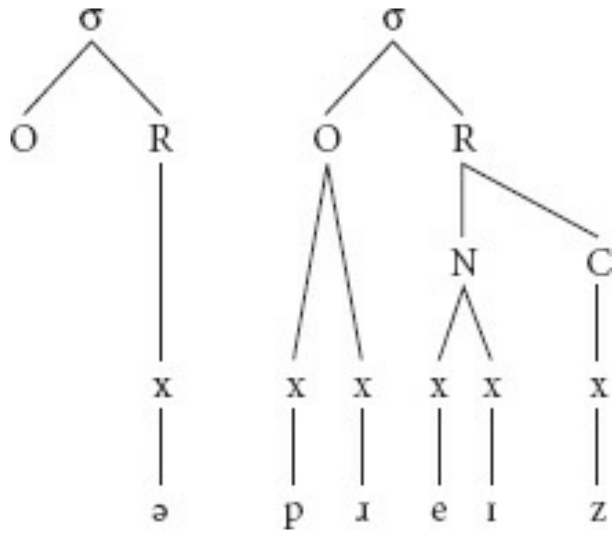
Such cases of articulatory weakening, often leading to complete loss of articulation, of coda consonants abound in the world’s languages. They suggest that coda consonants are somehow less salient in perception than onset consonants, and studies in the way that human beings retrieve phonological forms from mental storage suggest greater prominence for onset consonants than for coda consonants: if one is searching for a word in one’s lexical memory, one is more likely to search on the basis of onset consonants than of coda consonants.

Fifthly, there are no known languages which have VC-type syllables but lack CV-type syllables, whereas the reverse is not the case. This strongly suggests that CV syllables are more basic than VC, or indeed any other, syllable type.

This generalization about CV syllable structure probably has a basis in both articulation and perception. If you try to produce a word with an empty onset in isolation (e.g. the word), you will find it hard to do without uttering some kind of consonantal articulation (typically, a glottal stop) before you utter the vowel. Preference for filled, rather than empty, onsets is probably rooted in the nature of our articulatory apparatus and also tied to greater perceptual salience of onset consonants.

Given the principle of Maximal Onset, it is clear that, in a syllable such as the first syllable in , the rhyme contains a short vowel (dominated, of course, by a single skeletal slot) and does not contain a coda, thus:

(11) :



Syllables such as the first syllable in , in which there is no branching within the rhyme, either at the level of the rhyme node itself, or within the nucleus, are called **light syllables**. And syllables which have branching anywhere within the rhyme constituent are called **heavy syllables**. This distinction in **syllable weight** is said by some to be important in understanding the nature of word stress in English.

There are two generalizations about word stress in English which some phonologists make. The first is that any stressed syllable in English is very likely to be a heavy syllable. The second is that monosyllabic words may not end in one of the short vowel phonemes (/ʊ/, /ɪ/, /æ/, /ɛ/, /ɑ/ or /ʌ/), since a nucleus containing only one of those vowels, with no coda consonant, is light, and if a monosyllabic word is to be stressed, there is no choice as to which syllable it will be stressed on.

7.4 Language-Specific Phonotactics

Let us now consider some language-specific constraints on the sequences of segments which may be combined in syllable structure, known as language-specific **phonotactic constraints** (phonotactics, for short).

We have allowed that, in English syllables, onsets, nuclei, rhymes and codas may branch. But we have not said whether there is any on the number of branches they may have. Only one sort of English onset exceeds binary (two-way) branching: /s/ + consonant + {/j/, /w/ or /ɹ/} onsets, as in and . Note that the range of segments which may form the third element in such sequences is even more restricted than those in binary branching onsets.

As we have seen, onsets may branch in English, but if they do, there are phonotactic constraints on the form they may take. Ignoring the /s/ + consonant cases, we may say that the first segment must be a stop or a fricative and the second must be /ɹ/, /l/, /j/ or /w/. Thus /pɹ/, /pl/, /pj/, /bɹ/, /bl/, /bj/, /tɹ/, /tw/, /dɹ/, /dw/, /kɹ/, /kl/, /kw/, /θɹ/, /θw/, /fɹ/, /fl/, /fj/, /sl/, /sj/ and /sw/ are all permissible. This list reflects other onset phonotactics. For instance, /t/, /d/ and /θ/ may not be followed by /l/, and none of the voiced fricatives may occur in branching onsets.

Among the phonotactic constraints on rhymes in English, we may note the following. Firstly, /h/ does not occur in rhymes in English. Secondly, in many accents of English, /ɹ/ does not occur in rhymes either; so that words like and arguably have phonological forms such as /fɑːm/ and /kɑː/, without an /ɹ/. Accents which lack /ɹ/ in rhymes are referred to as **non-rhotic** accents; they include Australian English, New Zealand English, RP, South African English, most of the accents of the North of England, and the Southern and Eastern accents of the United States. These accents were rhotic at one stage; [ɹ] has been lost in rhymes in those accents. Rhotic accents, which have not undergone this historical change, include GA, the accents of English spoken in Scotland, and some accents spoken in the South West of England. We will discuss such accents in more detail later.

The overall shape of syllables in a language often acts as a major factor in adult second language acquisition. For instance, simplifying somewhat, Japanese syllable structure does not allow for branching onsets. This often has the effect that, when native speakers of Japanese utter English words with complex onsets, such as , they will tend to insert a vowel after each of the first three consonants, thus rendering the word trisyllabic: [sɯkɯɾɯ]. This process of vowel insertion is known as vowel **epenthesis**, and such vowels are known as epenthetic vowels. Similarly, and again simplifying somewhat, Japanese syllable structure does not allow for word-final coda consonants, so that English loanwords, such as , which end in a coda consonant in English, tend to be uttered as bisyllabic words ending in a vowel: [keki].

Similar cases abound. For instance, Spanish, unlike English, does not have words beginning with an s + consonant onset. However, Spanish does have words such as which, in some cases, correspond to English words with an s + consonant onset (in this case, the word). One of the effects of this is that Spanish speakers tend to insert an [ɛ] before English words beginning with s + consonant clusters, as in [ɛspeɪn] ( ). Similarly, English loanwords in Spanish are pronounced with such an epenthetic [ɛ], as in [ɛsmokin] (‘smoking jacket’).

7.5 Syllabic Consonants and Phonotactics

It is possible for consonants to form the nucleus of a syllable in the speech of English speakers, particularly as the rate of speech increases. These consonants are called **syllabic consonants**. Three alternative pronunciations of the word , for many speakers, are [bɑtəɫ], [bɑtɫ̩] or [bɑʔɫ̩]. In the latter two pronunciations, the final unstressed vowel (schwa) has been lost, but the word still has two syllables, with the lateral becoming syllabified. Syllabic consonants are transcribed by means of the ‘syllabic’ diacritic, placed under the appropriate consonant symbol.

Syllabic nasals are common in many varieties of English. An example is the word , which has two syllables. For many speakers of English, it may be pronounced [bʌtən] or [bʌʔn̩], where the second pronunciation has a syllabic nasal. The second vowel in the first pronunciation is an unstressed vowel (schwa) which may be ‘lost’, particularly in faster or more casual speech. A similar example is the word , which, for many speakers, has (at least) the two pronunciations [hæpən] and [hæʔm̩]. Here, the nasal [n] assimilates to the ‘intended’ bilabial articulation [p], which in turn is articulated as a glottal stop. Other examples involving nasals are [ɹɛʔŋ̩] vs [ɹɛkən] ( ), and [kn̩u:] vs [kənu:] ( ).

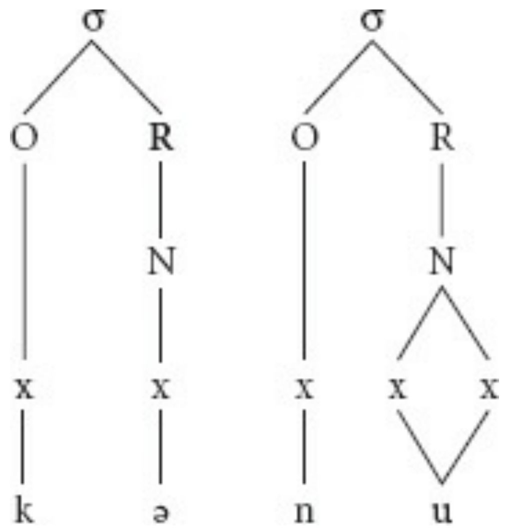
A similar example involving the approximant [ɹ] is the word

, which often has the alternative pronunciations [pəɹeɪd] and [pɹ̩eɪd]. It is typically nasals, laterals and [ɹ] which undergo syllabification in English, although fricatives may also be syllabified, as in some pronunciations of, for instance, , which may be [səp^{h}ɔ:t] or [s̩p^{h}ɔ:t] (as distinct from : [spɔ:t], with one, not two, syllables).

In English (but not in some other languages), for every case in which a syllabic consonant may occur, there will be an alternative pronunciation of the word with a vowel preceding or following the syllabified consonant.

All of these words have phonological representations containing a vowel in the nucleus of each syllable, as in :

(12) :



It is striking that, although English speakers frequently utter words such as with a syllabic nasal, as in [kn̩u:], when faced with non-English words such as , they will tend to insert an epenthetic vowel, uttering the word as [gənu], making it conform to English syllable structure, in which /gn̩/ is not a permissible English syllable. The important point to be made here is that constraints on English syllable structure are defined in terms of permissible sequences, rather than ones.

We have said that English does not allow for phonological syllabic consonants, but many other languages do. The Polynesian language Maori, and many Bantu languages, for instance, allow for phonologically syllabic nasals. Examples are Bantu names such as and , each of which begins, phonologically, with a syllabic nasal, as in [ŋkomo], which has three syllables. Native speakers of English will tend to utter such words with an epenthetic vowel placed adjacent to the relevant nasal, as in [nɪkomo] or [ɪŋkomo], thus making the phonetic sequence conform to the English phonological pattern. All of these cases provide evidence for the phonological vs phonetic distinction we have drawn, and show how profound an influence our phonological representations can have on our perception and production of non-native words.

7.6 Syllable-Based Generalizations

We said in 7.1 that some of the evidence for the existence of the syllable as a phonological constituent comes from the fact that there are significant phonological generalizations which cannot be adequately expressed without appeal to syllable structure. One such generalization concerns the distribution of velarized laterals (‘dark l’s) in many accents of English. For many speakers, the following sort of distribution between velarized and non-velarized /l/ may be found:

(13) Velarized and non-velarized /l/

[lʌɫ]	lull	[li:f]	leaf	[sli:p]	sleep
[bɒtɫ̩]	bottle	[pi:əɫ]	peel	[mɪɫk]	milk
[lɪli]	lilly	[lɪɫtɪŋ]	lilting	[fɑ:ɫtə]	falter

One might attempt to state the distribution of [l] and [ɫ] as follows: [ɫ] occurs when immediately followed by another consonant, or at the end of a word (i.e. when immediately followed by a word boundary). But one of the objections to this formulation is that it is not clear what, if anything, a following consonant and a word boundary might have in common. A simpler statement of the distribution, which does not entail appeal to this peculiar disjunction of environments, is to say that [ɫ] occurs in the   of a syllable, and [l] in the   . Indeed, we might take that syllable-based account of the distribution to help us diagnose the syllabic status of the second /l/ in : we might say that, because the second /l/ is not velarized in the speech of many speakers, this confirms our claim that it occupies onset position in the second syllable of , rather than coda position in the first syllable.

Let us consider another example, from London English, of the syllable-based nature of some phonological generalizations. The vowels [ɒʊ] and [ʌʊ] are said by some phonologists to be in complementary distribution in the speech of many speakers of London English. The following table exemplifies this:

(14) [ɒʊ] and [ʌʊ] in London English

[ɹɒʊl]	roll	[kʌʊlʌ]	cola
[stɹɒʊl]	stroll	[lʌʊd]	load
[ɒʊld]	old	[tɒmbʌʊlʌ]	tombola

The claim is that the [ʌʊ] phoneme is realized as [ɒʊ] when it is followed by an /l/ which is in the same syllable as the [ʌʊ] (i.e. when followed by a **tautosyllabic** /l/), and as [ʌʊ] elsewhere. Thus, in the words in the left-hand column, we have [ɒʊ]. In the right-hand column, the vowel in clearly lacks a following tautosyllabic /l/; as for and , we would want to argue, from the principle of Maximal Onset, that the /l/s there occupy onset position in the following syllable, and thus that the [ʌʊ] there also lacks a following tautosyllabic /l/.

7.7 Morphological Structure, Syllable Structure and Resyllabification

The case just cited is a little more complex than we have, thus far, suggested. Consider the following further data, from the same accent:

(15)

[ɹɒʊlʌ]	roller	[ɹʌʊlənd]	Roland
[hɒʊli]	holey	[hʌʊli]	holy

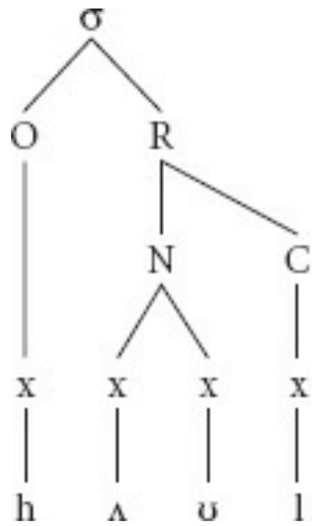
What is the phonological status of the diphthong in ? On the one hand, we have said that the [ɒʊ] allophone appears before a tautosyllabic /l/. On the other hand, it would appear that Maximal Onset would have us syllabify the /l/ into the onset position of the second syllable. But if the /l/ is indeed in that position, then we ought to get the [ʌʊ] allophone. Why, then, do we get that allophone? Let us consider two possible responses.

One response is to say that speakers of this accent originally had a straightforward phonological rule, of the sort we have given, for the realization of the [ʌʊ] phoneme, but that, as the accent has evolved, a **phonemic split** has emerged: the two vowels were in complementary distribution, but have come to occur in overlapping, parallel distribution. Evidence which is cited in support of this view is the emergence of minimal pairs, such as vs . Here, it is argued, we have clear evidence that a phonemic split has occurred.

We might make the following objection to the analysis just cited: it is failing to take note of an important generalization concerning the members of minimal pairs such as and , namely that those containing the [ʌʊ] vowel consist of only one morpheme (they are morphologically), whereas those containing [ɒʊ] consist of more than one morpheme (they are morphologically). Furthermore, in each case, the relevant vowel occurs before an /l/ which, in the morphologically complex cases, is morpheme-final.

One way of expressing this sensitivity of the phonological rule to morphological structure is to say that the rule applies to the affixation of the suffix in cases like , as follows:

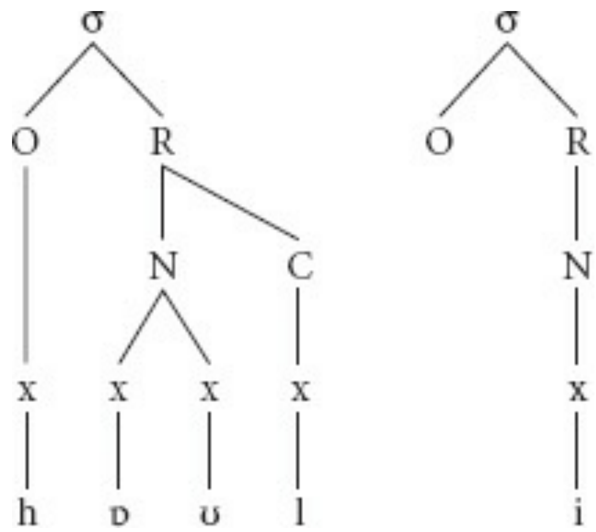
(16) root: /hʌʊl/



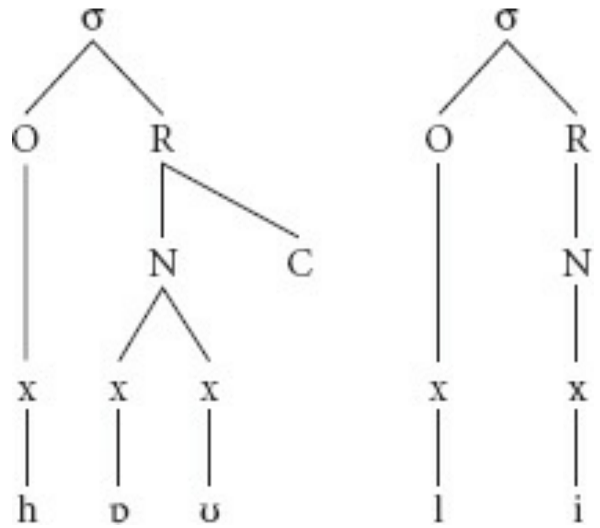
(17) Application of [ʌʊ] rule

[ʌʊ] → [ɒʊ] before tautosyllabic /l/

(18) Affixation



(19) Resyllabification



This analysis requires appeal to a further notion: that of **resyllabification**. The idea is that, while the /l/ is syllabified into coda position, it is syllabified, after affixation, and according to the Maximal Onset principle, into the empty onset in the suffix (the idea of resyllabification into an empty onset position being part of the motivation for postulating empty onsets).

There is an important point to be made with respect to cases such as this. It is that there may be cases where a phonetic distinction acts as the basis for minimal pairs (in this case, pairs like) where, nonetheless, we do not wish to postulate a phonemic distinction between the two segments. In such cases, the members of those pairs will almost always differ in their morphological structure (in this case, is morphologically simple, while is morphologically complex), and that difference will affect the application of generalizations which govern the allophones of phonemes.

Notice that, in the case of many non-London accents, the reverse ordering of phonological generalizations and resyllabification applies in the case of the generalizations which yield the [i:] and [i:ə] allophones of the /i:/ phoneme and the clear and dark allophones of the /l/ phoneme. We saw that, in many accents, in words like and , the allophone of /iː/ is [iːə], and the /l/ allophone is [ɫ], as in [fiːəɫ] ( ). In cases like , the /l/ is resyllabified into the empty onset of the suffix -ɪŋ, thus: [fiːlɪŋ]. We want to say that the rule (as defined on p. 64) which governs the allophones of the /l/ phoneme, and the rule which governs the allophones of the /iː/ phoneme, apply resyllabification.

We will take the view that, while the phonemic principle, in distinguishing between contrastive and non-contrastive distinctions, embodies an important insight into the nature of phonological organization, its conception of phonological contrast is at times overly restrictive, since it is defined without reference to the influence of morphological and syntactic factors on phonological organization. We will therefore, at times, allow our analyses to, as it were, override the phonemic principle. In doing so, we are not abandoning the notions which find a place in that principle; rather, we are allowing that morphological factors may influence phonological processes.

7.8 Summing Up

We have adopted here an account of the syllable as a phonological constituent. We have said that the sub-parts of syllables have differing degrees of perceptual salience, so that the nucleus is more salient than the other parts of the syllable. This is perhaps why, in many languages, coda consonants so often diminish in degree of stricture, to the point of 'fading away' altogether. This is what has happened to the realizations of the /ɹ/ phoneme in many accents of English, for instance: at one stage in their history, /ɹ/ was realized in coda position. It is also what happened to many case endings (suffixes which consisted of a vowel and one or more coda consonants) in the history of English nouns (English nouns used to have case endings such as , all of which have disappeared over time).

The loss of case endings in English nouns involved another factor, however: not only are nuclei more salient perceptually than coda consonants, but, in languages like English, some nuclei in a word (nuclei) are more salient than others. It is to the subject of word stress in English that we now turn.

Exercises

1 For each of the following words, say how it is syllabified, and why alternative syllabifications are disallowed (e.g. the word /kwɒntɪti/ ( ) is syllabified as /kwɒn.tɪ.ti/; the syllabification /kwɒ.ntɪ.ti/ violates English phonotactics, since /nt/ is not a permissible branching onset; the syllabification /kwɒnt.ɪ.ti/ is ruled out by the principle of Maximal Onset):

(a)
/ɹəvaɪz/
( )
(b)
/pɹədɪkʃən/
( )
(c)

/ɹɛzɪdɛnʃəl/

( )

(d)

/ɛmpəɹə/

( )

(e)

/dʒæpəni:z/

( )

(f)

/kɒndʌkt/

( )

2 Of the following monosyllabic phonological representations, say which are English, and which are non-English, representations for monosyllabic words. For each non-English form, say why it is not possible:

(a) /pnɪt/

(b) /psɪt/

(c) /pɹɒt/

(d) /plɛt/

(e) /pɪh/

(f) /ɹpæt/

3 Examine the following data from the Cockney variety of London English:

(a)

[lɔɪʔ]

light

(b)

[laɪdi]

lady

(c)
[fiːlɪn]
feeling
(d)
[fiɪw]
feel
(e)
[gɜːw]
girl
(f)
[bɒʔu]
bottle (bisyllabic)
(g)
[filʌ]
filler
(h)
[fiɪwʔʌ]
filter
(i)
[waɪ]
way
(j)
[weɪʔ]
wet

The sound [w] may be a realization of the /w/ phoneme, as in (i) and (j). However, it may also be an allophone of the /l/ phoneme, as in (d), (e) and (h). In (f), /l/ becomes a syllabic '[w]', namely [u]. That is, there is phonemic overlapping between /l/ and /w/. In which contexts does the [w] allophone of /l/ occur? Your answer should be expressed in terms of syllable structure.

Listen to sound files online

4 Further transcription practice

Listen to Track 7.1 at www.wiley.com/go/carrphonetics. Transcribe, with as much detail as possible, the words you hear on the recording, indicating syllable boundaries (with a full stop) and any syllabic consonants (e.g. : [bæ.ʔɫ̩], for some speakers):

8

Rhythm and Word Stress in English

8.1 The Rhythm of English

Human beings speak : they engage in the act of speaking by putting regular in the speech signal. You can hear those beats in an English utterance such as . Here, the beats are on , and . In most varieties of English, we do not necessarily place a beat on every single syllable. In this utterance, no beat falls on the preposition , or on the two occurrences of . This is because English, unlike certain other languages, is : the rhythmic beats fall only on syllables. In our example, only , and are stressed, so the beats fall only on those. English is unlike many other languages in this respect. Take the phrase , the name for a product sold by a well-known fast-food company. This is pronounced [ˌtʃɪkənməkˈnʌgəts]. There are two stressed syllables in this sequence. (The second is more prominent than the first: we'll come back to that.) The sequences [ˈtʃɪkənmək] and [ˈnʌgəts] form rhythmic units in the utterance. Those units are called . A metrical foot in English consists of a stressed syllable followed by zero or more unstressed syllables. In our example, the first metrical foot contains a stressed syllable and two unstressed syllables: [ˌtʃɪkənmək]. The second metrical foot contains a stressed syllable and one unstressed syllable: [ˈnʌgəts]. Notice that divisions between the metrical feet need not coincide with word boundaries: the word boundary falls between the words and the word . But the rhythmic boundary falls between [ˌtʃɪkənmək] and [ˈnʌgəts]. We call these metrical feet . This is an adjective derived from the noun . A trochee is essentially a stressed–unstressed sequence, such as [ˈnʌgəts]. We will examine these trochaic metrical feet in more detail in chapter 9. For the moment, we need only note that word stress patterns are part and parcel of the rhythmic structure of English.

8.2 English Word Stress: Is It Entirely Random?

We have already noted that the native speaker’s perceptual capacities allow him or her to say how many syllables a word has, in the absence of any conscious knowledge of what a syllable might be, or how it might be defined. Similarly, English speakers can tell which syllable in a word receives most stress, in the absence of any conscious knowledge of exactly what ‘stress’ might be. While the native speaker may not know consciously what stress is, it seems clear that, the more stressed a syllable is, the more salient it is, perceptually. For instance, most native speakers of English will agree that, in the word , it is the antepenultimate (third last) syllable (the one before the penultimate, or second last, one) which is most stressed. Equally, most speakers will know that, in , it is the last of its three syllables which receives most stress, and so on. It is equally striking that the native speaker can judge that, while the syllable in receives more stress than either of the others, the antepenultimate (third from last) syllable in turn receives more stress than the penultimate (second from last) syllable. The final syllable and the penultimate syllable of are **unstressed**, as is the syllable before the ante-penultimate syllable. They are therefore less salient than the antepenultimate syllable, which has primary stress. The penultimate syllable of is unstressed and is the least salient syllable in that word. Let us say that the syllable in a word which receives most stress has **primary stress**, and that syllables such as the ante-penultimate syllable in have **secondary stress**; while syllables which have neither primary nor secondary stress are unstressed syllables. We could therefore say that any given word will have a : in the case of , a final syllable with primary stress, preceded by a penultimate unstressed syllable, preceded by an antepenultimate syllable with secondary stress. We can informally represent primary stress by placing a superscript diacritic (ˈ) immediately before the start of the appropriate syllable, and secondary stress by using the subscript diacritic (ˌ), leaving any unstressed syllables without a diacritic, as follows: [ˌkæŋgəˈɹuː]. This notation is used in pronouncing dictionaries.

It seems clear that knowledge of the stress patterns of words does not normally require instruction, and children acquiring their native language are not normally given explicit instruction by their parents as to where the stresses in a word are placed. Having noted that, let us consider the following question: how does the speaker know what the stress pattern of a given word is? It seems reasonable to suggest that the child who is acquiring English simply has to the stress pattern of each word as it is learned. After all, for any given word, the speaker has to memorize the sequence of phonemes which make up the phonemic form of the word. Whether one is a child acquiring English or an adult learning English, one might just as well memorize the stress pattern while one is at it. One might object that this means that an average speaker has a vast number of stress patterns to memorize (as many stress patterns as there are words in her or his vocabulary), but we know that human beings are very good at storing large amounts of information of this sort in memory. Again, we can point to the vast number of phonological forms which the speaker clearly must have in mental storage; it is, surely, not too tasking to an organism which has that kind of storage capacity to store the stress patterns of those words along with the sequence of phonemes which partly make up its phonological form.

But this is not to say that there are no unconsciously stored generalizations governing stress patterns in English. We know that, for some languages, such as Modern Greek, the stress pattern of most words is entirely arbitrary. We also know that some languages have fixed stress: the stress always falls on a given syllable (in French, for instance, it always falls on the final syllable of the word, and in Polish on the second last syllable of the word). Let us consider some evidence in favour of the idea that the native speaker of English has unconsciously formed generalizations concerning word stress patterns in English. Take the following two sets of bisyllabic English words (words with two syllables), all of them morphologically simple (i.e. containing only one morpheme: no prefixes, no suffixes):

(1)

[ˈsɪmpɫ̩]	simple	[ˌbæmˈbuː]	bamboo
[ˈkʰɑːpət]	carpet	[dʒɪˈɹɑːf]	giraffe
[ˈgæðə]	gather	[pəˈɹeɪd]	parade
[ˈsʌdən]	sudden	[ænˈtiːk]	antique
[ˈɪnsɛkt]	insect	[kɹiˈeɪt]	create

The words in the first column have primary stress on the penultimate syllable, those in the second column have primary stress on the final syllable. On the face of it, it looks as if bisyllabic words in English may have the primary stress on either the final or the penultimate syllable. Perhaps it's all entirely arbitrary: perhaps there are no rules. Consider, however, the following words as uttered by speakers of French and by speakers of English with a noticeably English pronunciation of French (you may ignore any unfamiliar phonetic symbols in the French transcriptions: it's the stress patterns that matter here):

(2)

		French speaker	English pronunciation
manger	'to eat'	[mɑ̃ˈʒe]	[ˈmɑ̃ʒeɪ]
chercher	'to look for'	[ʃɛʁˈʃe]	[ˈʃɛɹʃeɪ]
bateau	'boat'	[bɐˈto]	[ˈbætəʊ]
français	'French'	[fʁɑ̃ˈse]	[ˈfɹɑ̃seɪ]
lointain	'distant'	[lwɛ̃ˈtɛ̃]	[ˈlwʌnteɪn]

The word stress rule in French could not be simpler: stress the final syllable of the word. So why do so many speakers of English have this strong tendency to stress the penultimate syllable in so many French words when they try to speak French? It’s not as if the French rule is hard to grasp. And, after all, English speakers have countless words in their language in which bisyllabic words are stressed on the final syllable: it’s not as if they are not in the habit of placing primary stress on the final syllable of bisyllabic words in their own language. So what is ‘typically English’ about such stress patterns in French bisyllabic words?

Consider too the following trisyllabic French words and a typical English mis-stressing of them:

(3)

		French speaker	English pronunciation
bâtiment	‘building’	[bɐti'mɑ̃]	['bætimɑ̃]
fermeture	‘closure’	[fɛʁmə'tyʁ]	['fɛʁmətyʁ]
soigneusement	‘meticulously’	[swɐɲøz'mɑ̃]	['swɐɲøzmɑ̃]
sortilège	‘spell’	[sɔʁti'lɛʒ]	['sɔʁtilɛʒ]
consacrer	‘to devote’	[kɔ̃sɐ'kʁe]	['kɔ̃sɐkʁe]

Why should there be a tendency among English speakers to mis-stress such trisyllabic words on the antepenultimate syllable? What is it that makes such a mis-stressing ‘typically English’?

Finally, consider the following trisyllabic nouns in English:

(4)

['ɛləfənt]	elephant	[pə'tʰeɪtəʊ]	potato	[ˌkæŋgə'ru:]	kangaroo
['sɪnəmə]	cinema	[bə'nɑ:nə]	banana	[ˌræfjʊ'dʒi:]	refugee

The words in the first column have primary stress on the antepenultimate syllable, those in the second column on the penultimate syllable, and those in the third column on the final syllable. Given that the words do not differ as to syntactic category (they are all nouns) and do not differ in terms of total number of syllables, it looks as though there is no generalization concerning primary word stress in English trisyllabic nouns: it all looks entirely arbitrary, as if there were no rules. Consider, however, the following trisyllabic nouns:

(5)

Gigondas	moussaka	Zaventem	tavola

The first of these words is French (it is the name of a town, and a wine, in the Southern Rhône Valley), and is stressed, like all French words, on the final syllable. The second is Greek, and is stressed by Greek speakers on the final syllable. The third is Dutch (it is the name of a town in Belgium, and the name of Brussels airport); it is stressed, in Dutch, on the antepenultimate syllable. The fourth is the Italian word for ‘table’ and is stressed, in Italian, on the antepenultimate syllable.

What is striking about English speakers who know little or no French, Greek, Dutch or Italian, and have never heard the words before, is that they show a very strong tendency, on first encountering them, to mispronounce these words by stressing them on the syllable, as follows: [dʒɪˈgɒndəs], [mu ˈsɒ:kə], [zəˈvɛntəm], [təˈvəʊlə]. If the English speaker has no word stress generalizations, this tendency is deeply puzzling, since that would mean that, given a word one has never encountered before (especially a foreign word), one should display no tendency to prefer placing the stress on any particular syllable. One might expect a given individual to utter each word variably on different occasions, with each of the three possible stress patterns. And even if a given speaker alighted, arbitrarily, on a given stress pattern and stuck to it thereafter, one would expect variation from speaker to speaker. But this does not seem to happen: the pronunciation in which the second syllable is stressed is the one which they tend to opt for. You will probably agree, especially if you know any French, Dutch or Italian and have heard English speakers mispronouncing words in those languages, that this kind of pronunciation is 'typically English'. But, once again, is 'typically English' about it? We can only answer such questions if there word stress generalizations, or at least word stress tendencies, in English, and if we know what they are. We will now look at the form those generalizations take.

8.3 English Word Stress: Some General Principles

A first general principle (Principle 1: The End-Based Principle) is that the placement of primary stresses in English words is calculated by counting from the of the word: the primary stress in a word will tend to fall on either the final syllable of the word, the penultimate (second last) syllable or the antepenultimate (third last) syllable (though it fall earlier than that). This reflects the fact that most varieties of English have word stress patterns which are essentially . Recall that the adjective is derived from the noun , and that a trochee is a stressed syllable (whether primary stressed or secondary stressed) followed by zero or more unstressed syllables. We will say that stressed monosyllabic words (such as), words with penultimate stress (such as and) and words with antepenultimate stress (such as and) all exhibit trochaic stress patterns. The rhythm of most varieties of English is trochaic: there is a tendency to place the rhythmic beat on the stressed syllables of trochaic feet.

A second general principle (Principle 2: The Rhythmic Principle) is that, while it is possible for English words to with as many as four unstressed syllables (as in '), English words cannot with more than unstressed syllable. Principle 2 is directly related to Principle 1: the reason why we do not begin words with sequences of two or more unstressed syllables is that, if we place a secondary stress within such sequences, we can create a trochaic foot, which is desirable from the point of view of the rhythmic structure of English. When we derive from , the primary stress shifts from the final syllable of onto the final syllable of (we will examine such stress shifts shortly). But, having shifted the primary stress, we cannot leave the word with a sequence of two unstressed syllables preceding the primary stressed syllable: we must place a secondary stress on one of the two preceding syllables: ˌ ' .

When this happens, a third general principle (Principle 3: The Derivational Principle) comes into play: there is a tendency to place the secondary stress on the syllable which had primary stress in the deriving word (the word which we are deriving the more complex word from). For instance, the word exhibits a shift of primary stress from its deriving word ˈ (which itself is derived from ˈ). Since Principle 2 dictates that we cannot leave the word with a series of four unstressed syllables prior to the syllable with primary stress, we must place a secondary stress somewhere in that sequence of syllables. The Derivational Principle says that we ought to place it on the syllable which had primary stress in the deriving word: ˌ ˈ .

However, Principle 3 may be overruled by a fourth general principle (Principle 4: The Stress Clash Avoidance Principle), which states that one should try to avoid having two adjacent stressed syllables. In the case of ˌ ˈ , both principles are adhered to. But note that, in the case of ˌ ˈ , this is not the case: the primary stress in the deriving word ˈ falls on the final syllable. Principle 3 dictates that we place the secondary stress on that syllable, but Principle 4 dictates that we do not, since this would result in a stress clash, with two adjacent stressed syllables: ˌ ˈ . Cases like this demonstrate that, where the Derivational Principle and the Stress Clash Avoidance Principle come into conflict, and we have an option as to which syllable to place the secondary stress on, it is the Stress Clash Avoidance Principle 4 which predominates: in a word like ˌ ˈ , we place the secondary stress on the antepenultimate syllable in order to satisfy the Stress Clash Avoidance constraint. The Stress Clash Avoidance Principle is a strong general in English. But there words in English which violate that principle, such as the verb ˌ ˈ , and the nouns ˌ ˈ and ˌ

ˈ . We will look at these in more depth in the following chapter. As far as the Derivational Principle is concerned, there are words such as and , which are derived from ˈ and ˈ . In such cases, we have no choice but to place the secondary stress on the penultimate syllable: ˌ ˈ , ˌ ˈ . But where we have a choice between two or more syllables, as in ˌ ˈ and ˌ ˈ , we both avoid a stress clash and also create a sequence of two trochaic feet within the word. So Principle 4, the Stress Clash Avoidance Principle, often works together with Principle 2, the Rhythmic Principle, to create extra trochaic feet at the beginnings of words. Why, despite the avoidance of sequences of unstressed syllables at the of a word, is it nonetheless possible to have as many as four unstressed syllables at the of a word, as in ? Because they end with certain kinds of suffix. We will consider such suffixes shortly. For the moment, let us look further at words stress in morphologically simple words.

8.4 Word Stress Assignment in Morphologically Simple Words

English is a Germanic language which has borrowed a huge amount of vocabulary from Latinate languages, notably French and Latin, many of them with Latinate suffixes and prefixes. The effect of this has been to make the word stress patterns more complex than they would otherwise have been, and non-native speakers will testify to the difficulty they often experience in trying to master the stress patterns of English words. Nonetheless, there is considerable regularity in English word stress patterns. Let us begin by considering words which clearly do not have prefixes or suffixes, in present-day English (though some of these had prefixes historically). We will distinguish words of a lexical category from words of a non-lexical category. Words of a lexical category are nouns, verbs, adjectives and adverbs. Words of a non-lexical category include prepositions, determiners (such as), pronouns (such as) and the conjunction . Words of a non-lexical category, often referred to as function words, are not normally stressed. Among the words of a lexical category, primary stress placement may vary, depending on the syntactic category of the word.

Monosyllabic words of a lexical category (such as), are unproblematic: there is only one syllable for the primary stress to fall on. Let us therefore move on to morphologically simple bisyllabic words, and then proceed to morphologically simple polysyllabic words (words with three or more syllables).

8.4.1 Morphologically Simple Bisyllabic Words

The basic pattern here is the native Germanic pattern, that is, with primary stress on the syllable, as in

,

. However, there is a substantial class of exceptions to this basic pattern, in which bisyllabic nouns have final stress. Among these are nouns which have been borrowed from other languages and are written with a double vowel letter, as in

. Another set of bisyllabic nouns stressed on the final syllable have been borrowed from French; they contain the French endings - , - , - , - , - , - , - , - , as in ,

Since word stress in French always falls on the final syllable, words such as these have mostly retained the French stress pattern. Other French bisyllabic loanwords which have retained final stress include : [həʊˈtʰɛɫ]. Had this word been fully nativized, it would be pronounced [ˈhəʊtəɫ], with the native Germanic trochaic stress pattern. We assume here that words such as , etc. are mostly morphologically simple in contemporary English.[1]

The basic Germanic pattern is again trochaic, i.e. with stress on the penult, as in

. The trochaic pattern can also be found in morphologically simple bisyllabic adjectives ending in - , as in .[2] However, there are bisyllabic adjectives with final stress, such as . These contain historical prefixes which are Latinate in origin: they come from Latin or French. Most of these historical prefixes no longer count as productive prefixes in contemporary English (they cannot be freely combined with roots to form new words). Large numbers of these words are now morphologically simple, for the vast majority of speakers of English speakers. In addition to these, bisyllabic adjectives which contain what were, historically, French suffixes, such as and , have final stress. As with nouns like and , we assume here that adjectives such as and are morphologically simple.

Once again, the basic Germanic pattern is trochaic. Many bisyllabic adverbs end in - , as in and . We will deal with those in 8.5. Other adverbs which do not end in - include and . These also have a trochaic stress pattern.

The basic trochaic tendency in most varieties of English is much less evident in bisyllabic verbs: there are many with final primary stress, as we have seen:

,

. As with adjectives, many of these historically had prefixes in the languages they were borrowed from: French or Latin (e.g.

). The tendency to avoid stressing the historical prefix is strong: it is striking that such bisyllabic verbs differ so often from their bisyllabic noun counterparts, as in the verb–noun pairs ˈ (verb) but ˈ (noun), ˈ (verb) but ˈ (noun), ˈ (verb) but ˈ (noun), ˈ (verb) but ˈ (noun), ˈ (verb) but ˈ (noun), etc., where the verb takes final stress but the noun takes the normal trochaic penultimate stress. (But there are exceptions to this pattern, as in ˈ where the noun and verb both take the verb pattern, and ˈ , where the noun and verb both take the noun pattern.)

It is also worth noting that bisyllabic verbs ending in - will take primary stress on the - , as in

. These have to be distinguished from - words with three or more syllables (see below). Note that, in GA, some of these words take penultimate stress, such as and .

There are, however, bisyllabic verbs with the basic trochaic stress pattern, such as

.

8.4.2 Morphologically Simple Polysyllabic Words

The basic Germanic pattern is trochaic. For words of more than two syllables, this means having primary stress on the syllable, as in

However, there is a substantial class of exceptions to this basic pattern, in which polysyllabic nouns have final stress. Among these are nouns which have final syllables which are written with a double vowel letter, as in . Another set of polysyllabic nouns stressed on the final syllable have been borrowed from French; they contain the French endings - , as in Words such as these are mostly morphologically simple in contemporary English. That is, they do not really have a suffix: while there is a morpheme , a cigarette is not a small cigar, for instance.[3]

There is a set of nouns which have consonant clusters after the penultimate vowel, and these have primary stress, as in . There is also a set of nouns which have three or more syllables and which end in - . These too tend to have penultimate stress, as in

.

Finally, there is a set of loanwords ending in a vowel which depart from the basic antepenultimate pattern, and take penultimate stress, such as

.

Again, the basic pattern is , with stress, as in . However, there is a set of adjectives with a consonant cluster after the penultimate vowel, and these take penultimate stress, as in

These include words which have an 'rC' cluster in rhotic accents, and which used to have an 'rC' cluster in non-rhotic accents, as in

Polysyllabic adjectives ending in - have antepenultimate primary stress, as in . These are therefore unlike bisyllabic verbs ending in - , which have primary stress on the - (but American speakers have penultimate primary stress in some of these words, as we have seen).

As with bisyllabic verbs, polysyllabic verbs often flout the basic trochaic English word stress pattern: there are many verbs with three or more syllables which have primary stress, as in . As with bisyllabic verbs, many of these have etymological prefixes which are Latinate in origin (such as - and).

Unlike bisyllabic verbs ending in - , polysyllabic verbs ending in - follow the basic antepenultimate primary stress pattern, as . These verbs thus have the same stress pattern as polysyllabic adjectives ending in - .

We have now identified four factors which may play a part in word stress assignment in morphologically simple words in English. Firstly, the syntactic category of the word may play a role: we saw, for instance, that many verbs depart from the basic trochaic pattern. Secondly, we saw that the presence of historical (etymological) prefixes of a Latinate origin can affect the stress pattern of a word: we saw that Latinate affixes such as , although they are mostly no longer productive prefixes in present-day English, typically fail to take primary stress. Thirdly, we saw that spelling plays a role in word stress. For instance, words such as and all exhibit primary stress on a final syllable containing a double vowel letter. Connected with this is a fourth factor, i.e. the presence of loanwords in English: words such a and have retained the stress pattern of the language they were borrowed from. Note too that loanwords the existence of Latinate affixes are at work in the stress patterns of words such as and .

We may now return to the stress patterns we considered in 8.2. Recall that we noted that English speakers tend to mis-stress bisyllabic French words such as ('to eat') and ('boat') by placing the stress on the penultimate syllable. We can now see the reason for this: it is because that is the basic pattern for bisyllabic words in English. We also noted that English speakers tend to place primary words stress on the antepenultimate syllable of polysyllabic French words such as ('closure') and ('building'). The reason is that this is the basic pattern for English polysyllabic words. That is what is typically English about those mis-stressings. In the pronunciation of polysyllabic words such as and , the spelling plays a role: the existence of two consonant letters in the spelling of the penultimate syllable of the word in English often shifts the stress away from the basic antepenultimate stress pattern to a penultimate stress pattern. Finally, polysyllabic words such as and fall within the class of loanwords ending in a vowel which exceptionally take penultimate primary stress.

Let us now consider word stress in morphologically complex words.

8.5 Word Stress Assignment and Morphological Structure

English has both suffixes (as in , with the - suffix) and prefixes (as in , with the prefix). Let us begin with suffixes. These may be subdivided into **inflectional** and **derivational** suffixes. The addition of an inflectional suffix is often said to produce ‘a different form’ of the word one would have if the suffix had not been added. For instance, when the suffix - is added to the verb , the resulting word, (as in), is said to be a of that verb; when the plural suffix is added to the noun , the resulting word is a of that noun. But when a derivational suffix is added to a word, it is said to produce not a different form of the same word, but another word. Thus, when the suffix - is added to an adjective, say , the result, the adverb , is a distinct word. Similarly, when the suffix - is added to an adjective, as in , the result is a distinct word. Other examples of derivational suffixes in English are - (as in), - (as in), - (as in), - (as in), - (as in), - (as in), - (as in), etc. Inflectional suffixes are not stressed, and have no effect on word stress in English words, as can be seen from pairs such as ' / ' (plural suffix), ' /' (past tense suffix), ' /' , (progressive suffix), and ' /' (present tense suffix).

Among the English derivational suffixes, some have no effect on stress when added to a word, while others do affect the stress pattern. These two classes of suffix are referred to as the **stress-neutral** and **stress-shifting** suffixes, respectively. The stress-shifting suffixes are all of Latinate descent, i.e. from French or Latin. But not all Latinate suffixes are stress-shifting. Let us begin by considering stress-neutral suffixes.

8.5.1 Stress-Neutral Suffixes

English is a Germanic language; all native Germanic suffixes are stress-neutral. These include the adverbial suffix - , as in the words ' , ' , ' , ' , ' , ' , ' , etc. We can see here that the stress does not shift when the suffix is added to the monosyllabic adjective from which the adverb is derived ((' , ' , ' , etc.). The stress pattern of these adverbs is therefore the basic trochaic pattern for English bisyllabic words: stress on the penultimate syllable. Polysyllabic adverbs ending in - include ' , ' , ' , ' , ' , ' . Again, the primary stress remains where it was in the adjective which the adverb is derived from: ' , ' , ' , ' , ' , ' . In all of these cases, the resulting adverb is a polysyllabic word with the basic trochaic pattern for polysyllabic words, i.e. antepenultimate stress. However, the primary stress will fall earlier than the antepenultimate syllable if the deriving adjective has primary stress on its antepenultimate syllable or earlier, as in ' , derived from ' which has three syllables in RP, but four in GA. Note that adjectives such as and , which have a syllabic /l/, as in [ˈaɪdl̩], will lose a syllable when - is added: is pronounced [ˈaɪdli].

The native Germanic adjectival suffixes such as - also have no effect on stress, so when they are added to monosyllabic roots, the result is a trochaic stress pattern, as in ' /' , ' /' , ' / ' , ' /' , ' /' , ' /' . When these stress-neutral suffixes are added to bisyllabic words with the basic trochaic stress pattern, the result is an antepenultimate stress pattern, as in ' /' , ' , ' , ' /' , ' /' , ' / ' . Once again, in cases where a syllabic consonant is possible in the deriving word (e.g. : [bʌbḷ]), a syllable is lost when - is added, as in ' /' , ' / ' , ' /' , ' /' , ' /' . The result is therefore a bisyllabic adjective with penultimate stress.

In addition to the - suffix which marks the comparative form of adjectives (as in), there is a stress-neutral native Germanic - suffix which can be used to form nouns from verbs, as in ' /' , ' /' , ' / ' , ' / ' , ' /' , ' /' . We can see from these examples that the addition of this - suffix has no effect on stress.

Other native Germanic suffixes, used to form nouns, are - , - , - , - , - , and - , as in ' /' ,

' /' , ' /' , ' /' ,[4]
' /' , ' /' , ' /' ,
' /' , ' /' , ' /' , ' /
' , ' /' , ' / ' .

Once again, the addition of the suffix does not shift the stress.

There are Latinate suffixes which fail to shift stress. Among these are the adjective-forming bisyllabic suffix , pronounced [əbḷ], as in ' , ' , ' , ' , ' . We can see that, when is added to the morphemes ' , ' , ' , ' , ' the stress does not shift. The result, in these cases, is an adjective with antepenultimate primary stress. If is added to a word which does not have primary stress on the final syllable, such as ' , ' , ' or ' the result is a word with primary stress placed prior to the ante-penultimate syllable: ' , ' , ' , ' .

8.5.2 Stress-Shifting Suffixes

Among the stress-shifting derivational suffixes, we may distinguish between those on which the primary stress falls, and those which shift the stress the base form to which the suffix is attached. Let us begin with suffixes on which the primary stress falls.

The suffixes and all take primary stress, as in ' /ˌ ' ' /ˌ ' and ' / ˌ '). Notice that, in polysyllabic words containing one of these suffixes, we cannot leave the derived word with a sequence of unstressed syllables at the beginning: a secondary stress must occur. This is the Rhythmic Principle introduced in 8.3. The Derivational Principle and the Stress Clash Avoidance Principle may also play a role here. Take the pair : the suffix takes primary stress, but the Rhythmic Principle says that we cannot leave the resulting word with a sequence of unstressed syllables at the beginning of the word. The Derivational Principle says to put that secondary stress on the syllable which had primary stress in the deriving word, in this case on the syllable which is the final syllable of the verb . Thus the stress pattern ' ˌ . There is no violation here of the Stress Clash Avoidance Principle, since the secondary and primary stresses do not fall on adjacent syllables. Recall, however, that in a word such as , there is a clash between the Derivational Principle and the Stress Clash Avoidance Principle. Once we have placed primary stress on the suffix, the Rhythmic Principle insists on a secondary stress. The Derivational Principle says that this secondary stress should fall on the syllable containing the primary stress in the deriving word ' . But if we were to place the secondary stress there, this would violate the Stress Clash Avoidance Principle. As we have seen, when there is a conflict between those two principles, it is the Stress Clash Avoidance Principle which predominates. Thus the stress pattern ˌ ' . The same situation arises for words such as ˌ ' in which the secondary stress does not fall where the primary stress falls on the word ' .

We have argued that many words borrowed from French which have an ending in the spelling, such as and , are not really morphologically complex: it is hard to argue that they contain the morphemes and . However, we allowed that there are some words which are clearly morphologically complex, as in the word . Our grounds for doing so were that there is clearly a morpheme in a word like this, and a kitchenette is indeed a small kitchen. Given that we have allowed this, we can say that, at least in some words, there is an suffix which takes primary stress, and is thus parallel to and . We saw that, where has a clear meaning, it can mean either 'little' (as in and) or 'female' (as in the word). There are some words in which it is not entirely clear whether there is a proper suffix or not. While it might seem reasonable to say that an usherette is a female usher, it is not clear that a maisonette is a small maison (the French word for 'house'), although it's certainly a small house.

Examples of stress-shifting suffixes which do not themselves take primary stress are and , as in ˈ /ˌ ˈ , ˌ ˈ /ˌ ˈ , ˈ / ˈ , ˈ / ˈ . In each case, the primary stress falls on the syllable immediately preceding the stress-shifting suffix.

Further examples of suffixes which shift stress are (as in ˈ /ˌ ˈ) and (as in ˈ /

ˈ). Again, the primary stress falls on the syllable immediately preceding the stress-shifting suffix.

You may already have noted that, when the stress in a word shifts as a result of the addition of a stress-shifting suffix, this can have the effect of changing the pronunciation of the vowel in the affected syllable. Thus, in , the final syllable, in the suffix, has a schwa vowel ([ə]) while, in , the antepenultimate syllable, again in the suffix, has an [æ] vowel, since that syllable bears primary stress in . Consonantal changes in the base form may also occur when suffixes are added. For instance, when is added to the adjective , the resulting form has an [s], rather than a [k], at the end of the base.[5] Such changes are not limited to stress-shifting suffixes. For instance, when the suffix is added to , the resulting form, , has an [s], rather than a [t], at the end of the base. Clearly, the more variation there is, in terms of both stress pattern and consonant and vowel realization, between base form and affixed form, the less evident it will be that the base and affixed forms are actually related: while the relationship between, say, and (with no stress, vowel or consonant changes) is transparent, that between, say, and is much less so. It is for this reason that many people prefer to add affixes of the sort (which tend to be native to the Germanic family of languages to which English belongs, unlike the sort, which have their roots in the Latinate languages) to base forms.

A striking property of unstressed syllables in English is that they often contain the schwa vowel, transcribed as [ə], which is ‘less distinct’ perceptually than most other vowels. It is a common feature of related words such as and that, when the stress shifts from the first to the second syllable by virtue of the addition of the suffix, the vowel in that syllable changes from being a ‘full’ vowel to being a ‘reduced’ schwa. We will look at this kind of phenomenon in a little more detail in the following chapter.

8.5.3 Word Stress Patterns and Prefixes

Let us now turn to the stressing of prefixes. We will take the view that most separable monosyllabic prefixes bear secondary stress. By ‘separable’, we mean that, if the prefix is removed, we are left with an existing English word, as in the verbs ˌ -ˈ , ˌ -ˈ , ˌ -ˈ , ˌ ˈ , ˌ ˈ . Other monosyllabic prefixes include:

co- ('together')	as in *ˌco-conˈspirator, ˌco-ˈedit*
de- ('get rid of/reverse')	as in *ˌde-reguˈlation, ˌde-ˈlouse*
dis- (negative)	as in *ˌdisaˈppear, ˌdisˈpleasure*
ex- ('former')	as in *ˌex-adˈministrator, ˌex-ˈboss, ˌex-ˈserviceman*
in- (negative)	as in *ˌincoˈrrect, ˌinˈactive* (see too *il-*, as in *ˌilleˈgality, ˌiˈlliterate, im-*, as in *ˌimperˈturbable, ˌimˈproper*, and *-ir*, as in *ˌirreˈsistible, ˌiˈrregular*)
mal- ('badly')	as in *ˌmalaˈdjusted, ˌmaˈlodorous*
mis- ('wrongly')	as in *ˌmis-aˈddressed, ˌmis-ˈspelled*
pre- ('before')	as in *ˌpre-eˈxist, ˌpre-ˈpay*
pro- ('in favour of')	as in *ˌpro-ˈhunting, ˌpro-ˈchoice, ˌpro-ˈlife*
re- ('again')	as in *ˌre-aˈppear, ˌre-ˈfill* (verb)
sub- ('beneath')	as in *ˌsub-aˈtomic, ˌsub-ˈhuman*
trans- ('across')	as in *ˌtrans-Atˈlantic, ˌtranˈsexual*
un- (negative)	as in *ˌunaˈttractive, ˌunˈfair*

It is striking that some of these can be used as independent words, as in 'I'm having dinner with my ex tonight' and 'Are you with the pros or the antis?'

Bisyllabic prefixes can form a trochaic foot, and so, in accordance with the Rhythmic Principle, will have the penultimate syllable of the prefix bearing secondary stress, as in ˌ ˌ , ˌ ˌ , ˌ ˌ , ˌ ˈ , ˌ ˈ , ˌ - ˈ , etc. While these all have secondary stress, there are some cases where there is primary stress on bisyllabic prefixes, as in and . It is perhaps wise to consider these latter cases as compounds (words made from two or more words), to which we turn shortly. Equally, while can have secondary stress, as in ˌ ˈ , there are clear cases where it has primary stress, as in ˈ ˈ ˈ . These too may be used as independent words, as in 'The antis are out in force' or 'That film was absolutely mega!' Other bisyllabic prefixes are , as in , and , as in .

Prefixes such as those just listed are of Latinate origin: they occur in words which have been borrowed directly from Latin, or from words borrowed from French, which is derived historically from Latin. As we have seen, there are a good many bisyllabic noun–verb pairs in which the verb is stressed on the final syllable, whereas the noun is stressed on the Latinate prefix. Examples are ˌ ˈ (verb) vs ˈ (noun), ˌ ˈ (verb) vs ˈ (noun), ˌ ˈ (verb) vs ˈ (noun). It is not clear that such words are still morphologically complex in contemporary English: for example, one cannot separate the - of and arrive at a verb . But these words certainly have elements which are, etymologically, prefixes. There are exceptions to the generalization that the verbs in such pairs are stressed on the final syllable, and the nouns on the etymological prefix: some pairs, such as and , have the verbal pattern, while others, such as and , have the noun pattern. However, the differential verb-vs-noun stress pattern still seems to be productive in contemporary English. This can be seen from neologisms. Take the verb ˈ : it conforms to the pattern. The noun has existed for some time, but a more informal term has been coined: an ˈ . Interestingly, when the new noun was derived from the existing verb, the stress shifted to conform with the noun pattern, suggesting that speakers have access to the noun-vs-verb stress patterns.

8.6 Compound Words

Compound words are, put simply, words which can be analysed as consisting of two (or more) words, rather than as containing a base and an affix. For instance, while is a compound, is not (the form is a suffix, not a word). We will focus on two-part compounds here. The Compound Stress Rule in English says, of two-part compounds: of the two elements, the first is the most prominent. Two-word thus have the opposite pattern to two-word , such as the noun phrase (a bird which is black), the adjective phrase , verb phrases such as , adverb phrases such as and prepositional phrases such as . These all exhibit the English Phrasal Stress Rule (which we will return to in chapter 9): in all of these, it is the second element which is most prominent. Examples of compounds which have the regular compound stress pattern, with the first element the most prominent, are:

How can we tell whether a given two-word sequence is a compound or a phrase? When the two parts are written as one word (e.g.) or with a hyphen (e.g.), it is easy to see that one is dealing with a compound. But if the two parts are written as separate words (e.g.), it is less easy. (There is also variation in how compounds are written: one may find, for example, the written forms

.) In many cases, compounds have a different kind of semantics (meaning) from phrases. Take the phrases and . Compare their meanings with the compounds and

. While all (male) blackbirds are black birds, not all black birds (phrase) are blackbirds (compound): ravens, jackdaws and cormorants are black birds (phrase), but they are not blackbirds (compound). While all darkrooms are rooms which are normally dark, not all dark rooms are darkrooms (places for developing photographs): if I close the shutters and switch off the lights in my study, it becomes a dark room (phrase), but not a darkroom (compound). A green house (phrase) is a house which is painted green, but a greenhouse (compound) is not a house, and may be painted white. It seems likely that, in the history of English, compounds started off as phrases: a (male) blackbird is indeed black, a darkroom is indeed normally dark, and a greenhouse is a house-shaped structure where one grows green things. But such phrases have made the transition to becoming single words.

8.6.1 Exceptions to the Compound Stress Rule

(a) two-part place-names, such as

These include street names, e.g.

. The only exception to this are large numbers of street names ending in the word : in my home town of Edinburgh, London Road (where is most prominent) is not far from London Street (where is most prominent).

There are sets of compounds which systematically violate the Compound Stress Rule. We now list these.

(b) compounds with a participial second element, ending in or . The compounds are more common than the latter. Many of these are based on parts of the body, and some of these are more metaphorical than others. Examples are:

' , ' -' , ' -' , ' -' ,
' -' , ' -' , ' -' , ' -
' , ' -' , ' -' , ' -' .

(c) compounds in which the first part expresses what the object is made from. Examples are:

. Notice that these are distinct from similar compounds in which the first part does not express what the object is made from, as in , which is not made from paper, , which is not made from cotton, and , which is not made of olives. Note that, in American English, the first element can be the most prominent in such compounds.

(d) compounds in which some kind of concrete or abstract positioning is involved. Examples are:

.

(e) compounds which are two-part colour words. Examples are:

.

(f) compounds derived from phrasal verbs. Examples are compound nouns derived from phrasal verbs:

(from), (from),
(from), (from); and compound adjectives derived from phrasal verbs:

(from), , (from

), (from), ,
(from).

8.7 Summing Up

We have covered a fair amount of detail in this chapter. It may prove helpful to the reader to have the main points summarized here, ignoring many details and exceptions: if the reader can grasp the take-home message for each of the points covered in this chapter, the details can then be mastered by consulting each section of the chapter.

- English word stress is not random.
- English rhythm is trochaic, as in and .
- Primary stress is calculated from the of the word, not the beginning.
- English words cannot begin with more than one unstressed syllable.
- When one English word is derived from another, and the primary stress shifts as a result of the derivation, there is a tendency to place the secondary stress on the syllable that had primary stress in the deriving word, as in ˌ ˈ .
- There is a tendency to avoid placing primary and secondary stresses next to each other, as in ˌ ˈ .
- While English nouns, adjectives and adverbs mainly follow the basic trochaic pattern, there are many verbs which do not.
- English suffixes may be divided into those which affect

the primary stress (such as - and -) and those which do not (such as -).

• Among the suffixes which affect primary stress, some take the primary stress (such as -), while others do not (such as -).

• Separable prefixes normally take secondary stress, as in ˌ -ˈ .

• The basic pattern for two-part compounds is: the first element is the most prominent, as in .

Notes

1 To the extent that any such endings may be productive, and thus may result in neologisms, then the resulting words will indeed be morphologically complex. For instance, the fairly recently coined word (in Britain, a female lad: a young woman who behaves in a loud, foul-mouthed, heavy-drinking, sexually promiscuous manner) is morphologically complex (here means 'female', rather than 'little').

2 It might be argued that words such as are morphologically complex, given the existence of the word . We take the view that - does not constitute a morpheme in contemporary English, and that words such as and are therefore morphologically simple.

3 There is, however, a case for saying that words such as are morphologically complex, since is clearly a morpheme, and a kitchenette is indeed a small

kitchen.

4 Note that is typically stressed on the final syllable in American English.

5 The word also exhibits an alternation, between [eɪ] and [æ], in the stressed vowel of the base. This is one of a set of vowel alternations which we do not examine here.

6 The expression (a bunch or sachet of herbs used in cooking) is pronounced with primary stress on the penultimate, rather than the final, syllable of . This is a result of the process of iambic reversal, described later in our discussion of metrical structure.

Exercises

Listen to sound files online

1 Listen to Track 8.1 at www.wiley.com/go/carrphonetics. For each of the bisyllabic words on the recording, say which ones have the default trochaic stress pattern for primary stress. For those which deviate from that pattern, explain why. The words are:

(a) famine

(b) Maltese

(c) migrate

(d) trainee

(e) winter

(f) explain

(g) silly

(h) compact (noun)
(i) compact (verb/adj.)
(j) export (noun)
(k) export (verb)
(l) stumble
(m) fancy
(n) differ
(o) taboo
(p) gazette
(q) arcade
(r) burlesque

2 Listen to Track 8.2. For each of the sets of polysyllabic words on the recording, say which have the default Germanic trochaic pattern. For those which deviate from the default pattern, explain why. The sets are as follows:

(a) factory
America
family
academy
stimulus

(b) develop
inherit
complicit
explicit
inherit

(c) kangaroo
employee
engineer
seventeen
mountaineer

(d) interrogate
investigate
accommodate
demonstrate
co-ordinate

(e) mathematics
physics
periodic
linguistics
alcoholic

(f) autumnal
sentimental
orchestral
horizontal
universal

(g) hostility
austerity
modernity
humility
ambiguity

(h) banana
bikini
karate
martini
piano

(i) adventure
amalgam
consensus
November
advantage

(j) momentary

secretary
literary
laboratory
military

3 Listen to Track 8.3. Each of the nouns on the recording (all of them loanwords into English) has a stress pattern which deviates from the default Germanic pattern. In each case, say in what sense that stress pattern is exceptional with respect to the rules of word stress assignment in English. Primary and secondary stresses are marked as follows: [ˌhəʊˈtɛɫ] (, with primary stress on the final syllable and secondary stress on the penult). Where the GA and RP pronunciations differ, this is indicated, as in [ˌhəʊˈtɛɫ]/[ˌhoʊˈtɛɫ], which gives the RP pronunciation followed by the GA pronunciation.

(a)
hotel
([ˌhəʊˈtɛɫ]/[ˌhoʊˈtɛɫ], not [ˈhəʊtɛɫ]/[ˈhoʊtɛɫ])
(b)
bouquet
([ˌbʊˈkʰeɪ], not [ˈbu:keɪ])[6]
(c)
bamboo
([ˌbæmˈbu:], not [ˈbæmbu:])
(d)
champagne
([ˌʃæmˈpʰeɪn], not [ˈʃæmpeɪn])
(e)
bikini

([bɪˈkʰiːni], not [ˈbɪkʰini])
(f)
martini
([mɒːˈtʰiːni]/[mɒɹˈtʰiːni], not [ˈmɒːtini]/[ˈmɒɹtini])
(g)
chorizo
([tʃəˈɹiːzoʊ], not [ˈtʃɑɹɪzoʊ])

4 For non-native speakers: listen to Track 8.4 and repeat each utterance. For native speakers and non-native speakers alike: indicate primary stresses and secondary stresses on the transcriptions of those utterances, given below. For example:

Mary finds Bill’s book uninterpretable. ˈmeəɹi ˈfaɪndz ˈbɪłz ˈbʊk ˌʌnɪnˈtɜːpɹətəbł̩

(a)
Mathematics is incredibly difficult. mæθəmætɪks ɪz ɪŋkɹɛdɪbli dɪfɪkəłt
(b)
My car was made in America. maɪ kɒː wəs meɪd ɪn əmɛɹɪkə
(c)
His computer is Japanese. hɪz kəmpjuːtəɹɪz dʒæpəniːz
(d)
Academic conversation is dull. ækədemɪk kɒnvəseɪʃn̩ ɪz dʌł
(e)
The police will interrogate the detainees. ðə pəliːs wɪł ɪntɛɹəgeɪt ðə diteɪniːz
(f)
They don’t produce many exports. ðeɪ dəʊnt pɹədjuːs

mɛni ɛkspɔ:ts

(g)

They don’t export much produce. ðeɪ dəʊnt ɪkspɔ:t mʌtʃ pɹɒdju:s

(h)

I found the film rather sentimental. aɪ faʊnd ðə fɪłm ɹɒ:ðə sɛntɪmɛntl̩

5 For non-native speakers: listen to Track 8.5 and repeat the words you hear. For native speakers and non-native speakers alike: explain the word stress patterns of the following groups of English bisyllabic words, as heard on the recording.

(a) happen
woman
fancy
echo
father

(b) deny
inspect
comply
expand
inflect

(c) trainee
bamboo
bazaar
taboo
shampoo

(d) create
migrate

locate
frustrate
narrate

(e) produce (verb)
export (verb)
discharge (verb)
object (noun)
contract (noun)

(f) produce (noun)
export (noun)
discharge (verb)
object (noun)
contract (noun)

6 Listen to Track 8.6. Explain where the metrical foot boundaries fall in the following sentences, as heard on the recording. (Draw vertical lines where the foot boundaries fall.)

(a) Leave me alone!

(b) Leave me a slice!

(c) She left in a hurry.

(d) She lives in America.

(e) Put it in the refrigerator!

(f) John’s a modern metrosexual.

(g) Clinton opposes militaristic solutions.

(h) Aude is a flexitarian.

9

Rhythm, Reversal and Reduction

9.1 More on the Trochaic Metrical Foot

We said, in chapter 8, that the rhythm of English is : the basic rhythmic pattern consists of a stressed syllable followed by zero or more unstressed syllables. For instance, in the phrase , the metrical structure is [ˈmeɪdɪnəˈfæktəɹi]. The two trochaic feet here are [ˈmeɪdɪnə] and [ˈfæktəɹi]. We assumed too that syllables with secondary stress also form trochaic metrical feet, as in the word : [ˌækəˈdɛmɪk]. The two trochaic metrical feet here are [ˌækə] and [ˈdɛmɪk]: the secondary stress in [ˌækə] forms a trochaic metrical foot with the following unstressed syllable, and the primary stress in [ˈdɛmɪk] forms a trochaic metrical foot with the following unstressed syllable.

But what is the evidence for the metrical foot? And what evidence is there for our claim that all feet in English are trochaic? We will now address these questions.

9.1.1 Evidence for the Trochaic Metrical Foot (a): Rhyming

Although we have identified a constituent within the syllable widely known as the rhyme, the term is a misnomer: this constituent is the unit on which rhyming in English is based. While it is true that rhymes with , and that both contain the rhyme [æd], we must not be misled into thinking that two words rhyme only if they have identical rhyme constituents in the syllable, in this case [æd]. Consider the words and : they rhyme because they both have a trochaic metrical foot of the same sort: [ˈwɪti] and [ˈsɪti]. Clearly, onset consonants play no role in rhyming, but metrical structure , and the rhyme constituent does not. The reason why does not rhyme with either or is that the metrical foot structure of is [ˈɛntɪti]: the word does not contain a metrical foot of the shape [ˈɪti]. Rhyming in English is based on identity of stressed vowels in two or more trochaic metrical feet, and identity of all subsequent phonetic segments.[1] The word does not have the same stressed vowel as the words and .

Similarly, the phrase ( ), which is pronounced [ˈfəʊnje] in RP and [ˈfoʊnje] in GA, rhymes with : [njuːˈməʊnje] in RP and [nuːˈmoʊnje] in GA. In this case, the rhyme is based on trochaic metrical feet which do not map directly onto word boundaries: [ˈfəʊnje] contains two words, while [ˈməʊnje] is part of a word.

9.1.2 Evidence for the Trochaic Metrical Foot (b): Expletive Insertion

Expletives such as and * are frequently used in informal spoken English by many speakers. Each consists of a standard Germanic bisyllabic trochaic metrical foot: [ˈblʌdi] and [ˈfʌkɪŋ] (also pronounced with a final [n] by many speakers of English, which may be syllabic: [ˈfʌkn̩]).

The patterns of use of words such as these in syntactic structure is, contrary to popular belief, quite complex. We will ignore that syntactic complexity and focus here on the fact that they can be inserted into the internal structure of words, as in , where is inserted into the word . This word is an adverb derived from the adjective , which has the basic antepenulti-mate word stress pattern of polysyllabic words in English: [ˈæbsəlu:t]. There is, however, an emphatic pronunciation with final primary stress and a secondary stress on the antepenult: [ˌæbsəˈlu:t]. It is in this emphatic form that the expletive can be inserted into, as in:

A:

B:

If you are a native speaker of English (and some non-native speakers will see this too), you will agree that it is not possible to reply , or . But why not? The answer lies in the existence of the trochaic metrical foot. The expletives are bisyllabic trochaic metrical feet. The words into which they can be inserted will contain trochaic metrical feet. In inserting expletives, one must respect the trochaic metrical structure of the word one is inserting an expletive into: contains a sequence of three trochaic bisyllabic metrical feet. It respects the trochaic bisyllabic metrical foot structure of , inserting another such foot in between those two.

9.1.3 Evidence for the Trochaic Metrical Foot (c): Neologisms

Many neologisms are based on the trochaic metrical foot. Here are some examples: . All of these neologisms are based on analogy with one or more previously existing words. If we assume that are all based on an analogy with the word , then we must ask what the basis of the analogy might be. In this case, it cannot be the morphological structure of the word , which contains the root and the suffix (which, as we have seen, shifts the stress from the word , pronounced [ˈæłkəhɒł]). If the morphology were the basis for the analogy, then the neologisms would be and .

The analogy in these cases is based on the metrical structure of the word ([ˌælkəˈhɒlɪk]), which contains two bisyllabic trochaic metrical feet: [ˌælkə] and [ˈhɒlɪk]. The reason why the forms and have to be written with an <a> is that this letter represents a schwa vowel ([ə]) which is present in the first trochaic metrical foot of , on which the analogy is based. Native speakers of English produce such neologisms because they have a (perhaps not entirely conscious) sense of the metrical structure of their native language.

Consider contemporary neologisms such as (a vegetarian who is prepared to be flexible, and eat meat from time to time) and (someone who doesn't eat meat, but who does eat fish, or seafood in general). These words have been formed by analogy with the word , pronounced [ˌvɛdʒəˈteəɹiən] in RP. The two trochaic metrical feet in this word are [ˌvɛdʒə] and [ˈteəɹiən]. It is this metrical structure that drives the analogical process which results in and : [ˌflɛksɪˈteəɹiən] and [ˌpɛksɪ ˈteəɹiən].[2]

Consider too the recent neologism (a metropolitan heterosexual man who is overly concerned with his physical apearance). In this case, the trochaic metrical foot plays a role: the word is pronounced [ˌhɛtɹəʊ ˈsɛkʃəɫ]. The rhyming process, based on the trochaic metrical foot, is again at work here: speakers know that the morpheme may be bisyllabic ([ˌhɛtɹəʊ]), as is the morpheme ([ˈsɛkʃəɫ]).[3]

The recent brand name (a cleaning product) is based on the word : [dɪˈtʰɜ:dʒənt] (RP), [dɪˈtʰɜɹdʒənt] (GA). The foot structure of this word has an initial extrametrical unstressed syllable, leaving [ˈtʰɜ:dʒənt] (RP)/

[ˈtʰɜɹdʒənt] (GA) as the trochaic metrical foot. We can then add the morpheme ([seɪf]) to this metrical foot to produce : [ˌseɪfˈtʰɜːdʒənt] (RP)/ [ˌseɪfˈtʰɜɹdʒənt] (GA).

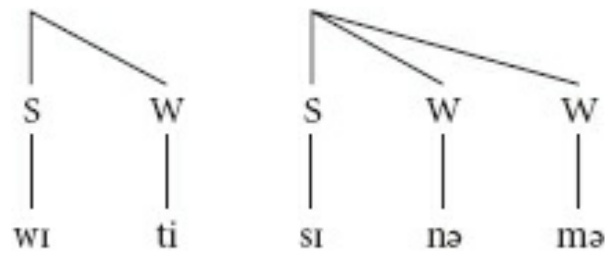
Similarly, the words ([ˈkɪsəgɹæm]: someone who delivers a telegram-type message with a kiss) and ([ˈstɹɪpəgɹæm]: someone who delivers such a message and strips) are formed by analogy with the word , whose metrical structure conforms to the basic trochaic structure for words of more than two syllables in English in that it has antepenultimate primary stress: [ˈtʰɛləgɹæm]. Although the morphological structure of is + , the neologisms are not and . The reason for this is that the neologisms are modelled on the metrical structure of . The English trochaic metrical foot clearly plays a role in all of the neologisms we have considered here.

9.2 Representing Metrical Structure

We have represented primary stress with a superscript diacritic, as in [ˈdɛmɪk], and secondary stress with a subscript diacritic, as in [ˌækə] in the word . These conventions will suffice if we confine our interest to the level of the word. But they will not suffice if we wish to represent the way levels of stress and relative perceptual salience operate when words are combined into phrases. Take the phrase , for instance. When appears in the phrase , the secondary and primary stresses in that word switch round. The single syllable of the word has more stress than any of the stressed syllables in , but in the word , the [kʰæŋgə] foot is less salient than the [ɹuː] foot. So we are dealing here with different levels of salience. That is not easy to represent using only the two diacritics we have used for word stress. We need a further mode of representation.

We represented syllable structure in terms of branching tree structures. Many phonologists also represent foot structure in terms of branching trees. We will represent any syllable which has any degree of stress with an 'S', indicating that it is strong with respect to weak unstressed syllables, which we label with a 'W'. A stressed syllable and any unstressed syllables with which it forms a foot may then be represented as follows:[4]

(1)



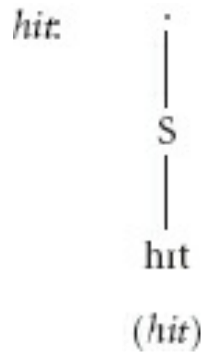
(*witty*) (*cinema*)

The bottom-most level of representation in this diagram is the level of the segment. The next level up is the syllable. At that level, the S labels represent strong (stressed) syllables and the W labels represent unstressed syllables. It is important to bear in mind that stress levels are : rather than a stressed syllable being definable in absolute terms, one syllable is more or less stressed .

The next level of representation up from the syllable (the lines above the S and W labels in this diagram) is that of the foot. Each word in (1) consists of a single foot, the first word consisting of a binary-branching foot and the second word consisting of a tertiary-branching (three-way-branching) foot.

Monosyllabic words contain, by definition, a single stressed syllable. We will take it that they contain a single, foot. We will therefore represent such words as having a single S-labelled syllable (indicating that it is stressed), dominated by a non-branching foot node, thus:

(2)



This diagram represents two claims. The first is that the syllable in question is stressed (labelled S). The second is that, because it is stressed, it constitutes a foot which happens not to have a branching structure, since there are no unstressed syllables following it.

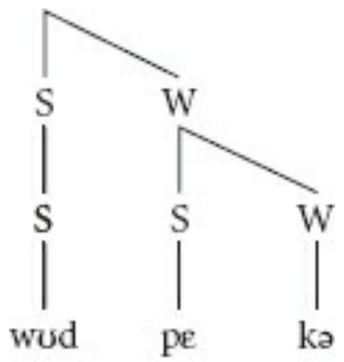
Monosyllabic non-lexical words, such as pronouns (e.g.), prepositions (e.g.), articles ( ) and conjunctions (e.g.) are typically unstressed. We will therefore represent them with a W-labelled syllable, but no foot structure above that level (since a foot by definition must contain a stressed syllable), thus:

(3)

(*it*)

The relational nature of stress levels can be seen clearly in words which have both a primary stressed syllable and a secondary stressed syllable, such as . It is clear that the antepenultimate syllable in this word has more stress than the penultimate (it is strong with respect to the penult). It is equally clear that the final syllable has less stress than the penultimate: it is unstressed, and thus weak. We may therefore represent the foot structure of the word as consisting of two feet, the first of which is stronger than the second, as follows:[5]

(4)

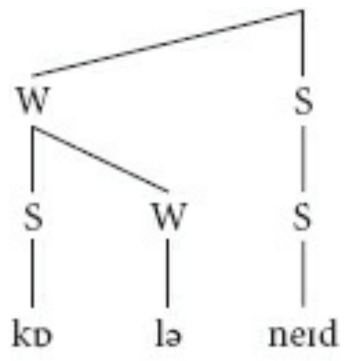


(*woodpecker*)

Note that the S/W notation is used to represent the relative strength of syllables within a foot of sequences of feet: the notation shows that the first syllable is strong with respect to the second, and also shows, at a higher level, that the first   is strong with respect to the second.

In words such as   and   , on the other hand, it is the second of the two feet which is the stronger, since it has a secondary stressed syllable followed by an unstressed syllable followed by a primary stressed syllable:

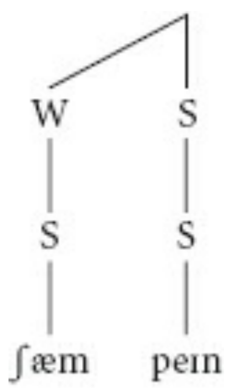
(5)



(*colonnade*)

Words such as   , which have a primary stressed and a secondary stressed syllable, but no unstressed syllables, contain two feet, each of which contains only a strong syllable. However, one of those feet is strong with respect to the other, thus:

(6)



(*champagne*)

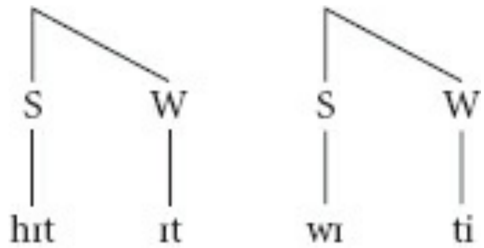
In this word, the second of the two feet is the stronger, whereas, in a word such as the noun   , it is the first of the two feet which is the stronger:

(7)

(*export*)

Recall that monosyllabic function words are typically unstressed (and are thus simply labelled with a W). Monosyllabic words of a lexical category may form branching feet with such words, as in the phrase . Thus a phrase such as contains exactly the same foot structure as a single word such as :

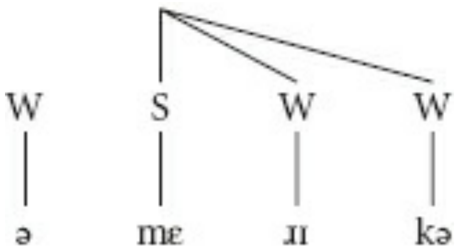
(8)



(*hit it*) (*witty*)

Thus, the constituent we have called the foot does not map directly onto the word: there may be more than one foot within a word, and a foot may extend beyond the span of a single word. Furthermore, a word may not be exhaustively divisible into feet. For example, words such as contain a foot consisting of the stressed antepunultimate syllable and the two unstressed syllables which follow it; the word-initial, unstressed syllable is a 'stray' unstressed syllable, which is part of the word, but is not integrated into the foot structure formed by the three syllables which follow it (rather in the same way that an /s/ preceding an onset consonant may be part of a word without being integrated into syllable structure):

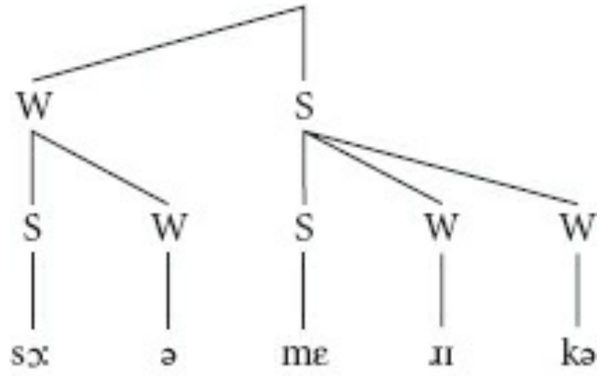
(9)



(America)

The word-initial, 'stray' (extrametrical) unstressed syllable here is parallel to the monosyllabic function words discussed above: at the level of the word, it is not integrated into foot structure. On this way of analysing English foot structure, words which begin with an unstressed syllable, such as and , do not consist of a single foot which begins with a W syllable, since we are denying that there are W-S feet in English. It is only at the level of larger units such as the phrase that such unstressed syllables may be integrated into foot structure,[6] as in the verb phrase :

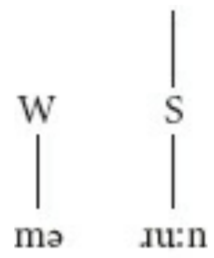
(10)



(*saw America*)

We saw that words like and are bisyllabic and contain two feet. These are fundamentally different from words such as , which are also bisyllabic but beging with a 'stray' unstressed syllable and contain only one foot, which consists of the stressed syllable:

(11)

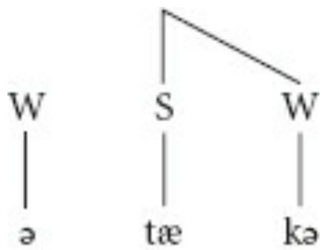


(*maroon*)

9.3 Phonological Generalizations and Foot Stucture

One of the reasons for postulating the foot as a phonological constituent is that, just as some phonological generalizations are sensitive to syllable structure, so some phonological generalizations are sensistive to foot structure. Take the generalization, or rule, of Flapping, in many dialects of American English. Under this generalization, /t/ and /d/ are realized as an alveolar tap (also known as a flap) between vowels, as in and ([bɛɾi]] and [bɛɾiŋ]). But the rule does not apply if a foot boundary occurs adjacent to the /t/ or /d/. Thus, the generalization does not cover cases such as , or , since, in those cases, a foot boundary intervenes between the first vowel and the /t/, thus:

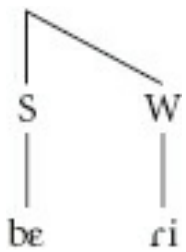
(12)



(*attacker*, *a tacker*)

If a foot consists, as we have said, in a stressed vowel followed by any immediately following unstressed syllables, then a word such as contains a single foot (which begins with the stressed syllable) preceded by a 'stray' unstressed syllable, as in the word-initial syllable of . Thus the word-initial syllable is not a part of the foot in which the /t/ appears, whereas in a word such as , it is:

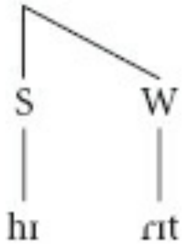
(13)



(*Betty*)

Note that Flapping also occurs in feet which are formed across word boundaries, as in ([hɪɾɪt]):

(14)

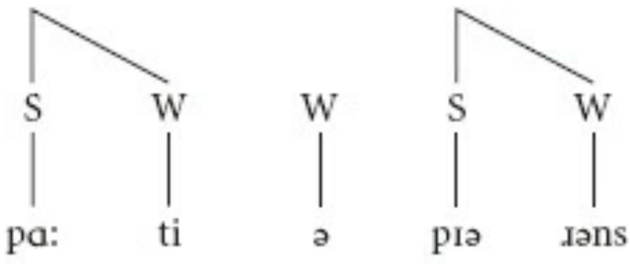


(*hit it*)

That is, Flapping is not sensistive to word boundaries or to the morphological structure; rather, it is foot structure which matters in the application of Flapping: Flapping only applies .

Another example of a generalization which is often said to be sensitive to foot structure is the rule of Aspiration. We have already noted that we must acknowledge that there are degrees of aspiration of voiceless stops in English. However, aspiration is at its strongest when the voiceless stop in question is in foot-initial position, as in and , whose foot structures are given in (15):

(15)



(*party*) (*appearance*)

Aspiration applies foot-initially (although there may be some degree of aspiration in other positions).

9.4 The Rhythm of English Again: Stress Timing and Eurhythmy

We saw, in chapter 8, that the rhythm of English is **stress-timed**. What this means is that the regular recurring beats found in the speech of English speakers (the rhythm of English speech) fall on stressed syllables. That is, stressed syllables in English occur at more or less equal intervals. Languages like English are often said to be distinct from languages like French in this respect in that, in languages like French, each syllable is said to occur at a more or less equal interval (languages of that sort are often, therefore, said to be).

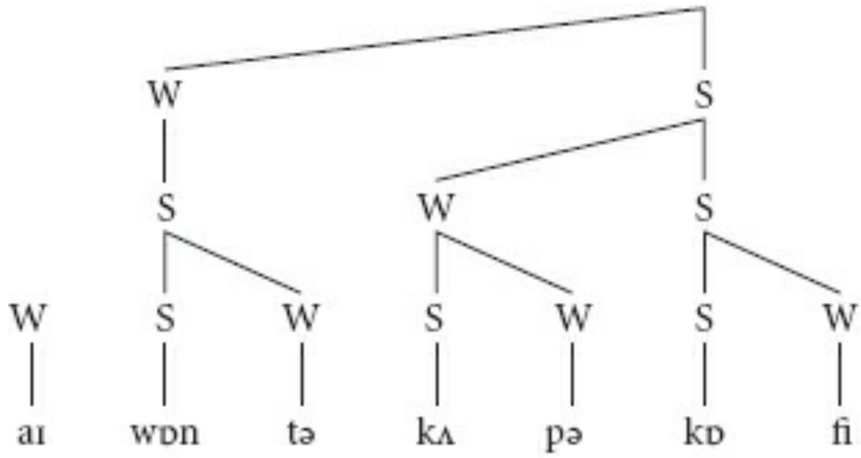
One of the consequences of this kind of rhythm is that English feet may consist of a stressed syllable followed by a sequence of unstressed syllables, as in the phrase , in which the stressed syllable in is followed by two unstressed syllables, or the phrase , where is followed by three, or the phrase , where it is followed by four.

Having said that English allows for really quite extensive sequences of unstressed syllables, it has to be said that the ‘ideal’ or optimal rhythmic structure is one in which strong and weak syllables alternate, in an S-W-S-W pattern. It appears to be the case that such sequences of ‘alternating opposites’ are optimal in a perceptual sense: they seem to make the speech signal more easily decoded. Such optimal rhythmic structures are often referred to as **eurhythmic** stuctures. It follows from this that the optimal, most eurhythmic, foot structure is a simple S-W structure, with only one unstressed syllable to the right of the stressed syllable. Foot structures with more than one W syllable are therefore less eurhythmic, less optimal, than those with only one, and the greater the number of unstressed syllables, the less eurhythmic or optimal the foot.

This preference for eurhythmy extends to sequences of feet: sequences of S and W feet are also more eurhythmic than other sequences. For instance, in the sentence

, there is an S-W-S sequence of three feet in the verb phrase, each of which is itself an S-W sequence of syllables; it is eurhythmic both at the level of sequences of syllables and at the level of sequences of feet:

(16)



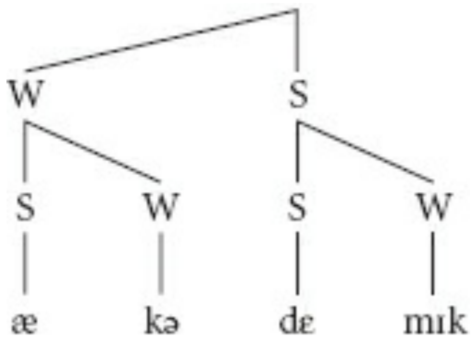
(*I want a cup of coffee*)

In many cases, however, a given combination of words may potentially create a phrase which is less than eurhythmic, and indeed may potentially result in adjacent S-labelled feet. This results from the fact that, in most English phrases, it is the final word which is most stressed, as in the phrase ,
discussed earlier. This **Phrasal Stress Rule** seems to hold for

most types of phrase in English, as in (verb phrase), (adjective phrase), (prepositional phrase) and (adverb phrase). It also seems to apply at the level of the sentence, as we can see from the example just given: the predicate verb phrase is more salient than the preceding subject noun phrase. Where the Phrasal Stress Rule brings about adjacent S-labelled feet, it appears that ‘evasive action’ can be taken. Let us consider some examples.

Take the words and . Clearly, has primary stress on the penultimate syllable and secondary stress on the first syllable; the other syllables are unstressed. The foot structure of the word is as follows:

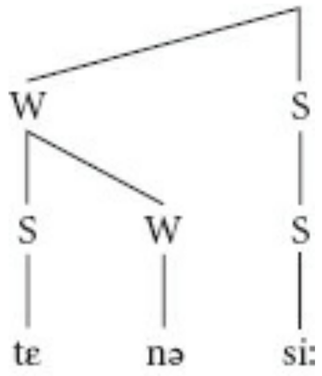
(17)



(*academic*)

also contains two feet, the second stronger than the first. However, the second foot consists simply of a stressed syllable, with no unstressed syllables following it:

(18)



(*Tennessee*)

also has two feet, as we have already seen, the first of which consists of a syllable with secondary stress and the second of which consists of a syllable with primary stress:

(19)

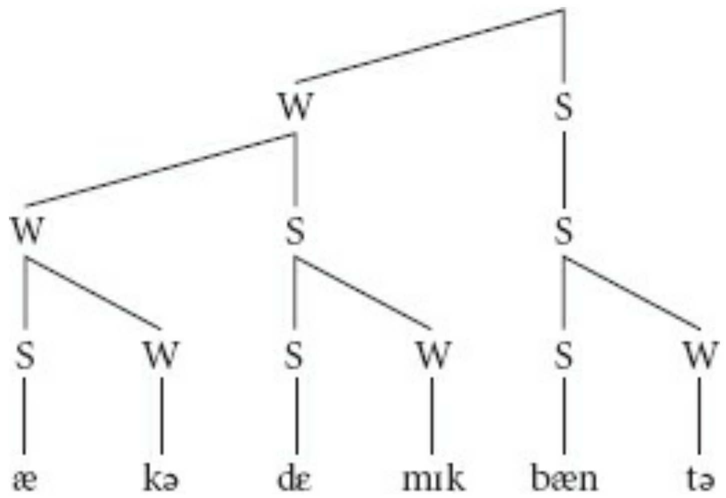
(*champagne*)

In each of these three cases, the word consists of two feet, the second of which is strong with respect to the first. However, when these words appear in phrases where the stronger of the two feet is immediately followed by the stressed syllable of another foot, and where that syllable must be more heavily stressed than the preceding one, a kind of 'stress clash' results, in which, rather than a eurhythmic sequence of S and W feet, an S-S sequence of feet occurs. In situations such as this, a rule of **rhythm reversal** applies. Consider some such phrases, e.g.

. Note that, in each case, the rule for phrasal stress assignment means that the second of the two words must have greater stress than the first. Note too that the primary and secondary stresses in the words

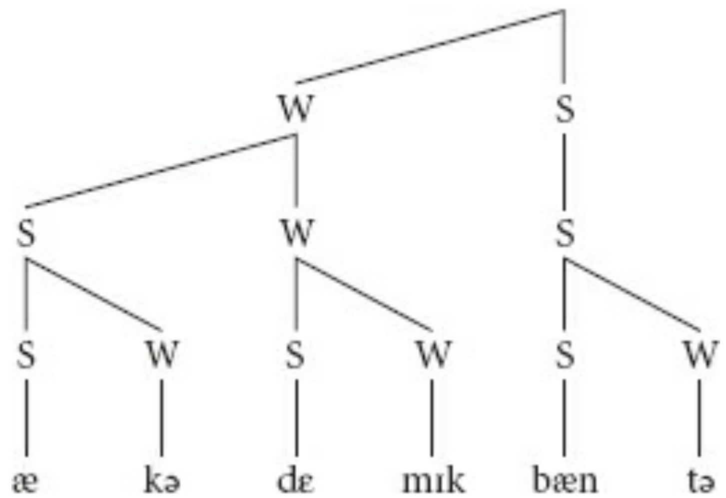
and have reversed. That is, the offending structure (exemplified in (20) below) is altered to the more eurhythmic structure exemplified in (21).

(20)



(*academic banter*)

(21)



(*academic banter*)

This process of rhythm reversal is quite regular in English.

Other examples are easily found; consider vs

vs vs

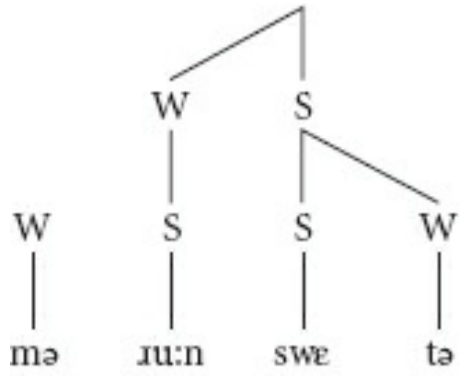
vs

vs , and so on. As we have seen, in English phrases, it is the head, rather than a preceding modifier, which bears the most stress. Rhythm reversal occurs whenever a word containing a weak–strong sequence of feet is combined, to form a phrase or compound, with a word whose first syllable is the first syllable of a foot (i.e. is stressed). That is, rhythm reversal operates, within the context of phrases and compounds, on feet, not syllables, reversing weak–strong sequences of feet, rather than weak–strong sequences of syllables. Another way of putting this is to say that the reversal process reverses a sequence of a secondary stressed syllable and a primary stressed syllable when it is followed by a primary stressed syllable within a phrase.

Reversal does operate on a sequence consisting of an unstressed syllable and the first syllable of a foot, as in

. The word contains a single foot, which consists only of a stressed syllable with no following unstressed syllables; that foot is preceded by a 'stray' unstressed syllable (just like the unstressed syllable in , shown above), thus:

(22)

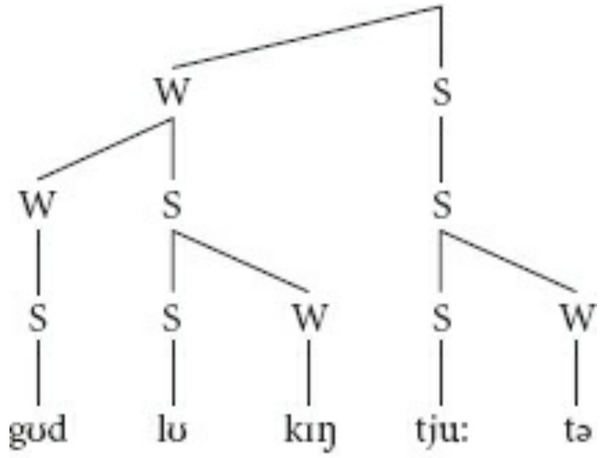


(*maroon sweater*)

While contains an S-S sequence of , it does contain an S-S sequence of , and it does not therefore undergo rhythm reversal. The pervasive effects of the Rhythm Reversal Rule are said to constitute evidence for the existence of the foot as a consistuent in the phonology of English. Furthermore, the fact that words like do not undergo reversal can be taken as evidence for our claims (a) that English feet always begin with a stressed syllable, and (b) that English words are not necessarily exhaustively divisible into feet. In other words, a word like is to be analysed as consisting of a foot with a W-S sequence of syllables.

The claim that rhythm reversal operates at the level of sequences of feet, rather than at the level of sequences of syllables, is supported by the fact that it operates in phrases such as , which, prior to reversal, has the following structure:

(23)



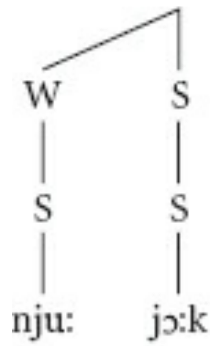
(*good-looking tutor*)

If reversal operated at the level of sequences of S and W syllables (rather than feet), it would not affect the sequence and , which have an alternating S-WS-W structure. It is at the level of the foot that the S-W sequencing is violated.

The most striking aspect of reversal is that it demonstrates the interaction of syntax and phonology. The conditions under which reversal operates are partly determined by a syntactic fact about English: the fact that modifiers typically precede heads in English phrases. This, combined with the fact that it is the head which receives more stress than the modifier, brings about the reversal phenomenon.

Reversal also interacts with morphological structure: it operates within words which contain a suffix which itself takes stress. Take the word . It consists of two feet, the second stronger than the first:

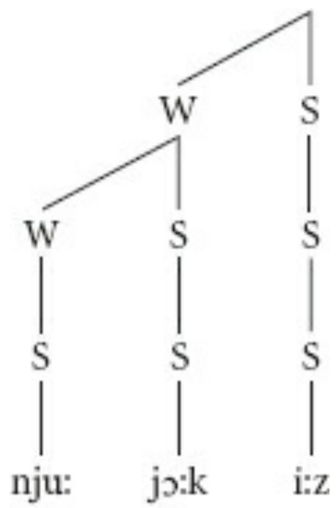
(24)



(*New York*)

The suffix is one of those English suffixes which takes stress. It consists of a single non-branching foot, and when it is added to , the resulting word consists of three feet in a W-S-S sequence:

(25)



(*New Yorkese*)

This structure meets the conditions for reversal, which then applies to yield the S-W-S sequence of feet in .

It is true, of course, that one can, in fact, utter phrases such as with primary stress on the penultimate syllable of . When one does this, one is usually contrasting some aspect of the phrase with some other possibility. One might be stressing, for instance, that one means banter and not some other kind of banter, such as

banter. The most important fact about this phenomenon, often referred to as contrastive sentence stress or contrastive intonation (see chapter 10 for details), is that it is a

: one could not make sense of such a stressing if it were not accompanied by an appropriate discourse setting in which one’s interlocutor stands a chance of understanding what other possibilities one is contrasting the rhythm-reversed adjective with. The following exchange is an example:

(26)

A:

I do enjoy academic banter, you know. (rhythm reversal)

B:

What kind of banter?

A:

ˌAcaˈdemic banter. (contrastive stress; no rhythm reversal)

The stress pattern here is nonetheless distinct from that given in (20) above, since the stressed syllable of is less stressed than the primary stressed syllable of .[7] What examples such as these suggest is that reversal is a metrical phenomenon which interacts with morphology and syntax, and can be described , whereas contrastive intonation is a phenomenon which be described independently of discourse context. This suggests that it is possible, and perhaps necessary, to distinguish those phenomena which can be analysed independently of context of utterance from those which cannot.

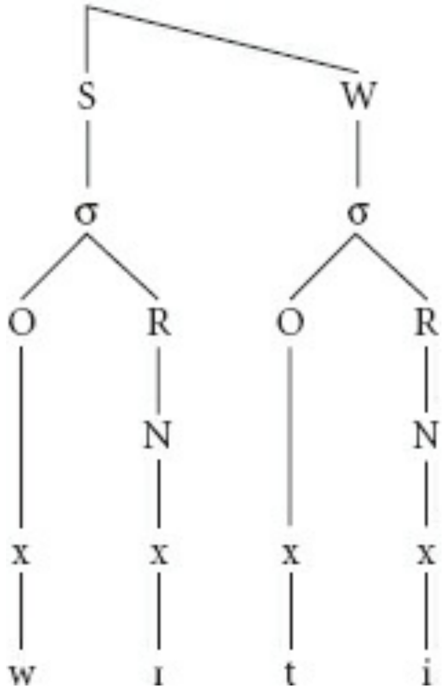
Notes

1 There are also imperfect rhymes, where the phonetic segments following the stressed syllable are similar but not identical. The imperfect rhyme between the words ([ˈmænə]) and ([ˈbæŋə]) are examples of this: the place of articulation of the nasals is not the same, so would rhyme better with . But it's not that bad a rhyme, after all: both [n] and [ŋ] are nasals, so they sound very similar.

2 There are other factors at play here, including the availability of the morpheme (as in , itself formed by analogy with ; itself comes from the adjective e, in which the prefix is). The term depends on the availability of knowledge of the Latin word for 'fish'.

3 In this case, the morphology and the metrical structure coincide.

4 The metrical trees we present here are abbreviated, for reasons of lack of space. Because metrical structure is determined by syllable structure, we ought, strictly speaking, to show metrical trees built upon syllable structure trees, as follows:



The abbreviated trees used in this chapter do not actually show that it is the structure of the rhyme, and not the entire syllable, which is crucial in determining metrical structure. However, they will suffice for our purposes.

The abbreviated trees used in this chapter do not actually show that it is the structure of the rhyme, and not the entire syllable, which is crucial in determining metrical structure. However, they will suffice for our purposes.

5 We label the non-branching foot with an ‘S’, consistent with our treatment of monosyllabic words of a lexical category.

6 The possibility of non-alignment of foot structure and word structure in English is often exploited by songwriters, who structure their songs on the basis of syntax (words and phrases), rhythm (metrical structure) and various types of repetition, such as alliteration, rhyming and repetition of metrical and phonemic structures. It is for this reason that it is possible to construct a repetition involving

([ˈmænəz]), ([ˈgɹæmə]), and the word followed by the first, stray, unstressed syllable of (yielding a foot whose form is [ˈdæmə]). Here, a repetition exploits syllable and foot structure while cutting across word and phrase structure. The sort of repetition known as ‘rhyming’ can therefore extend beyond the syllabic constituent we have called the rhyme, into foot structure. The rhyming of and in the song ‘Diamonds are a girl’s best friend’ is a further example.

7 It is often suggested that this follows from the fact that ‘banter’ is somehow ‘given’ in the discourse context.

Exercises

1 Draw metrical trees for each of the following words. Begin by drawing the foot structure. Where a word contains more than one foot, draw a superordinate S-W or W-S branching structure, showing which of the two feet is stronger.

(a) pretty
(b) collided
(c) sentiment
(d) bat
(e) nightingale
(f) kangaroo
(g) rabbi
(h) contract

Listen to sound files online

2 Listen to Track 9.1 at www.wiley.com/go/carrphonetics and draw metrical trees for the following phrases, as heard on the recording:

(a) sacked a worker

(b) delighted agents

(c) very pretty

(d) Piccadilly Circus (Show the tree for this last phrase both before and after rhythm reversal has applied)

3 Listen to Track 9.2 and explain the metrical structure of the following expressions, as heard on the recording:

(a) an ill-advised decision

(b) a well-decorated bedroom

(c) West Hampton Wanderers

(d) a broken-hearted man

(e) fifteen dollars

(f) fifty pounds

(g) dark-green trousers

(h) champagne cocktail

Listen to sound files online

4 Further phonetic transcription practice

Transcribe, with as much phonetic detail as possible, the following words as they are uttered on Track 9.3, indicating syllable boundaries, primary stress and (where applicable) secondary stress:

(a) rudimentary
(b) unfriendliness
(c) deconstructible
(d) opportunity
(e) Chinese

10

English Intonation

10.1 Tonic Syllables, Tones and Intonation Phrases

Although we often say that some speakers speak in a monotonous manner, the fact is that human beings do not utter speech which is monotone in nature: we all inflect our speech, creating intonational contours. But what is intonation, exactly? It is the use of pitch variation in discourse. What is pitch? We have seen that pitch is the auditory impression created by variations in the rate of vibration of the vocal folds. Intonation is the use of pitch contours over stretches of speech which often consist of more than one word. An example is the utterance . There are three syllables with primary word stress in this utterance: the penultimate syllable of , the single syllable of and the penultimate syllable of . But there is additional pitch movement on the primary-stressed syllable of . That stressed syllable is perceptually more prominent than the others, and will tend to be longer in duration, and louder, than the other stressed syllables in the utterance. That syllable is said to be the **tonic syllable**. The word ‘tonic’ denotes the fact that this syllable is where the **tone** falls. The tone is the extra pitch movement placed on that syllable. In our example, the tone is a **falling** tone: the rate of vibration of the vocal folds decreases as the syllable is uttered, resulting in a transition from a higher to a lower pitch. We will represent these as follows:

Listen to sound files online

(1) ˈMary ˈwent to the ↘doctor. (Track 10.1 at www.wiley.com/go/carrphonetics)

As we saw in chapter 8 on word stress, the diacritics on ˈMary and ˈwent indicate that the following syllables are stressed. The underlining on ↘___ indicates that it is the tonic syllable (and thus, by definition, stressed), and the preceding '↘' diacritic indicates a falling tone. This kind of tone is typical of declarative utterances, in which the speaker is making a statement, as opposed to, say, posing a yes/no question. Other tones are possible in English. In yes/no questions (questions which may solicit the responses 'Yes' or 'No'), it is common to find a **rising** tone in the tonic syllable, as in the question ? We will represent rising tones as follows:

(2) Is ˈMary ↗pregnant? (Track 10.2)

Here, the penultimate syllables of and are stressed, the tonic falls on the stressed syllable of , and the tone is a rise.

A third tone is the **rise-fall** tone, in which the pitch rises and then falls, as in the following exchange:

(3)
Wife:
Have you been ˈseeing ↗Mary?

Husband:
↗↘No! (Track 10.3)

The use of rise-fall tones conveys certainty, exclamation, strong conviction or strength of feeling on the part of the speaker. In this case, the husband is saying that he has certainly been seeing Mary: the intonation conveys a complete denial of the implied accusation.

A fourth tone is the **fall-rise** tone, as in the following exchange:

(4)
Wife:
Have you been ˈseeing ↗Mary?

Husband:
↗↘No! (Track 10.4)

Here, the pitch falls then rises in the second utterance. Use of such a tone conveys hesitation, lack of certainty, prevarication or reservation on the part of the speaker. In this example, the husband is being less than clear and straightforward in his response: he is denying that he's been seeing Mary, or trying to suggest that what he's been doing does not really amount to 'seeing Mary' in the romantic sense.

A stretch of discourse which contains a tonic syllable is called an **intonation phrase** (IP), otherwise known as an intonation group, intonation unit or tone group. These are also referred to as breath groups, since they constitute units of speech in which we expel air from the lungs. When we stop speaking to draw breath, we often do so at the end of a tone group: it is common for speakers to pause at the end of such units.

It is common to identify three main features of intonation:

- the chunking up of stretches of speech into IPs,
- the placing of the tonic on one of the stressed syllables of that chunk and
- the assignment of a specific tone on the tonic syllable

These are, to some extent, independent variables: what syllable we choose to put the tonic on can be independent of where the IP boundaries go, and what tone we place on the tonic syllable can be independent of where we choose to place the tonic.

In examples (1) to (4), the tonic falls on what is known as the **last lexical item** (LLI). Recall from our discussion of word stress that words can be classified into two broad groupings: words of a lexical category (typically nouns, verbs, adjectives and adverbs) and words of a functional, or grammatical, category (such as articles, conjunctions, prepositions and pronouns). The last lexical item in a syntactic unit is thus the last noun, verb, adjective or adverb. For example, in (1), the LLI is the noun , and in (2) it is the adjective . The following examples contain, in (5), an LLI which is a verb, and in (6), an LLI which is an adverb:

(5) My ˈhusband ↘cheats. (Track 10.5)

(6) His ˈlover ˈwalks ↘gracefully. (Track 10.6)

The LLI may not be the last item in an IP, as in:

(7) ˈBill ↘gave it to her. (Track 10.7)

Here, the last item is a pronoun, which is not a lexical item, and thus does not take the tonic. The second-last item is a preposition, and thus also fails to take the tonic. The third-last item is a pronoun, so that too fails to take the tonic. The tonic falls on , since it is the LLI, but not the last item. In cases like this, any syllables which follow the tonic syllable are said to constitute the **tail** of the IP: after the fall here, the pitch just trails off at a low level into the remaining syllables after the tonic.

10.2 Departures from the LLI Rule

The LLI rule is the default rule for the placement of the tonic. By ‘default’, we mean the point where the tonic is placed if no special circumstances prevail. Defaults in linguistics are rather like the default settings on your computer: they are the settings that are used unless one deliberately changes the set-up for some special purpose. It is common in English to shift the tonic away from the default position, for various purposes. We now examine some of those.

10.2.1 Contrastive Intonation

(8)

Speaker A:

ˈMary ˈgave ˈJohn a ↘camera.

Speaker B:

No, didn't give it to ; gave it to . (Track 10.8) (= ↘No | ↘she ˈdidn't ˈgive it to ↘him | ↘he ˈgave it to ↘her)

Here, the italics show that, in addition to , four pronouns receive the tonics (the vertical lines indicate the boundary between the IPs): it wasn't that gave a camera: it was that gave a camera. The referent of the word here is given, once speaker A has spoken. Consider the following possible intonational patterns for the sentence :

(9) ˈJohn is ˈtaking the ˈtrain to ↘London. (Track 10.9)

(10) ˈJohn is ˈtaking the ↘train to ˈLondon. (Track 10.10)

(11) ↘John is ˈtaking the ˈtrain to ˈLondon. (Track 10.11)

In (9), we have the default pattern for tonic placement, with the tonic on the stressed syllable of the LLI. In (10), we have contrastive intonation: the train is being contrasted with some other mode of transport, such as the plane. In (11), the speaker is stressing the fact that it is John, not someone else, who is taking the train to London.

This use of tonic placement relates to what is called **focus**. In (9), we have **broad focus**, associated with statements in which everything is news. These are statements which are said to come 'out of the blue': all of the information is announced as new information, so everything in the utterance is brought into focus.

In (10), the person being addressed already knows that John is going to London: what is news is the information concerning his mode of transport. This is called **narrow focus**.

In (11), the addressee knows that someone is taking the train to London, and is being informed that the person in question is John. This too is a case of narrow focus. Narrow focus relates to the **given/new distinction**. Given information is shared (mutual) knowledge, known to both the speaker and the hearer. New information is not previously known to the speaker and the hearer.

The tonic can be moved onto almost any syllable for contrastive purposes, including affixes. Here is a statement by the British prime minister David Cameron in 2010:

(12) It's not ↘unemˈployment that will be creˈated. | It's ↘emˈployment.

Normally, and have primary stress on the penultimate syllable, as in ˈtrying to creˈate ↘____ . Here, the prime minister is contrasting employment with unemployment. We see from this example that, given a context in which we wish to highlight a given word in order to contrast it with another, the tonic may be placed on something other than the LLI (in the first IP), and that even affixes may receive the tonic (the prefix in).

10.2.2 Given Information

Another situation in which the LLI rule is flouted concerns the notions of given and new information. Consider the following exchange:

(13)
A:
We need tomatoes. (We ˈneed to↘matoes)

B:
We've tomatoes! (We've ↘ got ˈmatoes) (Track 10.12)

The word is given in the first statement: it has already been mentioned, so the information it conveys is given (shared by the participants in the exchange). The tonic is therefore shifted away from the LLI ( ) onto the word . Now consider the following:

(14) In ↘ most ˈcases, | we aˈpply the ↘rule, | but in ↘some cases, | we ↘don't. (Track 10.13)

Here, the LLI in the first IP ( ) is given by the context of utterance: if we utter (14), the person we are speaking to already knows what the rule is about, and what kinds of cases are being spoken about.

Synonyms can count as conveying given information:
(15)
A:
She's ˈborrowed ˈJane's ↘frock.

B:
↘No, | it's ↘Mary's ˈdress (Track 10.14)

Here, the word isn't given, but its meaning is, via the uttering of the synonym .

Presuppositions can be conveyed via tonic placement:
(16)
A:

Have you ˈspoken to ↗John?

B:
I don't ↘speak to ˈracists (Track 10.15)

B is presupposing that 'John is a racist' is given information. So the tonic shifts from the LLI ( ) to the preceding lexical item. Speaker A can the presupposition that John is a racist, even if that is open to question.

Notice here that the verb is given, but nonetheless takes the tonic: contrastive intonation can lead to the tonic being placed on given information. Here is another example:

(17)
A:
He's ˈgoing to ↘Paris.

B:
He's not ˈgoing to ↘Paris. | He's ˈgoing to ↘London. (Track 10.16)

Given information can be shared by millions of people (e.g. the fact that Barack Obama was elected president of the United States) or by as few as two people (e.g. husband and wife).

10.2.3 Final Temporal Adverbials

It is common to find that LLIs which are in syntactic units which have an adverbial function, and which convey information relating to time, fail to take the tonic, as in:

(18) ˈJohn's ˈgoing to ↘London on ˈSaturday. (Track 10.17)

Here, is the LLI, but since the prepositional phrase is a final temporal adverbial, the LLI within that adverbial expression fails to take the tonic. If we to place the tonic on , that would constitute a case of contrastive intonation:

(19) ˈJohn's ˈgoing to ˈLondon on ↘Saturday (Track 10.18) (as opposed to some other day of the week).

When final adverbial expressions are fronted, they tend to form a separate IP:

(20) On ↘Saturday, | ˈJohn's ˈgoing to ↘London. (Track 10.19)

10.2.4 'Event' Sentences

These are rather curious. They are short statements which contain intransitive verbs, but the tonic fails to fall on the intransitive verb rather than the LLI:

(21)

(a)

The ↘kettle's ˈboiling.

(b)

The ↘baby's ˈcrying.

(c)

Your ↘house is on ˈfire.

(d)
The ↘sun's ˈcome ˈout. (Track 10.20)

One would have expected the tonics to fall on the LLIs , and fire, and on the particle (see above on LLIs, and below on intransitive phrasal verbs, where we expect the tonic to fall on the particle).

It has been observed that the subjects in such sentences do not denote human agents, but why that should affect the tonic placement is far from clear. There seems to be pragmatic foregrounding (selecting out) of the subject in such cases.

10.2.5 Non-Lexical Items which Often Take the Tonic

The negative equivalents of the non-lexical items ( ) often take the tonic:

(22)
(a)
I ˈsaw ↘no one.

(b)
I've ˈdone ↘nothing.

(c)
We're ˈgetting ↘nowhere.

(d)
This ˈinterests ↘nobody. (Track 10.21)

One place the tonic on the non-lexical items and , but only if the intonation is contrastive, as in:

(23)
A:
I ˈsaw the ˈneighbour in the ↘pine grove this ˈmorning.

B:
You ↘couldn’t have. | He’s in ↘Paris ˈright ˈnow.

A:
Well I ˈsaw ↘someone. (Track 10.22)
(That is, not the neighbour, but some other person).

Non-native speakers should be aware that pro-forms such as and are not lexical items: they convey given information, and thus do not normally take the tonic, as in:

(24)

(a) A: I 'went 'looking for a 'bottle of ↘wine.
B: Did you ↗get one?
A: ↘Yes.
(b) 'Mary 'drank some ↘wine | and ↘Bill did so too. (Track 10.23)

10.2.6 Cleft Sentences

Cleft sentences take the following form:

(25) It’s that wear kilts.

It was that did it.

Clefting is a way of highlighting, or bringing into focus, a syntactic constituent. One could say:

(26) I ˈlove ↘John.

But in the cleft version, the contrast between John and anyone else is more emphasized:

(27) It's ↘John that I ˈlove. (Track 10.24)

Although is the LLI, the tonic falls on the highlighted item. Here, the material after the highlighted material forms the tail of the IP: the hearer knows that the speaker loves someone, so that knowledge is given, and no further tonic is required.

10.2.7 Deictic Expressions

The word **deictic** means 'involved in pointing out', either by literally pointing at something with one's finger while speaking, or bringing something to someone's attention without physically pointing. Deictic expressions in English include the demonstrative words and . These count as function words, so they do not take the tonic when the LLI rule applies, as in the following questions:

(28)

(a)

Could you ↗give me that?

(b)

Can I have ↗five of those? (Track 10.25)

In such utterances, it is clear from the context of utterance what and are being used to refer to (in the first example, a parent might be asking a child to hand her a knife; in the second example, a shopper might be asking a shopkeeper to give her five oranges).

If the thing being referred to is explicitly mentioned, the LLI rule will assign the tonic to the noun in question, as in:

(29)

(a)

Could you ˈgive me that ↗knife?

(b)

Can I have ˈfive of those ↗oranges? (Track 10.26)

The tonic can fall on deictic expressions when they are being used contrastively, as in:

(30)

(a)

Could you ˈgive me ↗that ˈknife?

(as opposed to some other knife)

(b)

Can I have ˈfive of ↗those ˈoranges? (Track 10.27)

(as opposed to some other varieties of orange)

10.3 IPs and Syntactic Units

10.3.1 Syntactic Units which Normally Form a Separate IP

There are syntactic units which normally form separate IPs, such as relatively short main clauses, as in (1) to (3) in section 10.1: here, we can see that the intonation tracks the syntax. This is unsurprising, since both clauses and IPs convey coherent chunks of information. We will now examine a range of other syntactic units which normally form independent IPs.

Parenthetical information is extra, optional information offered by the speaker. If parentheticals are omitted from a syntactic structure, the structure in question remains grammatically well-formed. Let us look at some types of parenthetical.

Non-restrictive relative clauses

(31) The ˈguys in the ↘car, | who were ↘hungry, | ˈate some ↘sandwiches. (Track 10.28)

Here, the IP boundaries correspond to the commas in the written form of the sentence. The tonics fall on the LLIs in each IP: and . The meaning conveyed is that all of the guys in the car were hungry (thus the expression ‘non-restrictive relative clause’: the range of referents is not restricted).

(Restrictive relative clauses

Note that these do not normally count as parentheticals, and thus do not normally form a separate IP, as in:

(32) The ˈguys in the ˈcar who were ↘hungry | ˈate some

↘sandwiches. (Track 10.28)

In the second example, there are only two IPs, as opposed to three in the first. The subject noun phrase is sufficiently long to form a separate IP, but the restrictive relative clause within that noun phrase does not form a separate IP. The difference in meaning between (31) and (32) is that, in (32) it is not necessarily the case that of the guys in the car were hungry: the meaning is restricted only to the guys in the car.)

Noun phrases in apposition

Noun phrases are said to be in apposition when they are co-referential, that is, when they are being used to refer to the same person or entity, as in:

(33) ˈBarak O↘bama, | a ˈDemocrat poli↘tician, | is in↘telligent. (Track 10.29)

Other parentheticals

(34) ↘↗Mary, | you’re not ˈgoing to be↘↗lieve this, | but ˈJane is ↗↘pregnant! (Track 10.30)

Note that parentheticals are uttered on a lower pitch range than the preceding and following IPs: if you listen carefully to Tracks 10.28, 10.29 and 10.30, you should be able to hear this.

(35)

(a)

ˈMary ˈmoved to ↘Paris | but ˈJohn ˈstayed in ↘London.

(Sentence co-ordination)

(b)
ˈJohn ˈwent to the ↘pub | and ˈordered a ↘ beer. (Verb phrase co-ordination)

(c)
She's ˈvery ↘ tall | and ˈvery ↘ pretty. (Adjective phrase co-ordination)

(d)
His ˈvery ↘well | and ˈvery ↘quickly. (Adverb phrase co-ordination)

(e)
It's ˈeither ˈin the ↘fridge | or ˈon the ↘table. (Prepositional phrase co-ordination)

(f)
He ˈbought the ˈhouse on the ↘hill | and the ˈwoods in the ↘valley. (Noun phrase co-ordination) (Track 10.31)

However, when the constituents are short, separate IPs are not always required:

(36)
(a)
She's ˈtall and ↘lanky. (Adjective phrase co-ordination)

(b)
He ˈstopped and ↘stared. (Verb phrase co-ordination)

(c)
He ˈbought ˈmilk and ↘cheese. (Noun phrase co-ordination)

(Track 10.32)

This is especially noticeable when co-ordinated elements have been lexicalized (formed into lexical items), taken to be used to refer to single entities, such as:

(37)

(a)

ˈfish'n' ↘chips

(b)

ˈbeer and ↘skittles

(c)

ˈstrawberries and ↘cream (Track 10.33)

Other co-ordinated items that are used to refer to what is perceived as a single entity or unit are British pub names:

(38)

(a)

The ˈDog and ↘Duck

(b)

The ˈFox and ↘Hounds

and names of couples:

(c)

ˈBill and ↘Mary ˈJane and ↘Clive (Track 10.34)

(The couples here are considered to be 'an item'.)

Normally, each item on the list constitutes a separate tone group

(39) He bought ↗eggs, | ↗milk, | to↗matoes | and ↘ham. (Track 10.35)

Non-final items on a list often take a rising tone: this signals that the list is not yet complete.

When a sentence contains a subordinate clause, the clause boundary often corresponds to an IP boundary:

(40)

(a)

I'll ˈbuy the ˈfish'n' ↘chips | when I ˈgo to the ↘shops.

(b)

I ˈtold the ˈnew reˈcruit to the ↘company | that he was ↘fired. (Track 10.36)

If the material preceding the subordinate clause is relatively short, the subordinate clause need not form a separate IP:

(41) I ˈthink she's been ↘sacked. (Track 10.37)

Adverb phrases necessarily have an adverbial function. But other phrases, notably prepositional phrases, can have an adverbial function.

It is common to distinguish verb phrase adverbials from sentence adverbials, as in (42a) and (b) respectively:

(42)

(a)

John went to the interview hopefully.

(b)
John went to the interview, hopefully.

In (42a), John was hopeful. In (42b), the speaker is hopeful.

Notice that (42b) can be rephrased with the sentence adverbial in initial position:

(c)
Hopefully, John went to the interview.

In either case, sentence adverbials form a separate IP, and they have the fall-rise tone:

(43)
(a)
↘↗Hopefully, | ˈJohn ˈwent to the ↘interview.

(b)
ˈJohn ˈwent to the ↘interview, | ↘↗hopefully. (Track 10.38)

These take the syntactic form   It is common for these to form two separate IPs:

(44) What he ↘↗needs | is a ↘bath. (Track 10.39)

(45)
(a)
The ˈthing ↘↗is | is that she’s ↘pregnant.

(b)
The alˈternative ↘↗is | is that we'll have to beˈcome more in ˈvolved in ↘Europe.

(c)
My conˈcern ↘↗is | is that it's ˈgot ˈtoo ↘big | too ↘quickly.

(d)
The ˈfact of the ˈmatter ↘↗is | is that the ˈway it is ↘run | is too ↘complex.

(e)
The ˈgood ˈnews ↘↗is | is that they're ˈlending to ˈsmall ↘businesses. (Track 10.40)

This construction is very widespread in spoken English, both informal and formal (most of these examples are taken from formal interviews with British politicians on BBC TV). Although it is possible not to have an IP boundary after the first   , that is the norm.

10.3.2 Syntactic Units which Do Normally Form Separate IPs

These abound in novels, but also in everyday speech

(46) I'm ↘tired, he said. (Track 10.41)

Here, the falling tone on the tonic syllable keeps trailing off into the reporting clause ( )

It is possible, however, to form a separate IP around a reporting clause. Compare (47a) and (b):

(47)

(a)

He’s re↘<u>ti</u>red, I think.

(b)

He’s re↘<u>ti</u>red, | I ↘<u>th</u>ink. (Track 10.42)

Utterance (47b) conveys less certainty than (47a).

(48) The ˈold ˈman ˈkicked the ↘<u>dog</u>. (Track 10.43)

But, as we saw in earlier, the longer a subject noun phrase, the more likely it is that a separate IP will be possible.

As we have seen, these do not normally form a separate IP, as in example (32), repeated here (the long subject NP forms a separate IP, but not the relative clause):

(32) The ˈguys in the ˈcar who were ↘<u>hun</u>gry | ˈate some ↘<u>sand</u>wiches.

10.3.3 Syntactic Units which May, or May Not, Form Separate IPs, Depending on the Sense Conveyed

Reverse polarity tag questions

By **reverse polarity** we mean that the first part is in the positive, and the tag question in the negative, or vice versa, as in:

(49)

(a)

You’re going to do this, aren’t you?

(b)

You’re not going to do this, are you? (Track 10.44)

The intonation of reverse polarity tag questions works as follows: if we form a separate IP on the tag question, and place a falling tone on it, the tag invites agreement:

(50)

(a)

You’re ˈgoing to ↘do this, | ↘aren’t you?

(b)

You’re not ˈgoing to ↘do this, | ↘are you? (Track 10.44)

If there is a rising tone on the tag, it need not form a separate IP:

(51)

(a)
She’s ˈcoming to the ↗party, isn’t she?

(b)
She’s not ˈgetting ↗married, is she? (Track 10.45)

Here, the tag question forms part of the tail of the IP. But the tag question form a separate IP:

(52)
(a)
She’s ˈcoming to the ↗party, | ↗isn’t she?

(b)
She’s not ˈgetting ↗married, | ↗is she? (Track 10.46)

The differences in conveyed meaning are subtle: in (51a), the speaker is not entirely certain whether she’s coming to the party. In (51b), the speaker may be expressing surprise, or even astonishment, whereas in (52a) and (52b), the speaker is a lot less sure, and is posing more of a query than in (51a) and (51b).

10.4 Tonic Placement, IP Boundaries and Syntax

Phrasal verbs in English have two parts: the first part, which looks like a normal verb, and the second part, which looks like a preposition, and is often called a particle. They can be transitive (which means that they are followed by a direct object, as in , where is the direct object) or intransitive (which means that no direct object is required, as in). Learners of English as a foreign language are well-advised to learn the intonation of such verbs, since there are so many of them, and they occur with high frequency in spoken English.

If the direct object noun phrase is phrasal, the tonic falls on the head noun in the noun phrase (the head noun is the noun in a noun phrase which is semantically the most prominent):

(53)

(a)

He ˈchatted ˈup the ↘waitress.

(b)

He ˈchatted the ↘waitress ˈup.

If the direct object noun phrase is a pronoun, the tonic falls on the particle (which is normally shifted so that it follows the direct object):

(c) He ˈchatted her ↘up. (Track 10.47)

These take the tonic on the particle:

(54) He ˈbacked ↘down. (Track 10.48)

However, many short sentences with intransitive phrasal verbs are ‘event sentences’ (see 10.2.4 above), in which case the tonic is retracted:

(55)

(a)

The ↘plane ˈblew up.

(b)

The ↘car ˈbroke down. (Track 10.49)

10.4.2 Degree Adverbials

The most central example of a degree adverbial is the word . It functions to modify adjectives in adjective phrases, and adverbs in adverb phrases, as in:

(56)

(a)

He’s very ↘tall. (Adjective phrase)

(b)

He ˈtalks very ↘slowly. (Adverb phrase) (Track 10.50)

Other degree adverbials include , , the mild swearword and the stronger ‘f-word’, as in:

(57)

(a)

He’s so ↘stupid.

(b)

He’s inˈcredibly ↘arrogant.

(c)
He’s ˈbloody ↘good.

(d)
He’s ˈf*cking ↘good. (Track 10.50)

But the tonic can be placed on the degree adverbial, for emphasis, as in:

(58)
(a)
He’s ↗↘so ˈstupid

(b)
He’s in↗↘credibly ˈarrogant!

(c)
He’s ↗↘bloody ˈ

(d)
He’s ↗↘f*cking ˈgood! (Track 10.51)

There is a use of in colloquial English which acts as a verb phrase adverbial, and takes the tonic, with an exclamatory tone, as in:

(59)

(a) A (to C): Has 'anyone ever ↘told you | that you've 'wasted your ↗life?

B: You ↗↘so 'haven't!

(b) A: Why do 'men 'dress up as 'women at 'fancy ↘dress 'parties?

B: They ↗↘so do!

(c) I'm ↗↘so not 'shining 'shoes! (The speaker refuses to shine shoes)

(d) 'You two are ↗↘so 'going 'out with the 'wrong 'men.

(e) That's ↗↘so not 'cool, 'Carol. (Track 10.52)

10.5 Tones and Syntax

10.5.1 WH Questions

These normally have a falling tone:

(60) Where are you ↘going? (Track 10.53)

But not when used echoically:

(61)

A:

I'm ˈmoving to ↘London.

B:

↗Where are you moving to? (Track 10.54)

Speaker B here has either not properly heard what A said, or is expressing incredulity. B's WH question is said to be echoic in that it echoes part or all of what A has just said. Notice that the tone keeps on trailing upwards in the tail of the IP, just as it trails downwards when a tonic syllable before the tail has a falling tone.

10.5.2 Declaratives as Questions

We can use syntax to form yes/no question, in which the first auxiliary verb in the main clause is inverted around the subject noun phrase, as in:

(62) (a) Have they ˈfound your ↗mobile?

The corresponding declarative statement would have a falling tone:

(b) They've ˈfound your ↘mobile.

But we can retain the declarative syntactic structure and still ask a question by placing a rising, rather than a falling, tone on the LLI:

(c) They've ˈfound your ↗mobile? (Track 10.55)

Declarative structures can be uttered with a fall-rise:

(63)

A: They've 'found my ↗↘mobile!

B: They've 'found your ↘↗mobile? (Track 10.55)

In using a fall-rise, B is expressing surprise that the mobile has been found, or that it is the , as opposed to some other lost object, that has been found.

10.6 Tonic Placement and Discourse Context

10.6.1 Vocatives

Vocative expressions are used for addressing one's interlocutor, as in the following extracts from telephone messages left by a speaker calling a friend called Nick:

(64)

(a)

Nick, it's me.

(b)

It's me, Nick.

Here the speaker assumes that the addressee (Nick) can identify the voice of the person calling, or that Nick is expecting a call from that person.

vocatives form a separate IP:

(c) ↘Nick, | it's ↘me.

vocatives do not:

(d) It's ↘me, Nick. (Track 10.56)

If an IP is formed around the final , as in (65):

(65) It's ↘me, | ↘Nick. (Track 10.57)

then the word is not interpreted as a vocative; rather, it is interpreted as the name of the caller.

10.6.2 Other Meaning Differences Conveyed by IP Boundaries

As we've seen, the placement of IP boundaries, and/or the kind of tone we select, can convey differences in meaning. Consider:

(66)

(a)

He didn't marry her because she was French.

(= He ˈdidn't ˈmarry her because she was ↘French)

This means that he marry her, but not because she was French.

(b)

He didn't marry her, because she was French. (Track 10.58)

(= He ˈdidn't ↘marry her | because she was ↘French)

This means that he marry her, the reason being that she was French.

10.7 Summing Up

We have seen that there are three main structural aspects of English intonation: the dividing up of utterances into intonational phrases which are chunks of information, the placing of a tonic on one of the stressed syllables in each chunk, and the kind of tone we use in that tonic syllable. Intonation in English, we have seen, is connected to syntactic structure, the lexical vs functional distinction, the meaning expressed by the syntactic units in question, and aspects of discourse linked to the context of utterance and phenomena such as conveyed meaning and the speaker's attitude towards what she or he is saying. Perhaps the most striking aspect of

English intonation is the extent to which it is dynamic, in the sense that speakers of English frequently move the tonic away from the default LLI position, for a wide variety of purposes. For non-native speakers of English, some degree of mastery of this will result in a much more native-like speech style.

Exercises

1 In each of the following utterances, identify the last lexical item.

(a) John went to the pub.

(b) Mary put her finger on it.

(c) My father says he can’t understand that.

(d) He talks rather slowly.

(e) I want that pink one.

2 Which of the following questions can have a rising tone?

(a) Is Bush mad?

(b) What do you want?

(c) Have you eaten?

(d) How does this work?

(e) Isn’t it time for lunch?

3 Listen to Track 10.59. Where do the tonics fall in the following utterances in that sound file? Explain why.

(a) She chatted up the waiter.

(b) She chatted him up.

(c) She broke down.

(d) We’ve split up.

(e) I've put him off.

4 Listen to Track 10.60. Identify the IP boundaries in the following utterances in that sound file and say where the tonics fall. Explain why.

(a) Mary, you're fired.

(b) You're fired, Mary.

(c) He's mad, she said.

(d) It's an evil empire, said the president of the United States.

5 Listen to Track 10.61. Identify the IP boundaries in the following utterances in that file and say where the tonics fall. Explain why.

(a) Mary, a good friend of mine, is pregnant.

(b) The guys in the car, who were hungry, ate some sandwiches.

(c) The guys in the car who were hungry ate some sandwiches.

(d) Bill, you won't believe this, you've passed your exam.

(e) His new book, , is sure to be a bestseller.

6 What is the default tonic placement in the following utterances?

(a) He went to London on Thursday.

(b) I haven't seen her recently.

(c) He left for Paris in a hurry.

(d) She left her bedroom in a mess.

(e) He speaks quickly.

7 What are the intonation group possibilities in the following utterances? Discuss different tone possibilities in the tag questions. Listen to Track 10.62 and describe the intonational structure you hear, including IP boundaries and tones.

(a) You're not pregnant, are you?

(b) You like lasagne, don't you?

(c) You   play golf, don't you?

(d) We can sort this out, can't we?

(e) We'll never sort this out, will we?

8 Where do the tone group boundaries fall in the following utterances, and where do the tonics fall? If there are alternative intonational structures, say what you think they are. Listen to Track 10.63 and describe the intonational structure you hear, including IP boundaries and tones.

(a) 'You can't go', said Bill Smith, a good friend of mine.

(b) 'Obama can't win in Texas', claims Hillary Clinton, a woman whose husband Bill, ex-president, is from the South of the USA.

(c) 'Is Amy Winehouse in rehab?', asked Jonathon Ross on the Thursday after she sang out of tune at a concert in London.

(d) 'Dickens I can't stand', confessed the young recruit to a university lectureship in Victorian literature.

(e) 'What does George Bush, a devout Christian, have to say about the treatment of prisoners in Abu Ghraib?', asked the chair of the committee.

11

Graphophonemics: Spelling–Pronunciation Relations

11.1 Introduction

The relationship between spelling and pronunciation in English is complex: it can seem completely arbitrary. The complexities have historical sources: they result from changes introduced by scribes after the Norman Conquest, the adopting of many loanwords with foreign spellings, and, above all, changes in the phonological system of English as the language evolved from Old English, through Middle English, into Early Modern English and present-day English. Despite the complexities, there are some basic regularities which are worth learning, especially for non-native speakers of English.

We begin by distinguishing between letters and graphemes. There are twenty-six letters in the Roman alphabet, but there are more than twenty-six visual symbols (**graphemes**) for representing English phonemes and allophones, since combinations of letters can be used to represent a given phoneme or allophone. Examples are <ph>, which corresponds to the /f/ phoneme, as in the word , <th>, which corresponds to both the /θ/ and /ð/ phonemes, as in the words and , and <oa>, which corresponds to the RP and GA phonemes /oʊ/ and the SSE (Standard Scottish English) phoneme /o/, as in the word . We will refer to such graphemes as **digraphs**, since they contain two letters. We will refer to graphemes with three letters, such as <sch>, in words such as (‘overly sentimental’), as **trigraphs**. The distinction between letters and graphemes can be seen in the different writing conventions of English and French. For instance, in writing the initial for my first name in English, the convention is to use the first : ‘P. Carr’. The French convention is to select the first : ‘Ph. Carr’. The term **graphophonemics** is the name given to the study of the relationship between graphemes and phonemes (and some of their allophones). We will begin by examining vowel graphemes in English, and then proceed to consonant graphemes.

11.2 Vowel Graphemes and Their Phonemic Values

11.2.1 Vowel Monographs

Let us begin with the five vowel graphemes <a>, <e>, <i>, <o>, <u>. We will distinguish between two different phonemic values for these graphemes: their checked values and their free values. The terms **checked** and **free** derive, historically, from facts about syllable structure. Take the word : in Middle English, this word was bisyllabic: /bi:tə/, in which the first syllable is /bi:/, and the second syllable is /tə/. These are both open syllables: they contain no coda consonants. What we are referring to as free values derive, historically, from open syllables. The word , in Middle English, was pronounced /bit/, a monosyllabic word with a closed syllable, containing the coda consonant /t/. What we are referring to as checked values derive from closed syllables: a closed syllable is a checked syllable. Notice that, in contemporary English, the word has lost the final schwa, and is now monosyllabic: /baɪt/. But it retains a free value: /aɪ/, the historical descendant of the long /i:/ vowel. The word retains a checked value: /ɪ/, the historical descendant of the short /i/ vowel.

In the kinds of monosyllabic words we will consider here, four of these graphemes have a checked and a free value, and <u> has two checked and free values, as follows (examples from RP):

(1)

	Free value	Example	Checked value	Example
<a>	/eɪ/	made	/æ/	mad
<e>	/i:/	Pete	/ɛ/	pet
<i>	/aɪ/	hide	/ɪ/	hid
/o/	/əʊ/	note	/ɒ/	not
/u/	/ju:/ or /u:/	cute	/ʌ/ or /ʊ/	cut

While the final <e> remains in the spelling of the words in the first column of examples, the schwa which it once corresponded to was elided over time, as we have noted. We will refer to this grapheme as **mute e**. In stressed monosyllabic words, if there is a mute e, the preceding vowel grapheme corresponds to the free value. Note that the names of these graphemes have the free value: we call them /eɪ/, /i:/, /aɪ/, /oʊ/, /ju:/.

In a stressed monosyllabic word without a mute e, the preceding vowel grapheme corresponds to the checked value. This is true when there is more than one consonant grapheme at the end of the written word, as in:

(2)

apt, act

bell, text

Bill, width

accost, knots

butt, tuft

These free/checked values figure in alternations in pronunciation when certain affixes are added to bisyllabic words, as in the following examples, from RP:

(3)

<a>	sane	/seɪn/	sanity	/ˈsænɪti/
<e>	obscene	/əbˈsi:n/	obscenity	/əbˈsɛnɪti/
<i>	divine	/dɪˈvaɪn/	divinity	/dɪˈvɪnɪti/
<o>	verbose	/vɜ:ˈbəʊs/	verbosity	/vɜ:ˈbɒsɪti/
<u>	consume	/kənˈsju:m/	consumption	/kənˈsʌmpʃən/

In stressed monosyllabic words, there are additional values for these vowel graphemes in non-rhotic accents such as RP, where the historical loss of [ɹ] in coda position has resulted in changes to the preceding vowel:

(4) ‘Pre-r’ checked and free values in RP

	Free value	Example	Checked value	Example
<a>	/ɛə/ or /ɛː/	mare	/ɑː/	bard
<e>	/ɪə/	mere	/ɜː/	perk
<i>	/aɪə/ or [aː]	fire	/ɜː/	bird
<o>	/ɔː/	bore	/ɔː/	stork
<u>	/(j)ʊə/ or /(j)ɔː/	sure	/ɜː/	curt

In contemporary RP, the /(j)ɔː/ pronunciation of words like seems to be slowly replacing the previous /(j)ʊə/ pronunciation, so that and are homophones. Many words like are now pronounced /ɛː/, rather than /ɛə/. Monophthongal versions of /aɪə/ have a long [aː] in RP: there is variability here among RP speakers.

As we have seen, GA is rhotic, so the free and checked values in 'pre-r' position differ somewhat from the RP values (we include the /ɹ/ for clarity):

(5) 'Pre-r' checked and free values in GA

	Free value	Example	Checked value	Example
<a>	/ɛɹ/	mare	/ɑ/	bard
<e>	/iːɹ/	mere	/ɜɹ/	perk
<i>	/aɪɹ/	hire	/ɜɹ/	bird
<o>	/oɹ/	bore	/ɔ/	stork
<u>	/ʊɹ/	sure	/ɜɹ/	curt

The same pattern of checked and free values can be found in words with more than one syllable with the final syllastressed (examples from RP):

(6)

	Free value	Example	Checked value	Example
<a>	/eɪ/	ˌlemoˈnade	/æ/	forˈbad
<e>	/iː/	reˈplete	/ɛ/	reˈpent
<i>	/aɪ/	deˈride	/ɪ/	forˈbid
/o/	/əʊ/	deˈnote	/ɒ/	forˈgot
/u/	/juː/ or /uː/	deˈnude	/ʌ/ or /ʊ/	unˈcut

The same is true for 'pre-r' vowel graphemes in words with more than one syllable, stressed on the final syllable :

(7)

	Free value	Example	Checked value	Example
<a>	/ɛə/ or /ɛ:/	de'clare	/ɑ:/	re'tard
<e>	/ɪə/	ˌinter'fere	/ɜ:/	a'ssert
<i>	/aɪə/	re'tire	/ɜ:/	un'gird
<o>	/ɔ:/	dep'lore	/ɔ:/	re'tort
<u>	/(j)ʊə/ or /(j)ɔ:/	de'mure	/ɜ:/	curt

We have seen several examples of words with a stressed final syllable and a single consonant grapheme with no mute e, such as ': the vowel grapheme in the stressed syllable corresponds to the checked value of the grapheme. If we add inflectional suffixes to such words, and thus add a syllable, we must double the consonant grapheme to convey the checked value: . This is one sense in which the spelling of English words can actually be helpful as a guide to pronunciation: without the doubled consonant grapheme, the stressed vowel grapheme would have the free value, as in the following pairs:

(8)

	Free value	Example	Checked value	Example
<a>	/eɪ/	'tamer	/æ/	'tanner
<e>	/i:/	'meted	/ɛ/	'petting
<i>	/aɪ/	'diner	/ɪ/	'dinner
<o>	/əʊ/	'noter	/ɒ/	'hotter
<u>	/ju:/ or /u:/	'cuter	/ʌ/ or /ʊ/	'cutter

The words above are morphologically complex: they contain more than one morpheme. But even in words which are morphologically simple and have a prefinal stressed syllable, we find the same pattern. The word , for instance, does not contain a suffix, but the double consonant grapheme encodes the checked value for <a>, namely /æ/. The word is also morphologically simple, but the single consonant grapheme encodes the free (long/tense) value for <a>, namely /eɪ/ (in RP and GA). Unfortunately, there are irregular spellings. Take the word , which has the checked /æ/ value: we would expect it to be pronounced with /eɪ/, or to be written , just like and . These irregular spellings do not help the foreign learner, who often learns the spelling of a word before learning the pronunciation. But the regularities we have just described nonetheless cover a huge number of English words, and are therefore worth learning about.

A further regularity concerns stressed vowels in antepenultimate syllables (or earlier in the word than that) when they are followed by a single consonant grapheme: these typically have the checked value, as in the following words (RP pronunciations):

(9)

<a>	character	/ˈkæɹəktə/	family	/ˈfæmɪli/
<e>	enemy	/ˈɛnəmi/	federal	/ˈfɛdəɹəl/
<i>	cinema	/ˈsɪnəmə/	pitiful	/ˈpɪtɪfʊl/
<o>	moribund	/ˈmɒɹɪbʌnd/	positive	/ˈpɒzɪtɪv/

The <u> grapheme behaves differently: a stressed vowel in penultimate position (or earlier), when followed by a single consonant grapheme, has the free value for <u>, as in:

(10)

frugal

/ˈfɹu:gəl/

imprudent
/ɪmˈpɹuːdənt/

accumulate
/əˈkjuːmjəleɪt/

In addition to the <a>, <e>, <i>, <o>, <u> monographs, the <y> grapheme can represent vowels, as follows:

(11)

Grapheme	Phonemes/Allophones	Examples
<y>	/ɪ/, /aɪ/, [i]	myth, rhyme, happy

In stressed syllables, <y> has the same values as <i>. The checked value /ɪ/ occurs where we would expect it to: words like , consisting of a single stressed syllable, do not have a mute e, whereas words such as have a mute e, and thus have the free value /aɪ/. We have seen that, in 'pre-r' position, when followed by a mute e, RP <i> has /aɪə/, as in , which can be uttered as long [a:]. The same is true for <y> in this position: corresponds to /taɪə/, which can also be pronounced [ta:].

For word-final unstressed <y>, we have treated the short [i] vowel here as a positional variant (allophone) of the /i:/ phoneme in RP and GA, where it occurs in words such as , and in the many adverbs ending in <ly>, such as . Stressed word-final <y> corresponds to /aɪ/, as in and .

11.2.2 Vowel Digraphs

The vowel digraphs present problems for foreign learners, since they often correspond to more than one phonemic value, for reasons connected to the history of English. This is especially true for RP, since it is non-rhotic, which complicates matters. Let us list some of the most frequent vowel digraphs and discuss their phonemic values.

In SSE, <ai> corresponds to the phoneme /e/, as in (/pen/) and (/feɹ/). But, in RP, since it is non-rhotic, there is a separate 'pre-r' value: is /peɪn/, while is /fɛə/ or /fɛ:/. In this respect, SSE is less complex than RP for many learners of English. In GA, there is a complication: is /peɪn/, as in RP, but 'pre-r' words such as , while having the /eɪ/ phoneme, undergo the 'Marry Merry Mary' neutralization, in which the /æ/ vs /ɛ/ vs /eɪ/ oppositions are neutralized to [ɛ] in 'pre-r' position: is pronounced as [ˈmɛɹi] [ˈmɛɹi] [ˈmɛɹi], and is pronounced [fɛɹ]. This can cause confusion for speakers of varieties of British English when listening to Americans. For instance, when my son, while attending primary school in the United States for a semester, told me the name of a new American schoolfriend, I thought that the boy was called , since my son pronounced the name the way a GA speaker would: [fɛɹɫ]. In fact, the boy's name was ': schwa is often elided after /ɹ/ in GA, and the 'Marry Merry Mary' rule applies to the stressed vowel. As the playwright George Bernard Shaw once said, Britain and America are two countries divided by a common language!

In RP, this can correspond either to /ɔː/, as in , or to the LOT vowel /ɒ/, as in . Many speakers of GA lack the /ɒ/ phoneme, and have /ɑ/ instead, so that Austria has the value /ɑ/ in its stressed syllable (though there is variability here).

In SSE, this corresponds to the /i/ phoneme, as in (/si/) and (/piɹ/). In RP, words like have the /iː/ phoneme, while, in 'pre-r' position, there is a centring diphthong, as in (/piə/). In GA, <ee> corresponds to /iː/, but there is often neutralization of the opposition between /iː/ and /ɪ/ in 'pre-r' position, so that and can be homophonous. When this opposition is neutralized, the resulting vowel can sound like [ɪ] or [i]: [stɪɹɪt] or [stiɹɹɪt].

For historical reasons, this digraph has three main phonemic values in RP and GA: /uː/, as in , /ʊ/, as in , and /ʌ/, as in . In 'pre-r' position in RP, the centring diphthong /ʊə/ occurs, as in , but this is increasingly giving way to the long monophthong /ɔː/, so that and are homophones. In GA, the FOOT vowel can be unrounded, and the symbol often used to represent that resulting high, back unrounded vowel is [ɯ], as in : [ʃɯɹ].

This digraph can correspond to either /aʊ/, as in , or to /əʊ/ (GA /oʊ/), as in . This poses problems for learners of English, since there is no rule governing the occurrence of one or the other. Worse still, there are forms such as , in which the verb has /aʊ/, while the noun has /əʊ/.

In <ought> sequences, RP has /ɔ:/, as in , while many GA speakers have /ɒ/ in such words.

Word-final <ough> is complex: it can correspond to /u:/, as in , /əʊ/ (GA /oʊ/), as in , /ʌf/, as in , /ɒf/, as in , and /aʊ/, as in .

Finally, <ou> in words with <ouble> and <ouple> corresponds to /ʌ/, as in and .

This is one of the most difficult of the vowel digraphs for students learning English as a foreign language. It can correspond to /i:/ in RP (and GA), as in the word , or /ɛ/, as in . There are several 'pre-r' values in RP: /ɪə/, as in , /ɛə/ (or, more often in contemporary RP, /ɛ:/), as in , /ɜ:/, as in , and /ɒ:/, as in .

In GA, the 'pre-r' values include /ɜ/, as in (/dɜɹθ/), and /ɑ/, as in . GA also has phonetic [ɛ], as in , because of the ‘Marry Merry Mary’ neutralization rule.

11.3 Consonant Graphemes and Their Phonemic Values

11.3.1 Consonant Monographs

There are consonant monographs in English which have a single phonemic value. These are:

(12)
Grapheme
Phoneme
Example
<p>
/p/
pit
<k>
/k/
kit
<b>
/b/
bit
<d>
/d/
din
<j>
/dʒ/

joy
<f>
/f/
fun
<v>
/v/
van
<z>
/z/
zip
<l>
/l/
lip
<m>
/m/
mind
<w>
/w/
wet

There are also monographs with more than one phonemic (or phonetic) value:

(13)
Grapheme
Phonemes/Allophones
Examples
<y>
/j/
year
<n>
/n/, [m], [ɱ], [ŋ]
ip, input, inform, ink

<s>
/s/, /z/, /ʃ/, /ʒ/
sip, rose, compulsion, confuse
<g>
/g/, /dʒ/
got, gin
<x>
/ks/, /gz/, /kʃ/, /gʒ/
sex, exact, luxury, luxurious

For the grapheme <y>, we are treating words such as as beginning with a consonant: the palatal approximant /j/. The sounds [j] and [w] are often referred to as semi-consonants, since they are vowel-like in their articulation, but they occupy a consonantal position in syllables, namely the onset position.

As far as the grapheme <n> is concerned, we have seen that the phoneme /n/ undergoes nasal assimilation: if followed by a bilabial consonant, the realization is bilabial, as in ; if followed by a labio-dental consonant, its realization is labio-dental, as in ; and if followed by a velar consonant, its realization is velar, as in . These bilabial, labio-dental and velar realizations are all allophones of the /n/ phoneme.

The <s> grapheme corresponds to four phonemes for historical reasons: the palato-alveolar pronunciations /ʃ/ and /ʒ/ came about because the following consonant was a palatal approximant (known informally as a **yod**): the sequence [s] + [j] led to an assimilation process known as **coalescence**, in which a sequence of an alveolar and a palatal results in a palato-alveolar articulation. The same process can be seen in connected speech: a word-final /s/, when followed by a yod, can lead to assimilation. This can be heard in a song by the

Rolling Stones, entitled ‘Miss You’, pronounced [ˈmɪʃə] by Mick Jagger. Similarly, a word-final /z/ followed by a word-initial /ɛ/ can yield a [ʒ], as in , pronounced [hi ˈʒɜ:z]. Words such as , the past tense of , initially had a voiceless [s], often with a word-final schwa: [ro:sə]. It is common for voiceless segments to become voiced **intervocalically** (between vowels): since vowels are typically voiced, an assimilation process takes place which results in the intervocalic sound undergoing voicing. Thus [ro:zə], which later lost the word-final schwa. Much later, the /o:/ diphthongized in RP, resulting in the present-day [ɪəʊz] pronunciation.

The <g> grapheme corresponds to the /g/ phoneme, but also corresponds to the /dʒ/ phoneme, again for historical reasons relating to assimilation: since the vowel in words such as and is a high front vowel, the historical /g/ underwent palatalization, and became palato-alveolar. In words such as and , the vowel was the high-front /i:/ in Middle English, and thus induced the same palatalization process. Unfortunately, not all /g/ + high-front vowel sequences underwent this process, so that words such as and , while containing a high front vowel, have retained the velar pronunciation /g/. Some words, such as the proper name , can be pronounced either way, with the female name being pronounced with a [dʒ], and the male name being pronounced with a [g]. Similarly, the word , when meaning a part of a fish, is pronounced with a velar stop, whereas the unit of measurement known as a gill is pronounced with the palato-alveolar value.

The <x> grapheme corresponds to sequences of two phonemes. If we take the /ks/ sequence to be the ‘basic’ value, we can see that the same historical processes of voicing and palatalization have taken place to yield the other three values. The /gz/ sequences in words like have arisen from intervocalic voicing assimilation. In words like , the sequence /ksj/ gave rise to coalescence of the /sj/ sequence, resulting in /kʃ/. In words like , both coalescence and inter-vocalic voicing have resulted in the /kʒ/ sequence. Stress also plays a role here: if the preceding vowel is stressed, as in and , we have a voiceless value; if the following vowel is stressed, as in and , we have a voiced value.

11.3.2 Consonant Digraphs and Trigraphs

There are consonant digraphs and trigraphs with a single phonemic value:

(14)

Grapheme

Phoneme

Example

<ck>

/k/

clock

<ph>

/f/

photo

<sh>
/ʃ/
ship
<rh>
/ɹ/
rhubarb
<dg>
/dʒ/
edge
<d¡>
/dʒ/
adjunct
<ng>
/ŋ/
sing
<tch>
/tʃ/
itch

In addition to the consonant monographs with more than one value, there are consonant digraphs and trigraphs with more than one value. These are:

(15)
Grapheme
Phonemes
Examples
<gu>
/g/, /gw/
guard, anguish
<qu>
/k/, /kw/
unique, queen

<ch>
/tʃ/, /k/, /ʃ/
chair, chaos, chic
<gh>
/g/, /f/, zero
ghoul, tough, thigh
<th>
/θ/, /ð/
think, this
<sch>
/sk/, /ʃ/
school, schmaltzy
<sc>
/s/, /ʃ/, /sk/
scene, conscious, scour
<gg>
/g/, /dʒ/
egg, exaggerate
<cc>
/k/, /ks/
account, accent
<ss>
/s/, /ʃ/
kiss, mission

The <gu> digraph mostly corresponds to the /g/ phoneme, as in and , but can correspond to the /gw/ sequence. This can be found only in the middle of words, where there is no morpheme boundary, as in and , but not in, for instance, .

The <qu> digraph in words like was introduced by

French scribes after the Norman Conquest: words like this had previously been written with a <cw> sequence. Words such as have been borrowed from French with their French spellings.

The <ch> digraph replaced the <c> monograph in words like , which had previously been pronounced with a /k/, and is now pronounced with a /tʃ/. The value /k/ can be found in many words of Greek origin, such as and . The /ʃ/ value is found in fairly recent loanwords from French, such as and . We have not listed the value /x/ (the voiceless velar fricative), since it does not occur in most varieties of English. It does, however, occur in Scottish English, in words such as (/lɔx/), as opposed to (/lɔk/). While some non-Scottish speakers pronounce words such as with a [x], most produce a [k] when attempting Scottish words such as the place-name , German words such as or Spanish words such as .

The <gh> digraph used to represent the voiceless velar fricative /x/, which can be heard in Spanish words like (‘garlic’) or German words like (‘book’). The voiceless velar fricative no longer exists in RP or GA: it was elided in many words, like and , which is why we have represented its value as zero. The voiceless velar fricative can still be heard among older speakers of Scots, who pronounce as [dɔxtɪɾ] and as [θɔxt].

Most words spelled with <th> correspond to /θ/, since <th>

words which contain /ð/ are words of a non-lexical category (function words), such as and . While these are high-frequency words, there are very few of them, which is typical of function words.

Words spelled with the <sch> trigraph, such as and are often of Yiddish origin. Yiddish is a Germanic language (with vocabulary from Hebrew, Aramaic and Slavic languages) spoken by Ashkenazi Jews. The Yiddish words which have found their way into English mostly come from New York's Jewish community. Some of these words, such as (‘to carry around a heavy object’ or ‘to travel somewhere slowly’), are not known throughout the United States, and used not to be used in British English, but is beginning to appear in British newspapers. Other <sch> words are loanwords from German, such as (a German drink) and (meat fried in breadcrumbs). Words such as were borrowed from Latin, in which the <sch> sequence corresponded to /sk/, and has retained that sequence in English for over a thousand years.

Many words spelled with <sc> were borrowed from French. They initially had /sk/ sequences in Latin, but the /k/ had been elided in the French pronunciation by the time they were borrowed.

The sequences <gg> and <cc> correspond to the same values as <g> and <c>. Doubled graphemes in English do not correspond to long consonant phonemes (often called geminates), unlike in Italian, where spellings such as and correspond to the pronunciations /ˈmam:a/ and /

ˈpit:sa/, where the colon indicates length. Some sequences of identical consonant graphemes correspond to sequences of two identical consonant phonemes across a morpheme boundary, as in , which has the form /ˌʌn + ˈnɜ:v/, realized as [ˌʌnˈnɜ:v]. A sequence of two identical phonemes in English is not the same thing as a single, long (geminate) phoneme: there is no such phoneme as /n:/ in English. The fact that English speakers do not pronounce Italian loanwords such as with a geminate consonant reflects the fact that English has no geminate consonant phonemes.

The <ss> sequence has /s/ as its 'basic' value, as in and , but the kinds of palatalization process we discussed above have yielded /ʃ/ in words which have had a yod historically, like and .

11.3.3 Unpronounced Consonant Graphemes

Some word-initial consonants have been elided as English has evolved. These include /k/ and /g/ in branching onsets with /n/, so that the following digraphs no longer correspond to sequences of onset phonemes:

(16)

Grapheme	Phoneme	Example
<kn>	/n/	know

<gn>
/n/
gnome

Other words were borrowed from Greek, whose phonotactic constraints allow onset sequences such as /ps/ and /pt/. Since English phonotactics do not allow such sequences in onsets, the following correspondences have arisen:

(17)
Grapheme
Phoneme
Example
<ps>
/s/
psyche
<pt>
/t/
Ptolemy

Other graphemic sequences which reflect historical process of elision include the following:

(18)
Grapheme
Phoneme
Example
<wr>
/ɹ/
write
<wh>
/w/, /h/
whine, whole

Old English had an onset sequence /hw/ in words like .
In RP and GA, the /h/ has since elided, while the spelling still

indicates a cluster of two phonemes. In SSE, and in some varieties of American English, the /hw/ sequence has become [ʍ], a voiceless fricative which has both a bilabial and a velar articulation, rather like the sound one makes when blowing out a candle. Pairs such as and are minimal pairs in SSE: /ʍɪtʃ/ vs /wɪtʃ/, whereas they are homophones in RP and GA.

There are also word-final digraphs which correspond to a single phoneme, since the word-final consonant has elided during the evolution of the language:

(19)
Grapheme
Phoneme
Example
<mn>
/m/
hymn
<mb>
/m/
comb
<gn>
/n/
sign
<gm>
/m/
paradigm

In morphologically related words, the root-final consonant is no longer word-final, and has therefore often been retained, as in the following:

(20)
Spelling

Pronunciation
hymnal
[ˈhɪmnəɫ]
signatory
[ˈsɪgnətɹi]
paradigmatic
[ˌpæɹədɪgˈmætɪk]

In word-medial position, <t> is often not pronounced in the following sequences: <st> followed by a vowel grapheme, as in , and <stl>, as in .

11.3.4 Graphophonemics and Contrastive Phonemics

Some learners of English as a foreign language encounter problems which relate both to the English graphophonemic correspondences to the difference between the phoneme system of their native language and that of English, especially RP English. An example concerns mid vowels. It is common to find languages with a core monoph-thong system like the following:

(21)
/i/
/u/
/e/
/o/
/ɛ/
/ɔ/
/ɐ/

By **core** monophthong system is meant a system ignoring, for instance, nasalized monophthongs such as the French phonemes /ɛ̃/, /ɑ̃/ and /ɔ̃/, or front rounded vowels, such as the French phonemes /y/, /ø/ and /œ/, all of which are said by linguists to be 'marked', i.e. relatively uncommon in the world's languages.

In the above diagram, there are two high vowels (/i/ and /u/), two high-mid vowels (/e/ and /o/), two low-mid vowels (/ɛ/ and /ɔ/), plus a single low vowel which is neither fully front nor fully back, which is why we have chosen the /ɐ/ representation. Examples of languages with such a system are the Romance languages French, Spanish and Italian. If one speaks a language with this kind of core system, the vowel phoneme system of RP is quite a challenge. To begin with, it has high back rounded vowels (/u:/ and /ʊ/), rather than one, which is the norm in the world's languages. Its also has two 'a'-type vowels (/æ/ and /ɑ:/), or even , if we consider that the vowel /ʌ/ is rather like an 'a' sound, and thus often pronounced with an [ɐ] in languages which lack an /æ/–/ʌ/–/ɑ:/ contrast. The back mid vowels of RP are especially difficult, since RP has three 'o'-type phonemes: /əʊ/ (which used to be monophthongal /o:/ in Middle English), long/tense /ɔ:/ and short/lax /ɑ/. For speakers whose native language has only two 'o'-type phonemes, a phonological difficulty arises: how to expand the two perceptual categories into three.

The phoneme system of our native language, which we have in our heads, is a system of perceptual categories, via which we decode speech. It is difficult to perceive a three-way contrast in a foreign language if it corresponds to a two-way contrast in our native language. For instance, English has a two-way contrast between alveolar and palato-alveolar fricatives: /s/–/ʃ/ and /z/–/ʒ/. If an English speaker tries to learn

Polish, a difficulty arises with the fricatives, since Polish has a way contrast between alveolar, palato-alveolar and pre-palatal fricatives, transcribed as /s/, /ʃ/ and /ś/. Because we English speakers are so accustomed to perceiving only two categories of voiceless fricative in this region of the oral cavity, we have difficulty with both perceiving and producing the three-way Polish contrast.

For speakers of many languages who are learning English, the same difficulty applies to the RP contrast between /əʊ/, /ɔ:/ and /ɑ/. Standard Scottish English would be easier for such speakers, since its core monophthong system is very similar to the one depicted above: only one ‘u’-type sound, a two-way /e/–/ɛ/ distinction, a twoway /o/–/ɔ/ distinction and only one ‘a’ phoneme (though SSE does have the /ʌ/ phoneme). The phonological difficulty with ‘o’ sounds is compounded by the vagaries of English spelling since, as we have seen, there are various ways of spelling the three different ‘o’ sounds. We have seen that the <o> grapheme can correspond to either /əʊ/, /ɔ:/ or /ɑ/. However, there is some light at the end of the tunnel for non-native speakers learning RP: we have cited graphophonemic regularities above which determine, to a large extent, which values we find. In addition, the sequences <aught>, <ought> and <aw> systematically correspond to the RP long vowel /ɔ:/, as in and , and not to RP /əʊ/. This is a point worth bearing in mind for learners of English, since there are RP minimal pairs such as and , which have, respectively, /əʊ/ and /ɔ:/. An anecdote might serve to highlight the problem: when we first bought a house in France, a French colleague asked me a question about it. I thought the question was ‘Do you have a loan?’, to which

I answered 'Yes, of course: we're not wealthy enough to have paid cash for the house.' What he meant was 'Do you have a lawn?', but he produced /əʊ/ for the <aw> sequence in , rather than the /ɔː/ pronunciation.

To conclude on English spelling-to-pronunciation correspondences: the regularities we have given above show that these are not completely chaotic. There are, though, many exceptions to the 'rules' we have provided, and those make English graphophonemics rather more messy than it might otherwise have been. We have seen, for instance, that words such as really ought to be written , as in , with a double consonant grapheme, and my own name, , ought to be written , to indicate the checked value ([ɪ]) of the pre-final stressed vowel. However, it is best, perhaps, to emphasize the regularities, rather than make a very long list of exceptions: one cannot, after all, have irregular forms unless there are regular forms.

Exercises

Listen to sound files online

1 Listen to Track 11.1 at www.wiley.com/go/carrphonetics, made by an RP speaker. For each of the words on the recording (listed below), say whether the vowel corresponds to the free or the checked value of the vowel grapheme. Explain why. Provide the symbol, in slanted brackets, for each vowel phoneme.

(a) dam
(b) dame
(c) den
(d) dene
(e) Tim
(f) time
(g) dot
(h) dote
(i) dun
(j) dune

Listen to sound files online

2 Listen to Track 11.2, made by an RP speaker. For the stressed syllable in each of the words on the recording, say whether the vowel grapheme corresponds to the free or checked value and explain how the spelling encodes those values.

(a) taper
(b) tapper
(c) Peter
(d) petter
(e) pining
(f) pinning
(g) doting
(h) dotting
(i) astuter
(j) a stutter

3 Listen to Track 11.3, made by an RP speaker. For the stressed syllable in each of the following words, say whether the vowel grapheme corresponds to the free or checked value and explain how this relates to the way the word is spelled.

(a) sanity

(b) episode

(c) citadel

(d) opera

4 Listen to Track 11.4, made by an RP speaker. Which value does the <gh> digraph correspond to in each of the following words?

(a) rough

(b) through

(c) ghastly

12

Variation in English Accents

12.1 Introduction

In this chapter, we will consider some general aspects of accent variation. In chapter 13, a brief overview is given of several accents of English: London English, Tyneside English, Standard Scottish English (SSE), New York City English, Texan English, Australian English and Indian English, followed by an outline of the sorts of phenomena which give rise to divergence of accents over time.

Three of the accents we have referred to in this book (GA, RP and SSE) are viewed socially as ‘standard’ accents. The notion ‘standard’ is a social one: no linguist would claim that there is any coherent notion of inherent phonetic or phonological superiority, since such a notion simply does not make any phonetic or phonological sense. There can be no doubt that many people judge some accents to be superior to others, or take some accents to be standard accents and others to be non-standard accents. But those judgements are founded on non-linguistic factors, to do with social attitudes in the societies in question. From a strictly linguistic point of view, such judgements are, quite clearly, entirely arbitrary. For example, RP, the standard accent in England, is non-rhotic, and the non-rhoticity of RP is therefore judged by some (perhaps many) English people to be more prestigious than the rhotic accents found in many of the Western parts of England. But the standard accents SSE and GA are rhotic, and in the United States, it is the rhotic accents which are often judged to be more prestigious than the non-rhotic American accents (the judgement cannot arise in Scotland, where all native accents are rhotic).

Clearly, it is social attitudes which determine such judgements about accents, rather than the phonetic and phonological properties of the accents themselves. It is common to find social judgements to the effect that some accents are 'uglier' or 'harsher' than others. These judgements too are entirely arbitrary as far as phonetics and phonology are concerned. For instance, if the word , pronounced as [saɪ] in London English, is judged 'ugly' by an RP speaker, then the RP speaker's own pronunciation of the word , as [saɪ], ought also to strike the RP speaker as 'ugly'. Cases such as this show that it simply cannot be any phonetic properties of the sounds [saɪ] in and of themselves which induce the aesthetic judgement. Rather, such judgements also derive from and reflect social attitudes about what might, rather broadly, be called 'ways of life'. In Britain, most non-standard accents which are judged 'ugly' or 'uncivilized' are spoken in industrial or post-industrial urban areas; examples often cited are the working-class accents of London, Birmingham Liverpool, Belfast, Glasgow and Tyneside. Similar sorts of judgement are made in the United States, with respect to, for example, the broad New York City accent often referred to as 'the Brooklyn accent' (though it is not confined to the Brooklyn district of New York City).

Non-standard rural accents are, by contrast, often judged ‘quaint’, rather than ‘ugly’; examples are accents from the Highlands of Scotland, from the West Country in England, or from the US Southern states. It is highly likely that these aesthetic judgements arise from the conscious or unconscious association of accents with the real or imagined ways of life of those who speak them. We will therefore stand back from such social attitudes and attempt to examine accent variation from the point of view of the phonetician and the phonologist.[1]

12.2 Systemic vs Realizational Differences between Accents

Let us begin by considering one of the differences between many accents in the North of England and many of those spoken in the South of England. In the latter accents, there is a phonological contrast between /ʊ/ and /ʌ/, which can be observed in pairs such as and many others. That distinction is missing in many Northern English accents,[2] which have /ʊ/ in each member of the pair. That is, many Northern English accents simply lack the /ʌ/ phoneme, and thus the /ʊ/ vs /ʌ/ distinction: for them, pairs such as and are homophones, not minimal pairs. We will refer to this sort of difference as a **systemic** difference[3] between two accents: the set of phonological contrasts (specifically, in this case, the vowel contrasts) of the speakers differ.

Systemic differences are widely attested. In many Scottish English and Scots accents, for instance, there is no equivalent of the /æ/ vs /ɑː/ distinction, of the sort found in RP minimal pairs such as , etc. In Scottish English, each member of such pairs contains the same vowel phoneme, /ɐ/, which is realized as [ɐ]. Again, pairs which count as minimal pairs in non-Scottish accents are homophonous in Scottish English. Similarly, as we have seen, while RP has a three-way distinction between [ɒ], [ɑː] and [ɔː], many speakers of GA have only a two-way distinction between [ɔː] and /ɑ/. This systemic difference means that there is a difference in the sets of words which are distinguished by means of these phonemes, as follows:

(1)

Words of the type	RP	GA
palm	/ɑː/	/ɑ/
caught	/ɔː/	/ɔː/ or /ɑ/
cot	/ɒ/	/ɑ/
coffee	/ɒ/	/ɔː/

Systemic differences are not restricted to vowel systems. An example of a systemic difference in consonant systems is the contrast, found in many Scottish accents, between /ʍ/ and /w/, which is found in minimal pairs such as vs vs vs , etc. That contrast is absent in most non-Scottish accents, which have a /w/ phoneme, but not a /ʍ/ phoneme. In those accents, the above pairs are homophones, rather than minimal pairs.

There are differences between accents which do not amount to a difference in the systems of phonological contrasts. Consider our discussion of dark and non-dark /l/ in chapter 7: we said there (7.6) that there is an allophonic rule in RP, to the effect that /l/ is realized as a velarized ('dark') lateral when it occurs in rhymes, but is realized as a non-velarized ('non-dark') lateral when it occurs in onsets. In many accents of the South of Scotland, in Australian English, in GA and in some accents of the North of England, /l/ is realized as a dark lateral in all positions, so that /lʌl/ ( ), for instance, is realized as [ɫʌɫ], rather than [lʌɫ]. In Tyneside English, on the other hand, /l/ is consistently realized as a clear lateral in all positions, so that /lʌl/ ( )[4] is realized as [lʲʊlʲ]. There is no question of postulating, in these sorts of case, a difference in the underlying system of contrasts: all three accents have a contrast between /l/ and other consonant phonemes, such as /r/, but there is variation in the way that /l/ is realized. While there are languages in which the distinction between clear and dark laterals is contrastive, there can be no question of postulating such a phonological contrast in any of these accents of English: in the Southern Scottish and Australian cases, there are only dark laterals; in the Tyneside case, there are only clear laterals; and in RP and GA, while there are dark and non-dark laterals, the distinction is purely allophonic: the distinction lies at the level of realizations of a single phoneme. We will therefore refer to these sorts of difference as **realizational** differences.

Realizational differences involving vowels are very common. For instance, one of the differences between SSE and most non-Scottish accents is the presence of allophonic vowel length in SSE. This is absent in most non-Scottish accents, but this difference is a purely realizational matter. Take the contrast between /i/ and /u/ in SSE (as in): a parallel contrast is also present in most non-Scottish accents, including RP. In RP, the equivalent phonemes, /iː/ and /uː/, are, in all contexts, typically realized as [iː] and [uː] respectively, as in and . In SSE, however, /i/ is realized as either short [i] (as in) or long [iː] (as in), and /u/ is realized either as short [ʉ] (as in) or long [ʉ] (as in), depending on the phonological context: the long realization occurs word-finally and before the voiced segments [ɹ], [z], [v], [ð] and [ʒ]; the short realization occurs elsewhere.

Realizational differences can become systemic differences over time. For instance, at a stage in the history of RP when it was still a rhotic accent, the /iː/ phoneme was realized as [iːə] before /ɹ/ in coda position, in much the same way as /iː/ is currently realized as [iːə] before /l/ in coda position. At that stage in the history of RP, pairs such as and differed in two respects: one had [ɹ], while the other did not, and one had [iː], while the other had [iːə]. With the gradual loss of the [ɹ] articulation in coda position, pairs like this became minimal pairs: [biːəd] ( ) vs [biːd] ( ). It is reasonable to say that, at that stage, a new /iːə/ phoneme (the ancestor of present-day /ɪə/) emerged. This process is known as a **phonemic split**: a distinction which was once allophonic becomes phonemic. In terms of differences between accents, we want to say that, at a point when both RP and, say, SSE were rhotic, they both had the /i:/ phoneme, but there was a realizational difference, in that RP, but not SSE, had an [i:ə] allophone of the /i:/ phoneme. Now, however, there is a systemic difference: RP has an /i:/ vs /ɪə/ contrast which SSE lacks.

Another example of a phonemic split which occurred in the history of RP and many other varieties of English is the FOOT/STRUT split, mentioned above. The STRUT phoneme evolved historically from the FOOT phoneme: at one stage in the history of English, all realizations of the FOOT vowel were rounded. But unrounded realizations began to appear, and these eventually took on phonemic status, resulting in the emergence of the STRUT vowel, pronounced [ʌ] in RP and many other varieties of English. Many accents in the North of England failed to undergo the FOOT/STRUT split, so that words like are pronounced with an [ʊ]. The difference between these North of England accents and RP is therefore a systemic difference: RP possesses a phonemic distinction which is absent in those Northern varieties. As a result of this, pairs of words which are minimal pairs in RP, such as ([pʰʊt]) and ([pʰʌt]), are homophones in the Northern varieties: both and are pronounced [pʰʊt].

How can one tell, for a given difference, whether it is systemic or realizational? Let us examine a particular case, that of London English. By ‘London English’ (henceforth, LE), we do not mean the accent spoken by all natives of London; rather, we refer, albeit in a necessarily oversimplified way, to the speech of working-class natives of London.[5] It has been widely noted that speakers of LE often (but variably) utter words such as with an [aɪ] diphthong, and we know that, in RP, these have an [eɪ] diphthong. Is this a systemic or a realizational difference?

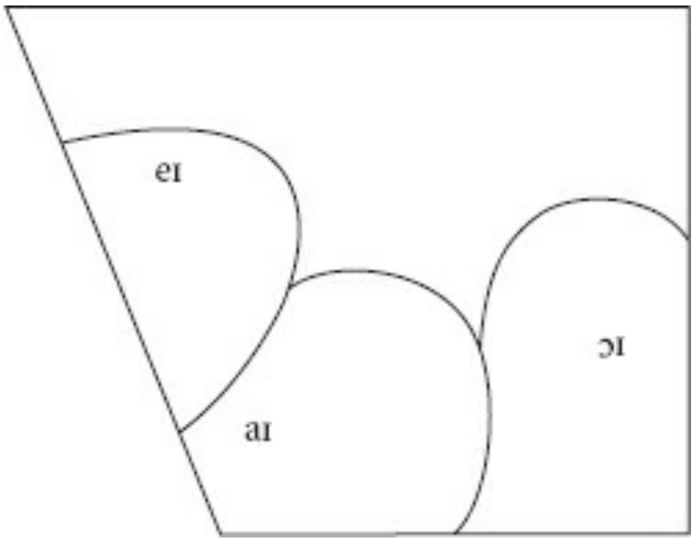
The answer is that we cannot tell on the basis of this evidence alone. In order to establish whether there is a systemic difference between the accents, we must consider the system of contrasts in each accent. Specifically, we must ask: is there a phonetic [eɪ]/[aɪ] distinction in either accent, and if so, is it contrastive? We have already established that, in RP, there is such a distinction, and that it is contrastive (cf. the minimal pairs vs vs vs , etc.). What we must then establish is whether these pairs are also minimal pairs in LE: if they were to turn out to be homophones in LE, then we could, all things being equal, reasonably conclude that there is a systemic difference here, in just the same way that we concluded, on the basis of evidence from minimal pairs and homophones, that RP has an /ʊ/ vs /ʌ/ contrast, which many Northern English accents (with /ʊ/ alone) lack.

What we find is that these are indeed minimal pairs in LE: while , etc. have an [aɪ] diphthong in LE, , etc. have an [ɔɪ] diphthong. In the absence of any further evidence to the contrary, we may conclude that there is no systemic difference here: the contrast which we found in RP is maintained in LE.

But the matter does not end there, since we also know that RP has a contrast between /aɪ/ and /ɔɪ/, as in vs . What we must now ask is whether this contrast is also sustained in LE, or whether these RP minimal pairs are homophones in LE. The answer is that the contrast is sustained in LE: while etc. have [ɔɪ], , etc. have [oɪ].

We have now noted a related set of non-systemic (non-contrastive), purely realizational differences between RP and LE, which we might conceive of in terms of the articulatory and perceptual 'space' in which the realizations of a phoneme are located. We may depict this in terms of the vowel space diagram. Consider the phonetic realizations of the RP vowel phonemes /eɪ/, /aɪ/ and /ɔɪ/. Let us imagine that speakers of RP typically have diphthongal realizations of those phonemes whose starting points fall within the following sorts of articulatory 'zones' in the vowel space:

(2)



The speaker might well vary in her or his realizations of these phonemes. However, so long as the articulation of a particular vowel does not encroach upon the space of any of the others, will still be distinguishable from , and from . If, in the course of the historical development of the accent, the articulations of /eɪ/ and /aɪ/ were to become so close as to be perceptually indistinguishable, such that pairs like vs were both uttered as [baɪ], then the contrast would be lost. That phenomenon, which is widely attested in the histories of human languages, is referred to as a **phonemic merger**: where once a phonemic contrast was present, it came to be collapsed.[6]

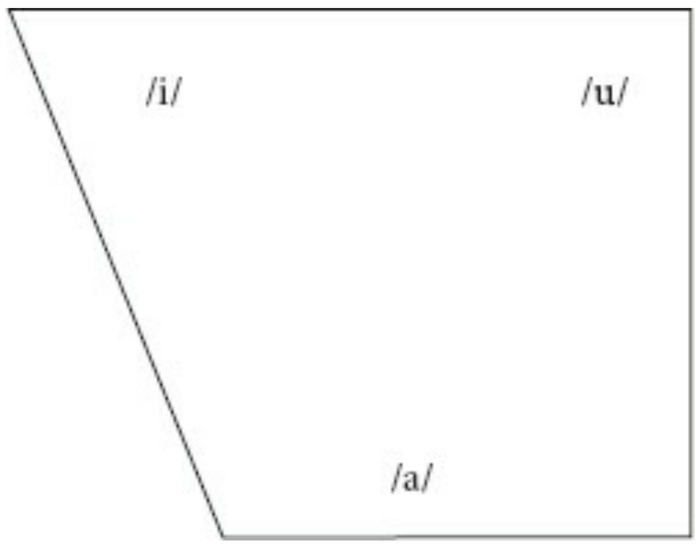
Clearly, no merger has occurred in the LE cases cited above: while the realization of /eɪ/ has indeed 'shifted' to [aɪ], the realization of [aɪ] has, in turn, shifted, to [ɔɪ], and the contrast is thus maintained. Similarly, while the realization of /aɪ/ has shifted to [ɔɪ], the realization of /ɔɪ/ has, in turn shifted, to [oɪ], and again the contrast has thus been maintained.[7] This kind of phenomenon, which is fairly widespread, is often referred to as a **vowel shift**. A parallel shift is evident in Australian English, where the /iː/ phoneme has diphthongized to [ɪi], thus encroaching, to some extent, on the space of /eɪ/. The /eɪ/ phoneme has in turn shifted to [ɐɪ], with a fairly low, central, unrounded starting point, thus encroaching on the perceptual space of /aɪ/. In turn, /aɪ/ has shifted to [ɔɪ], thus encroaching on the space of /ɔɪ/. However, the /ɔɪ/ phoneme appears not to have taken 'evasive action', so [ɔɪ] and [ɔɪ] are contrastive. The point to be borne in mind is that vowel shifts are purely realizational: they do not involve a change in the number of phonological contrasts.

We asked whether the realization of /aɪ/ as [ɔɪ] resulted in the destruction of a contrast which is present in RP. We might equally ask whether the realization of the /ɔɪ/ phoneme as [oɪ] in LE results in the destruction of a contrast found in RP; the answer is that, although RP speakers may well utter both [ɔɪ] and [oɪ], the phonetic difference between the two, while perceptible, is never contrastive: one cannot cite minimal pairs involving the two. So there is no phonemic contrast which would be collapsed by uttering /ɔɪ/ as [oɪ] (the RP speaker is free to do so, without risk of conflating a contrast). For the LE speaker, while the realization of /eɪ/ has encroached upon the ‘space’ of /aɪ/, and the realization of the latter has encroached upon the space of /ɔɪ/, the realization of the latter has not encroached upon the space of any other phoneme.

12.3 Perceptual and Articulatory Space

The simplest system of vowel phonemes[8] found in the world’s languages is the three-vowel /i/, /u/, /a/ system, typically depicted within the vowel space as:

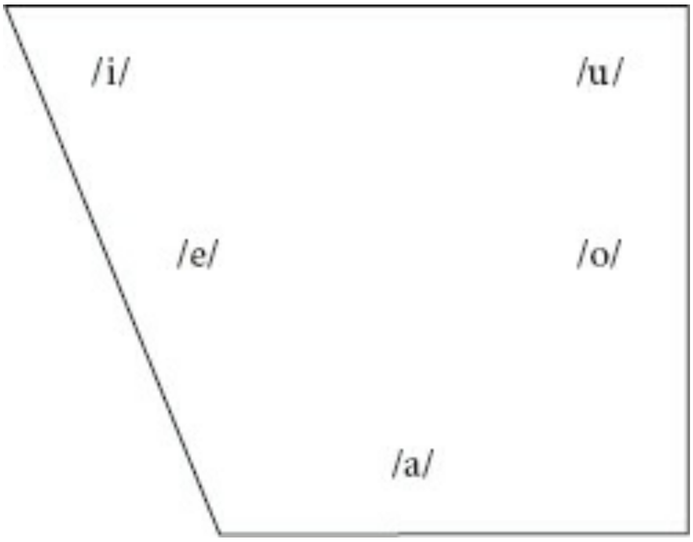
(3)



The realizations of the vowel phonemes in a system like this can vary considerably; it will matter little in terms of the hearer’s identification of a given vowel phoneme if /i/ is realized as an [e]-type vowel, so long as it is relatively front, unrounded and relatively high. Nor will it matter if /u/ is realized as an [o]-type vowel, so long as it is relatively high, relatively back and rounded. Similarly, it will not matter whether realizations of /a/ are front (like [a]), back (like [ɑ]), or central (like [ɐ]), so long as they are relatively low and unrounded. That is, each vowel phoneme has quite a large perceptual and articulatory space.

A slightly larger vowel system, frequently encountered in the world's languages, has mid vowels: /i/ /u/

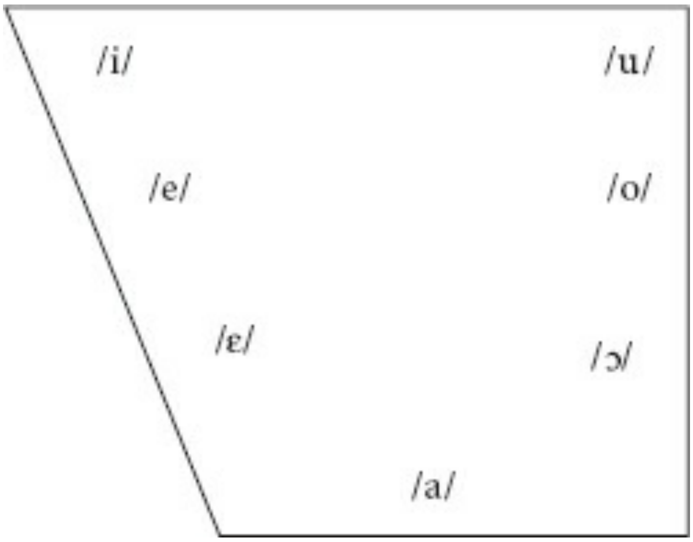
(4)



In a system like this, the vowel space is a little more crowded: it matter whether a realization of /i/ is [e]-like, or if a realization of /u/ is [o]-like. But it will not matter if a realization of /e/ is [ɛ]-like, so long as it is not [a]-like. Nor will it matter if a realization of /o/ is [ɔ]-like.

In a slightly larger system, there are contrasts between high-mid and low-mid vowels:

(5)



In this kind of system, there is even less articulatory and perceptual space for each vowel phoneme: it will matter if the realizations of /e / are [ɛ]-like, for instance.

Many English accents have relatively large vowel phoneme systems, often containing, for instance, as many as three or four ‘a’-type phonemes, as in RP’s front /æ/, central /ʌ/ and back /ɑ:/. This means that shifts in realization can easily result in one phoneme encroaching upon the space of another. And this notion of articulatory and perceptual ‘space’ has a bearing on the question of pairs of vowel phonemes we should consider when attempting to establish whether a given vowel difference is systemic or not. One way of answering this question is to say that one should consider vowel phonemes which are in some sense ‘adjacent’ in articulatory and/or perceptual terms. This is what we did when we considered the [eɪ] vs [aɪ] distinction: they are adjacent in that the starting points of both diphthongs are front, non-high and unrounded. We might equally have considered the [ɛ] vs [eɪ] distinction, since [e], the starting point for the latter, is close to [ɛ] in articulatory terms. Indeed, one finds that the /ɛ/ phoneme is often realized as an [eɪ] diphthong in LE (thus, can be pronounced as [weɪw]). This, however, does not cause merger of the /ɛ/ vs /eɪ/ contrast since, as we have seen, /eɪ/ is realized as [aɪ].

The main difficulty one encounters in considering adjacent vowel sounds is that a given vowel can be adjacent to many (perhaps most) of the other vowel sounds found in a particular accent. As we have just seen, [eɪ] is adjacent to both [aɪ] and [ɛ]. It is also adjacent to [ɜ:], since the starting point for [eɪ] is front, mid and unrounded, and [ɜ:] is central, mid and unrounded. As it happens, [ɜ:] has not shifted in LE, but it has in other accents. In the Liverpool accent, for instance, it has fronted to [ɛ:], so that and , which are pronounced [wɜ:k] and [bɜ:d] in RP, are pronounced [wɛ:x] and [bɛ:d] in Liverpool. This means that the realization of [ɜ:] has encroached upon the articulatory and perceptual space of the adjacent vowel phoneme /ɛ/. However, the phonemic distinction between [ɜ:] and /ɛ/ is maintained in Liverpool, since the former is long with respect to the latter, as in vs ([bɛ:d] vs [bɛd]) and thus distinguishable from it. For many Tyneside speakers, [ɜ:] has been retracted and rounded in some words, thus encroaching on the space of [ɔ:]. Thus is pronounced [wɔ:k] by many speakers. This has resulted in a loss of the [ɜ:] vs [ɔ:] contrast in a particular set of words, so that pairs such as and are homophones rather than minimal pairs.

It appears that vowel articulations are especially susceptible to this ‘shifting around’ phenomenon and it is because of this that the majority of differences between accents are differences in the articulation of vowels, rather than of consonants. This is perhaps because of the nature of vowel articulations, which all have a stricture of open approximation. Furthermore, the more open the vowel, the more open the approximation, and the less contact there is between the tongue and the other parts of the oral cavity. It seems clear that, for any given phonetic segment, we are unlikely to hit on exactly the same articulation each and every time we attempt it. Not every [d] we utter will have exactly the same part of the tongue closing against exactly the same part of the alveolar ridge for exactly the same length of time with exactly the same amount of vocal cord vibration beginning and ending at exactly the same time. So variation is inherent in speech.

But it appears to be even more inherent in vowel articulations than it is in, say, stops, since it is much harder to feel where one's tongue is in one's mouth when one produces a vowel than when one produces, say, a stop. One can feel the lips close together for the articulation of a bilabial stop, and one can feel the tongue against the alveolar ridge when one produces an alveolar stop, so that the articulatory difference between the two is easily discerned by the speaker. But when one produces, say an [e:] as opposed to an [ɛ], it is much harder to feel what one's tongue is doing. One can therefore easily 'overshoot' or 'undershoot' and end up with articulations which encroach upon the space of another, adjacent, vowel phoneme. Vowel articulations, in short, form a continuum; there are no abrupt, discrete divisions between them.

Parallel to this articulatory continuum, there is a perceptual one. Recall that our 'cardinal vowels' are merely reference points, based on an arbitrary carving up of the available vowel space and the available parameters of tongue height, frontness/backness and roundedness, all of which are a matter of degree, rather than a matter of absolutes. Once an [e]-type sound begins to lower, at what point does it become an [ɛ]-type sound? Once an [æ]-type sound begins to raise, at what point does it become an [ɛ]-type sound? The answer is that it is impossible to say with any certainty. Little wonder, then, that when some American English speakers say , speakers of many accents of British English think that they are saying : their [æ]-type sound has raised to what is perceived as an [ɛ]-type sound. We perceive vowels the way we perceive colours. If we are presented with an example of a classic example of green, we have no difficulty in identifying it as green. The same is true for a classic example of blue. But once we are presented with a colour that is half-way between the two, we often cannot tell whether we think it is greenish-blue or bluish-green. In such in-between cases, one person will judge, say an item of clothing as green, while another will judge it to be blue. So it is with vowels: we find it hard to categorize a vowel half-way between [e] and [ɛ]. If a vowel that was once close to prototypical [ɛ] begins to be produced closer to [e], hearers may categorize it as an instance of an [e], and thus may, in turn, begin to articulate it as an [e].

The most important point is that our perception of acoustic events is heavily dependent on the mentally represented system of phonological contrasts which we have in our native accent. It is also vital that, for a phonemic distinction to exist, it must be based on a phonetic distinction which is perceptible to human beings. Those phonetic differences can be minute, but they must be perceivable. If a language does not have a large vowel phoneme inventory, the realizations of each vowel phoneme are much more 'free to roam' in the available articulatory and perceptual space.

There are consonantal continua too, but there is a little more in the way of discrete divisions among consonants than among vowels. A bilabial articulation, for example, is radically distinct from an alveolar one, since the tongue is not implicated at all in the former, and the lips need not be involved in the latter. But even then, we have seen, particularly among consonantal articulations which are more vowel-like, that a gradual transition from alveolar to labial is indeed possible, as we have seen with [w] realizations of /l/. Once a secondary velar articulation is present in the pronunciation of an alveolar lateral, it can become the primary articulation, the original alveolar stricture can be lost altogether, and a new bilabial stricture can emerge.

12.4 Differences in the Lexical Distribution of Phonemes

There is a further kind of variation between accents in which the phonological representation of words is concerned, but which is not systemic variation. For instance, in many accents of the North of England, there is some kind of /æ/ vs /ɑː/ distinction, parallel to that found in RP. The actual vowel quality of the 'long a' often differs from that of RP. Many West Yorkshire speakers have a much more front articulation for 'long a' than RP speakers; very often the difference is between a short [a] and a long [aː], with the vowel qualities being the same, and the two vowels differentiated only with respect to length. Nonetheless, the phonemic distinction is there, and functions as the basis of minimal pairs such as and . We can say that West Yorkshire's /a/

vs /a:/ distinction is parallel to, or equivalent to, RP's /æ/ vs /ɑ:/ distinction.

But speakers of such Northern English accents (and indeed of GA) often utter a short vowel in words which would be uttered with a long vowel by the RP speaker, words of the lexical set BATH. Examples, for many Northern speakers, are and , all ending with a voiceless fricative. Other members of the set have a nasal-plus-consonant sequence, such as and . What we want to say about this sort of difference is not that the Northern English speaker lacks the 'long a' vs 'short a' phonemic distinction (in the way that Scottish speakers do) but that the phonological form of those particular words contains the short, rather than the long, phoneme. We can illustrate this as follows:

(6) Systemic difference vs lexical distribution difference

Phonemic
Phonological forms

'long/short a'?
of
RP speaker:
Yes
/ænt/, /ɑ:nt/, /bɑ:θ/
Northern speaker:
Yes
/ant/, /a:nt/, /baθ/
Scottish speaker:
No
/ɐnt/, /ɐnt/, /bɐθ/

Such differences are very susceptible to variation between different accents in the range of words which have one phoneme rather than the other in their phonological representations. They are therefore less general in nature, and more idiosyncratic, than systemic and realizational differences; however, such differences can cause major problems for mutual intelligibility between speakers of different dialects.

We have now begun to note some differences between accents and dialects, in a way which allows us more insight into the nature of those differences than merely noting different pronunciations, and which shows the extent to which theoretical considerations play a part in our analyses. While it is informative to note that many speakers in the North of England pronounce words like as [bʊs] and that LE speakers pronounce words like as [saɪ], or to note that some speakers utter an [ɹ] in some contexts rather than others, what we have done here is to say a little more than just that: we have sought to gain some insight into the nature of the phonological knowledge possessed by speakers of different accents.

Notes

1 This is not to deny that social attitudes to accents affect accent variation itself. Clearly, if we are investigating the way people speak, and if the way people speak is influenced by social attitudes to accents, then we are obliged to recognize those factors as part of the general picture.

2 Many, but not all. Some Northern accents have a distinction between /ʊ/ and /ʌ/, in which the realization of /ʌ/

is an unrounded version of [ʊ], which we might reasonably transcribe as [ɨ].

3 An alternative term would be ‘contrastive difference’.

4 Many Tyneside speakers lack the /ʊ/ vs /ʌ/ distinction.

5 The term ‘working-class’ is vague, and enormously difficult to define. We will, nonetheless, assume that it is meaningful and does serve to identify what are real, if complex, differences in social class (whatever problems may reside in the vague notion ‘social class’ itself). The term ‘London’ is equally vague; nonetheless, it also serves a useful function, since it too allows us to identify a genuine, if hard-to-define, geographical, and perhaps cultural, entity.

6 Clearly, if there is variation between one realization and the other, and this is not noted by linguists, then a contrast that might have been wrongly taken to have been merged could re-emerge. Additionally, a contrast may merge in some phonetic environments, but not in others, leaving some minimal pairs intact while collapsing others. See 12.4 below on lexical incidence.

7 We are assuming here that LE and RP have a common source, and that LE has innovated historically in its realizations of these phonemes in particular in a way which RP has not. This assumption happens to be justified in this case, but it must not be assumed that all cases are like this. There are cases in which it is RP which has innovated. Such is the case with the /ʌ/ vs /ʊ/ contrast, which is a Southern innovation: many Northern accents have not undergone that innovation.

8 By ‘vowel system’ here, we mean ‘system of monophthongs’, ignoring diphthongs. Generally speaking, monophthongs are ‘more basic’ than diphthongs, but

diphthongs are common, and the most common in the world's languages are the [ai]-type and [au]-type diphthongs.

Exercises

Listen to sound files online

1 Listen to Tracks 12.1, 12.2 and 12.3 at www.wiley.com/go/carrphonetics. For each sound file, answer the following questions:

(a) Is the speaker rhotic or non-rhotic? (Cf. chapter 7.) What evidence do you have for your answer?

(b) Does the speaker have a rounded or an unrounded vowel in words of the sort LOT?

(c) What is the pronunciation of words spelled <wh> and <W>?

(d) Does the speaker exhibit the phenomenon of Flapping? (Cf. chapter 9.) One of these speakers is an RP speaker, another is a GA speaker and the other is an SSE speaker. Which is which? Relate your responses to your answers to the questions above.

2 Where there are differences between the speakers with respect to the phenomena in (1a–d), say, for each difference, whether it is a systemic or a realizational difference, and explain why.

13

An Outline of Some Accents of English

13.1 Some British Accents

13.1.1 London English (Track 13.4 at www.wiley.com/go/carrphonetics and exercise 1)

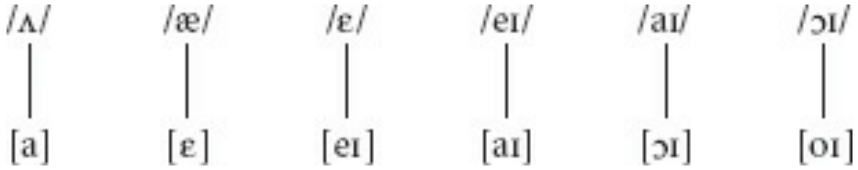
Listen to sound files online

By 'London English' we mean the vernacular, working-class London accent spoken in the boroughs of the East End of London. There are more and less broad varieties of this accent, ranging from Cockney at one end of the spectrum to accents which are closer to RP in some respects at the other end.

Let us postulate that the vowel system of LE is exactly parallel to that of RP, but with many major realizational differences, mostly of a ‘vowel shift’ nature, as discussed earlier in chapter 12.

Realizations of /ʌ/, /æ/, /ɛ/, /eɪ/, /aɪ/ and /ɔɪ/

In LE, the /ʌ/ phoneme is often realized as a short [a]-type sound. This then encroaches upon the /æ/ phoneme, which is often realized closer to [ɛ]. That in turn encroaches on the /ɛ/ phoneme, which in turn is often realized as a diphthong close to [eɪ], and that realization in turn encroaches on the perceptual space of /eɪ/. As we have seen, the /eɪ/, /aɪ/ and /ɔɪ/ phonemes also participate in this vowel shift, which we may depict as follows:



Realizations of /aʊ/, /ʌ/, /æ/ and /ɑ:/

Additionally, in the Cockney version of LE, the /aʊ/ phoneme is often realized as a long [a:], as in ([sa:nd])[1] and ([pa:t]). Since this vowel is distinct from the realizations of both /ʌ/ (as in : [pat]) and /æ/ (as in : [pɛt]), the distinction between the three phonemes is preserved. Note too that, since the realization of /ɑ:/ in LE has not shifted, there is a clear difference in vowel quality between it and the [a:] realization of /aʊ/.

Realizations of /ɔ:/ and /oʊ/

/ɔ:/ is typically realized as [ɔɐ] in open syllables (as in : [wɔɐ]) and [oʊ] elsewhere (as in : [ʃoʊʔ]). This means that one of the realizations of this phoneme encroaches upon the space of the /oʊ/ phoneme, leading to the possibility of phonemic overlapping. That phoneme in turn does not typically have an [oʊ] realization; rather, it tends to be realized as [ɒʊ] before a tautosyllabic /l/, and [ʌʊ] elsewhere.

Phonemic overlapping: /iː/ vs /ɪ/

The /ɪ/ phoneme is often realized as [ɪi] before a tautosyllabic /l/, as in : [fɪiw].

Similarly, the /iː/ phoneme is often realized as [ɪi] before a tautosyllabic /l/, as in : [fɪiw]. Thus pairs such as these are homophones in LE, but are minimal pairs in RP. Rather than concluding that LE has undergone a phonemic merger, and therefore that this difference is systemic in nature, we will say that it is a matter of phonemic overlapping between the two phonemes, since, in other contexts, /iː/ retains its [iː] realization, and /ɪ/ retains its [ɪ] realization, as in and : [biːʔ] and [biʔ]. Note that, when the /l/ at the end of and is syllabified into a following onset, as in and , the /l/ is no longer in the same syllable as the preceding vowel, which is realized in its normal way, thus: [fiːlɪn] and [fɪːlɪn].

Lowering of word-final schwa

Words such as and have a word-final schwa in RP. In LE, this vowel is often lowered to a central [ɐ] sound: is often pronounced [ˈlɛʔɐ].

LE differs from RP both in terms of the realization of consonant phonemes and, arguably, in terms of the consonant phoneme system.

Voiceless stop phonemes

The voiceless stop phonemes /p/, /t/, /k/ are often realized, before a primary stressed vowel, with heavy aspiration, and in the case of /t/ at least, often with affrication, : [kʰaʔpətsʰɪi]. Note too glottal realizations of these stops, as [ʔ], in a wider range of contexts than in RP. One such context is intervocalically (between vowels), noticeably when the first vowel is stressed, as in : [ˈmɛʔɐ]. Since the first vowel is stressed in such words, the glottaling is foot-internal.

/θ/ vs /f/ and /ð/ vs /v/ in Cockney

It has often been noted that RP minimal pairs such as are homophones for many Cockney speakers, both being [fɪn]. This is referred to informally as 'TH-Fronting' by Wells (1982; see Suggested Further Reading). Whether /θ/ may be said to be absent in Cockney depends very much on whether [θ] is uttered in contexts by Cockney speakers. If one found that [f], rather than [θ], invariably turns up in all other contexts (e.g. between vowels, as in , and word-finally, as in), then one could reasonably conclude that Cockney speakers simply lack the contrast, and that there is a systemic difference here between Cockney and RP. But many speakers are variable with respect to this phenomenon, so we cannot conclude that they lack the /θ/ phoneme. As far as the /v/ vs /ð/ distinction is concerned, it is rather difficult to find many minimal pairs involving the two (vs and are examples). However, it has been noted that words such as and , which have, respectively, intervocalic and word-final /ð/ in RP, are often uttered with [v] in Cockney. It is not clear, however, that word-initial /ð/, as in

, etc., is uttered with [v] in Cockney; if such words are uttered with [ð], then we must say that there is a purely realizational difference here, with Cockney /ð/ realized as [v] intervocalically and word-finally.

/l/ vocalization

The /l/ phoneme is often realized as a voiced alveolar approximant, but it may be overlaid by secondary articulations, such as the velarization we find in the rhyme position of the syllable in RP, and in both the onset and rhyme positions in accents such as Scottish English, GA and Australian English. When the secondary articulation becomes dominant, the alveolar articulation may be lost, resulting in a vowel-like articulation. When lip rounding is added to this articulation, the resulting realization sounds very much like a [w] sound, when in coda position in the syllable. This can be seen in Cockney pronunciations such as ([gɛw]), ([ˈbɛfnəw]) and ([ˈɛwfi]). This is known as ‘l vocalization’ because the /l/ is realized as a vowel-type sound. Cockney l-vocalization seems to be spreading to other towns in the UK.

Absence of /h/ in Cockney?

It has been widely noted that RP minimal pairs such as are homophones in Cockney and perhaps generally in LE. Since Cockney appears to lack even word-internal [h], as in , this looks very much like a systemic difference. Further evidence that /h/ is simply absent in Cockney comes from two sources.

Firstly, we know that [æn] (or [ən]) is the phonetic form of the indefinite article which occurs before vowel-initial words in English, as in , etc., whereas [æ] (or [ə]) is the form which occurs before consonant-initial words, as in

, etc. Cockneys select [æn] (or [ən]) before words such as , etc., which might be taken to suggest that such words are phonologically vowel-initial in Cockney.

Secondly, evidence from the phenomenon of **hyper-correction** is rather telling. When speakers hyper-correct, they 'correct' words (try to make them approximate to what is considered the 'proper' pronunciation found in a prestige accent) which do not require correction. For instance, many speakers who would normally utter [ɪn] rather than [ɪŋ] for the suffix may be nonetheless aware that [ɪŋ], rather than [ɪn], is the 'correct' pronunciation. Such speakers will often 'correct' words such as from [kʰɪkɪn] to [kʰɪkɪŋ], but may also mistakenly 'correct' words such as to [bædmɪŋtən]. The significance of this phenomenon is that the speaker has [ɪn] as his or her phonological form for the - morpheme, and overgeneralizes the 'correction' of his or her realizations to cases which do not even contain the suffix. Similarly, speakers of French, who lack an /h/ phoneme, will hyper-correct their English, resulting in pronunciations such as [hɛə] for both and . The problem for the French speaker is that, in the absence of an /h/ phoneme, she or he is not to know which of her or his /h/-less mental representations should have an /h/ added and which not. Cockney speakers have been observed to behave in just the same way as French speakers, hyper-correcting to [hɛə], to [hɪə] and so on. This strongly suggests that Cockney speakers, like French speakers, simply do not have an /h/ phoneme. However, this phenomenon is perhaps in decline in London English, suggesting that awareness of the /h/ phoneme has always been present among LE speakers.

13.1.2 Tyneside English (Track 13.5 and exercise 2)

By ‘the Tyneside English accent’ (otherwise known as ‘the Geordie accent’), we mean the accent spoken by the natives of the urban areas to the north and south of the last few miles of the River Tyne before it meets the North Sea, including, principally, Newcastle upon Tyne to the north of the river and Gateshead to the south.

/ʊ/ vs /ʌ/

Most Tyneside speakers are typically Northern in having no /ʊ/ vs /ʌ/ distinction: they have the former phoneme, but not the latter.[2] Accents which have this contrast have undergone what is referred to as the FOOT/STRUT split, a historical change in which the /ʊ/ phoneme developed unrounded realizations which eventually gained phonemic status, so that pairs such as (/pʌt/) and (/pʊt/) are minimal pairs. In North of England accents such as Geordie, these are typically homophones, both being pronounced [pʰʊt].

/e/ and /o/

The Tyneside equivalents of RP /eɪ/ and /oʊ/ are /e/ and /o/. The Tyneside /e/ phoneme is realized, in the speech of many Tyneside speakers, as a long monophthong: [e:]. This realization varies with a diphthongal realization ending in schwa: [e:ə]. The Tyneside /o/ phoneme may also be realized as a long monophthong by many speakers: [o:]. For some speakers the realization is a long monophthong which is a fronted [o:], of the [ø:] sort.

/ɔ:/, /ɜ:/ and non-rhoticity

Tyneside is non-rhotic and this has, of course, affected the development of the vowel system. Some Tyneside speakers lack a contrast between /ɔ:/ and /ɜ:/ in certain words, so that pairs such as and are homophones: [wɔ:k]. Speakers with a 'broader' Tyneside accent maintain the distinction between /ɔ:/ and /ɜ:/ in words like and , so that /ɔ:/ is realized as [ɐ:] and /ɜ:/ realized as [ɔ:]: [wɐ:k] ( ) vs [wɔ:k] ( ).

Schwa and non-rhoticity

Among the centring diphthongs, the Tyneside /ɪə/ phoneme is typically realized as [ɪʌ] or [ɪɐ]. The same kind of effect occurs in realizations of /ʊə/, as in : [pʰʊɐ]. The Tyneside /ɛə/ phoneme is typically realized as a long monophthong: [ɛ:].

Related to the Tyneside pronunciations of the centring diphthongs /ɪə/ and /ʊə/, in Tyneside, the schwa phoneme (/ə/) in word-final position can be rather [ʌ]- or [ɐ]-like (i.e. a low central unrounded vowel), but this depends on the history of the word. Where schwa was followed by an /r/ historically, it tends to be [ʌ] or [ɐ], as in .

Low unrounded vowels

The Tyneside /æ/ phoneme is typically realized as [ɐ], but often realized as a long [ɐ:] when it is followed by a voiced word-final consonant, as in ([ʲɐ:d]), but not in ([ʲæs]).

/ɑ:/

Although Tyneside speakers, like many Northern speakers, often have /æ/ rather than /ɑ:/ in some words (e.g.), this reflects neither a systemic nor a realizational difference between Tyneside and RP. It is a lexical-distributional difference: a matter of which of the two phonemes appears in a given word. The words in question are words belonging to what Wells (1982) refers to as the lexical set BATH; these include words such as and , where the vowel is followed by a voiceless fricative, and words such as and , where the vowel is followed by a nasal-plus-consonant cluster.

/aɪ/ realizations

Although subject to variation (both lexical and sociolinguistic), /aɪ/ is often realized as [aɪ] or [eɪ] word-finally and before voiced fricatives, but realized with a more central starting point, as [ʌɪ], elsewhere. The effect is similar to that of the Scottish vowel length generalization on /aɪ/ in Standard Scottish English.

/h/ and /l/ in Tyneside

Although almost every accent of English allows for non-realization of /h/ in unstressed words of a functional category (, etc.), only some allow for non-realization of /h/ in the stressed syllables of words of a lexical category. These accents can be found in many parts of England, but not in Tyneside: /h/ is almost always realized in stressed syllables in Tyneside.

Tyneside /l/ is realized as a ‘clear l’ in all positions, transcribed [lʲ]. The term ‘clear l’, in this context, denotes an alveolar lateral approximant with a secondary palatal articulation, in which the front of the tongue forms an articulation with the hard palate.

Glottal stop and glottalization of /p/, /t/ and /k/

The voiceless stops /p/, /t/ and /k/ often undergo glottalization between vowels, particularly when the first vowel has primary or secondary stress, as in . The resulting realization can be transcribed as [ʔp], [ʔt], [ʔk]. The articulations in the oral cavity occur simultaneously with the glottal closure. The [ʔt] realization of /t/ varies with [ʔɾ]: a glottalized tap. The term ‘glottal reinforcement’ is sometimes used to denote this kind of articulation. Sonorants may intervene between the vowels and the stop, as in

.

In the words cited here, we could define the context for this glottalization as foot-internal (i.e. between a primary or secondary stressed vowel and an unstressed vowel), so that /p/, /t/ and /k/ are aspirated at the beginning of a stressed syllable but glottalized if they occur foot-internally.

Speakers of Tyneside English also exhibit glottaling, in which voiceless stops, especially /t/, are realized as the glottal stop [ʔ]. This is distinct from what we have called glottalization, since the articulation here is a glottal stop with no additional closure in the oral cavity. Glottaling is found in many accents of English, including RP, but it is perceptually more salient if it occurs intervocalically after a stressed vowel, as in words such as , which in Tyneside English can be pronounced with glottalization ([ˈbʊtʔɐ]) or with glottaling ([ˈbʊʔɐ]). Glottaling in this position is socially stigmatized in Britain, but it is perfectly common in Geordie, London English, working-class Scottish English and many other accents of English in the British Isles. RP speakers who claim that glottaling is 'lazy', 'unclear' or 'slovenly' are, amusingly, unaware that they themselves utter glottal stops on a vast scale, but often in phonological contexts in which they are less salient perceptually, as in the case of word-final glottal stops in unstressed function words followed by a word-initial consonant, in sentences like , where the final /t/ of is very likely to be uttered as a glottal stop, even by RP speakers. Ironically, it is precisely because the intervocalic, foot-internal glottal stop in pronunciations of words such as is so clear (i.e. clearly audible) that it is the object of complaints by those who say it is unclear.

The 'r' realizations of /t/

Tyneside, like many North of England accents, has a realization of /t/ which is either an [ɹ]-type or an [ɾ]-type articulation. The phenomenon is known informally as 'T-to-R'. It is not entirely clear whether there is a stateable phonological context in which this occurs, or whether there is simply a stock of words, or even phrases, in which it typically occurs. The phenomenon is probably sociolinguistically variable. The realization is reminiscent of Flapping in GA, in that it seems to occur inter-vocalically, but the Tyneside phenomenon is lexically much more sporadic. Typical cases seem to involve a word-final /t/ preceded by a short vowel and followed by a word beginning with a vowel, as in

. It can, however, occur word-internally, as in , and after long vowels, as in . The 'r' realization also seems to vary with the glottal and glottalized realizations, so that a given pronunciation of, say, can have [t], [ʔ], [ʔt], [ʔɾ] [ɾ] or [ɹ] as the realization of the /t/ in .

13.1.3 Standard Scottish English (Track 13.6 and exercise 3)

Standard Scottish English (SSE) is the standard accent which many Scots speak when speaking the Standard English dialect. It is characteristic of university-educated, middle-class Scottish speakers. It is distinct from what is referred to as Scots, which is derived from the Northumbrian dialect of Old English. Scots is associated with working-class speakers in Scotland; representations of it can be found in the novel , written by Irvine Welsh, and in the film of that name. Many Scottish speakers can speak Standard English without any trace of Scots, but they can also mix Scots words, such as (child), (ear) and (church), into their Standard English.

The Scottish Vowel Length Rule

A major characteristic of SSE is that there is little evidence of phonemic vowel length: pairs such as /u:/ vs /ʊ/, /ɑ:/ vs /æ/ and /ɔ:/ vs /ɒ/ do not form part of the vowel phoneme system. But there is considerable evidence for vowel length. As we have seen (p. 142), some of the vowel phonemes have long allophones morpheme-finally or before voiced continuants, yielding long/short allophonic differences such as [ɫif]/[ɫi:v] ( ) and [hʉf]/[mʉ:v] ( ). This phenomenon is known as the Scottish Vowel Length Rule (SVLR), a major realizational difference between SSE and RP. Which set of vowels undergoes the SVLR is a matter of debate; but /i/, /u/ and /ai/ seem to undergo it for nearly all SSE speakers. The two realizations of /ai/ are [ɐe], as in and , and [ʌi], as in and . While /ai/ has these two allophones in exactly the same contexts as the long and short allophones of /i/ and /u/, the difference between the 'short' ([ʌi]) and 'long' ([ɐe]) allophones seems to be a matter of vowel quality, rather than quantity.

The /ʊ/ vs /u:/ and /ʊ/ vs /ʌ/ distinctions

One of the major systemic differences between RP and SSE is that SSE does not have the (or /u:/ vs /ʊ/) type of distinction. Since SSE does not have phonemically long vowels, there is no /u:/ or /ʊ/; instead, there is a single phoneme: /u/. This is realized as long [ʉ:] in the SVLR contexts, and realized as short [ʉ] elsewhere. In Scottish accents other than SSE, the realization of this vowel can be even more fronted than the high central [ʉ], sometimes approaching a French-type [y] sound (a high rounded vowel).

SSE, unlike many North of England accents, have the (/ʊ/ vs /ʌ/ in RP) type of distinction. The words and are therefore realized as [pʰʉt] (put), [pʰʌt] (putt) and [pʰʉɫ] (pull/pool), with words like realized as [pʰʉːɹ], with a long vowel, triggered by the SVLR.

Absence of the /ɔː/ vs /ɒ/ contrast

Another major systemic difference between RP and SSE lies in the fact that the /ɔː/ vs /ɒ/ contrast is missing in SSE, which has instead a single phoneme: /ɔ/. This means that, for many SSE speakers, RP minimal pairs such as and are homophones. The picture is complicated by the fact that many SSE speakers have borrowed the /ɔː/ vs /ɒ/ contrast from Anglo-English and thus have /ɔː/ in words such as and , but /ɔ/ in words such as and .

RP has the /ɔː/ phoneme in what Wells (1982) refers to as the lexical sets THOUGHT, NORTH and FORCE, but the phoneme /ɒ/ in words of the lexical set LOT. We have seen that SSE speakers whose speech has not been influenced by RP do not have this /ɔː/ vs /ɒ/ contrast, so that (which belongs to the set THOUGHT) and (which belongs to the set LOT) are homophones. Another difference between RP and SSE with respect to these lexical sets is that SSE has the /o/ vowel in words of the set FORCE, but the /ɔ/ vowel in words of the set NORTH. As a result, in SSE there are pairs such as (which belongs to the set NORTH) and (which belongs to the set FORCE) which are minimal pairs:

is pronounced [hɔɹs], while is pronounced [hoɹ; in RP, these are homophones, both being pronounced [hɔ:s].

Absence of the /æ/ vs /ɑ:/ distinction

Another striking systemic difference between RP and SSE lies in the fact that SSE does not have the /æ/ vs /ɑ:/ distinction. Instead, it has a single 'a' phoneme, realized as a low unrounded central vowel, [ɐ], so that pairs such as and are homophones in SSE: [ɐnt]. These are, of course, distinct in RP: is pronounced [ænt] in RP, while is pronounced [ɒ:nt]. Some educated SSE speakers have borrowed this contrast from RP, but are often variable in whether they produce it or not.

The /eɪ/ and /oʊ/-type phonemes

The SSE equivalent of the RP /eɪ/ phoneme is /e/, said by some to be realized as long monophthongal [e:] in the SVLR contexts, and as short [e] elsewhere, as in [bet] ( ) and [be:ɹ] ( ).

The SSE equivalent of the RP /oʊ/ phoneme is /o/, said by some to be realized as long monophthongal [o:] in the SVLR contexts, and short [o] elsewhere, as in [bot] ( ) and [bo:ɹ] ( ).

The diphthongs

As we have seen, the SSE /ai/ diphthong undergoes the SVLR, being realized as [ɐeɪ] in the vowel length contexts and [ʌi] elsewhere. The SSE /au/ diphthong is realized as [ʌʉ], as in : [mʌʉθ].

/ə/, /ʌ/ and /e/

Many words which, in RP, have word-final schwa and did

not historically end in an /ɹ/, such as ,
are uttered with an [ʌ]-type vowel in SSE. Where RP word-final schwa was historically followed by an /ɹ/, SSE retains the /ɹ/ and has either a schwa or an /ɪ/, as in and .

The short [i] vowel found, in RP and many other accents, at the ends of words such as is usually an [e] in SSE. This vowel used to be [ɪ] in RP, but has undergone what Wells (1982) refers to as happY Tensing, in which the [ɪ] has been raised, or tensed, to become [i]. Many varieties of English exhibit happY Tensing, but not SSE.

Rhoticity

SSE is rhotic; the SSE /ɹ/ phoneme is realized in all syllabic positions, including the coda position, so that words like and are always pronounced with an ‘r’. The phoneme is typically realized as [ɹ], sometimes as [ɾ], and very rarely as [r]. Some speakers exhibit an allophonic distinction between the [ɹ] and [ɾ] realizations, with the tap being realized in branching onsets, in words such as and . Note that, when we speak of a rhotic accent, we are not referring to the kind of ‘r’ sound a speaker has: we are denoting an accent in which the ‘r’, whatever its phonetic form might be, is realized in coda position.

The /ʍ/ vs /w/ distinction

The /ʍ/ vs /w/ distinction, as in and , as the spelling suggests, is a distinction which has been largely lost in RP, but is still present in SSE and some

American accents. The /ʍ/ phoneme is realized as a voiceless bilabial fricative, with a secondary velar articulation.

The /h/ vs /x/ vs /k/ distinctions

A major systemic difference between RP and SSE lies in the fact that SSE has retained the phonemic distinction between /k/ and /x/. The /x/ phoneme is realized as a voiceless velar fricative ([x]) in rhymes, after low vowels and back vowels, as in (lake). It occurs in many Scottish place-names, such as and . A fronted allophone ([ç]) may occur after high front vowels, but this tends to be restricted to certain words from Scots which have been incorporated into the speech of SSE speakers; an example is , pronounced [dɹiç] or [dɾiç], a word used to refer to cold, grey and wet weather. As with /ʍ/, the /x/ phoneme has been lost in RP; RP speakers often utter [k] instead in words like . Unlike speakers of some accents in England, SSE speakers do not elide the /h/ phoneme before stressed vowels.

‘Dark l’

The SSE /l/ phoneme is realized as a ‘dark l’, i.e. [ɫ], in all contexts. Recall that ‘dark l’ is an informal term for an alveolar lateral approximant with a secondary articulation of velarization. Recall too that the ‘dark l’ realization only occurs in the rhyme of the syllable in RP, so that the realizations of the two /l/ phonemes in a word such as are distinct in RP: [lʌɫ]. In SSE, this word is pronounced [ɫʌɫ].

13.2 Two American Accents

13.2.1 New York City English

The New York City English accent is fairly sharply defined in geographical terms, being largely confined to the boroughs of New York City (henceforth: New York). There is, however, considerable socially determined variation within New York, and this variation has been the subject of a good deal of sociolinguistic study. The New York accent is widely recognized in the United States and, like many urban accents, evokes mainly negative reactions. One of the questions addressed in the sociolinguistic studies conducted in New York is that of rhoticity. It seems clear that the accent has made the historical transition from rhotic to non-rhotic, since it has (most of) the set of centring diphthongs ending in schwa which are characteristic of nonrhotic accents. However, there is considerable sociolinguistic variation with respect to rhoticity, and it appears that [ɹ] in coda position is staging a comeback. Recall that the standard accent in the United States, General American, is rhotic, and that rhoticity is regarded as more prestigious than non-rhoticity.

The [ɜɪ] vowel

The [ɜɪ] realization of the /ɜː/ phoneme, as in [hɜɪd] (   ), is widely regarded as characteristic of New York speech, and is often said to characterize 'the Brooklyn accent' (although, as we noted earlier, it is by no means restricted to Brooklyn). However, it is highly stigmatized and is probably dying out. This realization occurs before a coda consonant, so that it does not occur in non-rhotic pronunciations of . Some speakers also have an [ɜɪ] realization of the /ɔɪ/ phoneme, also before a coda consonant, so that some minimal pairs have become homophones, as in and .

Allophones of the /æ/ phoneme

There are variable [ɛ:ə] and [æ:] realizations of /æ/ in certain environments, namely before a voiceless fricative, voiced stop or nasal when they occur in a word-final coda (although /ŋ/ behaves variably), as in and , but not in , with a final /l/, or , with a final voiceless stop. These diphthongal or long realizations are sometimes referred to as 'tense' realizations of the phoneme. It is not clear why these specific consonants in that position should trigger this tensing process.[3] There is considerable sociolinguistic variation in the exact phonetic form of the allophones, and, for some speakers, the [ɛ:ə] realization merges with realizations of the /ɛə/ phoneme, collapsing minimal pairs such as .

Realizations of /ɔ/

New York speakers often have [ɔə] and [oə] realizations of the /ɔ/ phoneme, as in [pʰɔə] ( ), and, with an even higher starting point, [ʊə], thus creating the possibility of partial merger with the /ʊə/ phoneme.

Rhoticity and non-rhoticity

The discussion above shows that New York speech has undergone the transition from rhoticity to non-rhoticity, and is reverting, for many speakers, to rhoticity. This phenomenon may well be resulting in greater occurrence of intrusive [ɹ], since a speaker who has been non-rhotic but is attempting to be rhotic may well insert intrusive [ɹ]s as part of a general strategy to utter [ɹ] where it may otherwise be absent.

Realizations of /θ/ and /ð/

The phonemes /θ/ and /ð/ are often realized as either affricates ([t̪θ] and [d̪ð]) or dental stops: [t̪] and [d̪]; the variation is sociolinguistically determined. Note that, for speakers who have dental stops, the phonemic distinction between alveolar and dental stops is maintained, as in [t̪ʰɪn] (   ) vs [tʰɪn] (   ). To many speakers of other varieties of English, the distinction may well be difficult to notice.

Realizations of /t/

New York speech, like General American, has Flapping of /t/ in intervocalic environments (where the first vowel is stressed), but it also has glottal stop realizations of /t/ in coda position on a greater scale than in GA. The /t/ phoneme is often heavily aspirated, to the point, at times, of being affricated, in syllable-initial position, as in [tsɪn] (   ).

13.2.2 Texan English (Track 13.7 and exercise 4)

Texas is a vast state, said to be larger than France. It is therefore unsurprising that there is much variability within Texan English. Many educated Texans speak both a Texan variety of English and General American, depending on the context of utterance. Texan English is rhotic in many parts of the state, but, as one approaches the eastern border with Louisiana, variable non-rhoticity can be found. Whether varieties of Texan English can be considered Southern varieties of US English depends on what one considers to constitute the linguistic American South: some dialect maps exclude Texas from the South, while others include eastern parts of the state. Many of the pronunciation features one encounters in Texas can be found in neighbouring states, but there are some pronunciation features that are said to be specifically Texan.

The /aɪ/ phoneme of the lexical set PRICE is often monophthongized in Texan speech. The phenomenon, referred to informally by Wells (1982) as PRICE smoothing, results in a long monophthong, as in pronounced [p^{h}ɹa:s]. PRICE smoothing can be heard in neighbouring states.

Monophthongization can also be heard in the Texan realization of the /ɔɪ/ phoneme in words of the lexical set CHOICE, so that words such as are pronounced [ɔ:l], though the phoneme is sometimes pronounced with a schwa off-glide, as in [ɔəl].

In addition to monophthongization, in Texas and in many Southern states, there is diphthongization of the phonemes /ɪ/, /ɛ/ and /æ/, in words of the lexical sets KIT, DRESS and TRAP. The diphthongs in question have a schwa off-glide, so that is pronounced [kʰɪət], is pronounced [dɹɛəs] and is pronounced [təɹæəp]. One can, at times, hear pronunciations of words of the TRAP set with what is either a triphthong or a bisyllabic pronunciation, as in [bæjənd] for . If such pronunciations have two syllables, then we might argue that they contain an [æj] diphthong followed by a schwa in the second syllable.

The /ɛ/ phoneme in words of the lexical set DRESS also undergoes diphthongization to [ei] before the phonemes /ʃ/, /ʒ/ and /ŋ/, so that the word , pronounced [speiʃəl], sounds just like the word .

The phonemic distinction between the /ɛ/ of the lexical set DRESS and the /ɪ/ of the lexical set KIT is neutralized before coda /n/ in the speech of many Texans, so that pairs such as and are both pronounced as [kʰɪn]. Neutralization is, as we have seen, defined as the suspension of a phonemic contrast in specific phono-logical environments (here, before coda /n/): it is not the case that speakers who exhibit this neutralization have lost the /ɛ/ vs /ɪ/ contrast altogether.

So, the /ɛ/ phoneme has a wide variety of realizations among speakers of Texan English, from [ɛə], to [ei], to [ɑ].

The /aʊ/ diphthong in words of the set MOTH often has a higher starting point, in Texas and in neighbouring states, so that the word is pronounced [mɛʊθ].

The phoneme in words of the lexical set LOT often has a diphthongal realization in Texan English, so that the word is pronounced [doʊg], often written informally as . Given this and the MOUTH vowel, a compound such as (dogs for herding cows) can be pronounced as [k^{h}ɛʊdoʊgz].

The /uː/ vs /ʊ/ distinction is often neutralized before coda /l/ in Texan English, so that and are both pronounced [p^{h}ʊl].

The consonants of Texan English are broadly parallel to those of General American, featuring ‘dark l’ in all contexts, Flapping and rhoticity. However, as one approaches the border with Louisiana, variable non-rhoticity can be observed, since the Southern accents to the east of Texas are largely non-rhotic.

We saw that SSE speakers have retained the /w/ vs /ʍ/ distinction in pairs such as and . This can also be attested among some speakers of Texan English.

13.3 Two Southern Hemisphere Accents

13.3.1 Australian English (Track 13.8 and exercise 5)

Descriptions of Australian English often distinguish between three socially defined varieties: Cultivated, General and Broad Australian. We do not examine the differences between these, which mostly concern vowel articulations. We will, however, give a brief overview of General Australian, which is spoken throughout Australia. General Australian English pronunciation has its origins in the speech of early nineteenth-century working-class speakers from the South East of England, and is therefore similar to present-day London English in some respects.

The main characterizing properties of General Australian are to be found in vowel articulations. The vowel system of General Australian is parallel to that of RP, but with many major realizational differences, mostly of a ‘vowel shift’ nature, as we noted earlier in chapter 12.

The /i:/, /eɪ/, /aɪ/, /ɔɪ/ vowel shift

We described this vowel shift in chapter 12. It can be depicted as follows:

/i:/	/eɪ/	/aɪ/	/ɔɪ/
[ɪi]	[ɐɪ]	[ɒɪ]	[ɔɪ]

The /u:/, /oʊ/, /aʊ/ vowel shift

Like the high front unrounded phoneme /i:/, the high back rounded vowel phoneme /u:/ has diphthongized, often resulting in a diphthong with a high front unrounded starting point and a high back unrounded finishing point; we will transcribe this as [ɪɯ]. This realization potentially encroaches on the space of the [əʊ]-type realizations of /oʊ/, which have shifted to [ɐʉ], with a low central unrounded starting point and a high central rounded end point, thus entering the space of /aʊ/, whose realization has shifted to [æo], in which the starting point is more front than that of [ɐʉ], and the end point lower and further back. This set of shifts can be depicted as follows:

/uː/	/oʊ/	/aʊ/
|	|	|
[ɨɯ]	[ɐʉ]	[æo]

The /ʌ/, /æ/, /ɛ/, /ɪ/ vowel shift

A vowel shift has also affected the short vowels /ʌ/, /æ/, /ɛ/ and /ɪ/, with /ʌ/ being realized as a low front articulation in the [a] area, close to the space of the /æ/ phoneme, which is realized as [ɛ]-like. In turn, /ɛ/ is realized as [e]-like, and thus close to the space of /ɪ/, which is rather [i]-like. This articulation is, of course, distinct from that of /i:/, which, as we have seen, is diphthongal. We may depict this set of vowel shifts as follows:

/ʌ/	/æ/	/ɛ/	/ɪ/
|	|	|	|
[a]	[ɛ]	[e]	[i]

The /ɑ:/ vs /ʌ/ distinction

The General Australian realization of /ɑ:/, like that of /ʌ/, is also fronted, but the distinction between the two is not merged, since there is a length difference between them, as in [P^{h}at] (   ) vs [P^{h}a:t] (   ).

We have already noted that General Australian, like SSE, has ‘dark l’ in all positions; the precise nature of the ‘darkness’ may entail, in both cases, retraction and lowering of the tongue body, rather than velarization as such.

General Australian is non-rhotic. It also has a process rather similar to that of Flapping in North American English: a /t/ will often be realized as a voiced articulation between vowels.

13.3.2 Indian English (Track 13.9 and exercise 6)

There are many languages spoken on the Indian subcontinent. They are divided into two language families. The languages of the North belong to the Indo-European language family. This is a vast family of historically related languages. In the nineteenth century, linguists were able to show that Sanskrit, an ancient language of the North of India, was ultimately related to Ancient Greek and to Latin. From this discovery, it was possible to show that the languages of the North of India were related to many of the languages spoken in Europe, such as the Romance languages (including French, Spanish, Portuguese and Italian) and the Germanic languages (including English, German, Dutch and Swedish). The Indo-European languages of the North of India include Hindi, Marathi, Gujerati and Punjabi. It is hard to believe, when listening to those languages, that they are historically related to English, but they are; it is has to be borne in mind that the time-span over which these languages have evolved is vast. The languages spoken in the South of India belong to an entirely distinct language family, known as the Dravidian language family. These include Tamil, Malayalam, Kannada, Telugu and Toda. The languages of India, whether Indo-European or Dravidian, are not mutually comprehensible. English has therefore taken on the status of a **lingua franca** in India: a language which can be used as a form of communication between people whose native languages are not mutually comprehensible. Many educated Indians have English as a second language, and thus use English as a lingua franca.

The vowels of the sets FACE and GOAT, which are diphthongs in RP and GA, are typically realized as the monophthongs [e] and [o] in Indian English, so that FACE is [fes] and GOAT is [got].

Many, but not all, speakers of Indian English are rhotic, and this has consequences for the vowel system: as we have seen, there are no centring diphthongs of the sort /ɪə/, /ɛə/, /ʊə/ in rhotic accents of English. The /ɜː/ vowel in words of the lexical set NURSE is often not present in either rhotic or non-rhotic varieties of Indian English. Even among non-rhotic speakers of Indian English, a vowel of the sort [ʌ], of the lexical set STRUT, can be found in words of the NURSE set, so that words such as (an English word of French origin, denoting a grape used to produce a red wine of that name), pronounced [ˈmɜːləʊ] in RP, is pronounced [ˈmΛlo].

The RP distinction between the long vowel /ɔː/ (of the lexical sets THOUGHT, NORTH and FORCE) and the short vowel /ɔː/ (of the lexical set LOT) is often absent in varieties of Indian English. RP minimal pairs such as ([ˈɔːfəɫ]) and ([ˈɒfəɫ]) are therefore homophones in the speech of many speakers of Indian English, both being pronounced [ˈɔfəɫ], with a short [ɔ], as in the word-play utterance ‘There’s an awful lot of offal being thrown away in British kitchens.’

Indian English has happY Tensing, so that the word is pronounced [hæpi], as in RP.

Some of the Indo-European languages of the North of India have a four-way phonemic contrast between voiceless unaspirated stops, voiceless aspirated stops, voiced stops and breathy voiced stops. We have already encountered voiceless unaspirated stops in English words such as and . We have also encountered aspirated voiceless stops in English words such as the [pʰ] in ; the initial stops in the words and are also aspirated. Voiced stops occur in words such as and . We have not yet encountered breathy voiced stops. In voiced sounds, the vocal cords are vibrating close together. In breathy voiced stops, the vocal cords are further apart, but still vibrate because of increased airflow through the glottis. The bilabial, alveolar and velar breathy voiced stops are transcribed as [b̤], [d̤] and [g̤]. They occur in Hindi words which feature in Indian cookery, such as (okra), (coriander seeds) and (clarified butter). Examples of minimal pairs featuring this four-way contrast are [bal] (‘hair’) vs [pal] (‘take care of’) vs [p̤al] (‘knife blade’) vs [b̤al] (‘forehead’). You can hear these, and minimal pairs at other places of articulation, on Track 13.1.

It is common for speakers of Indian English to produce words such as and with unaspirated voiceless stops in initial position. On the face of it, this is puzzling, since many speakers of Indian English have a series of aspirated voiceless stops in their native Indian language: why not just use the aspirated voiceless stops of one's native language when uttering words like and ? Some insight into this phenomenon may be gained if we consider that many of the native Indian languages lack dental fricatives; speakers of Indian English therefore engage in 'TH-stopping', in which they produce aspirated voiceless dental [t̪] for /θ/ and the breathy voiced dental stop [d̤] for /ð/. Perhaps it is because a member of the aspirated voiceless stop series ([tʰ]) is being used for /θ/ that the unaspirated series is used in words which would otherwise have an aspirated voiceless stop in RP.

Many of the native languages of India have a series of retroflex stops. Retroflex sounds are produced with the tip and blade of the tongue curled backwards, so that the underside of the tongue forms an articulation with the alveolar ridge. Stops formed this way include the voiceless retroflex stop [ʈ] and the voiced retroflex stop [ʈ]. Speakers of Indian English often realize /t/ and /d/ in this retroflex manner; this can often be heard in the pronunciation of waiters in Indian restaurants, who will often produce words such as (an Indian flatbread) with a retroflex [ʈ], and (a thin, crispy Indian bread) with a retroflex [ɖ]. You can hear this on Track 13.2.

The /l/ phoneme is not realized as a 'dark' (velarized) lateral in most Indian English. The /r/ phoneme is often realized as an alveolar tap, so that the word is pronounced as [ɾap]. The phonemes /v/ and /w/ are not always consistently distinguished by speakers of Indian English. Many of the native languages of India have an approximant which is intermediate between the approximant [w] and the fricative [v]. The sound they have is a labio-dental approximant, in which the lower lip forms a stricture of open approximation with the upper teeth. This sound is transcribed as [ʋ]. It is like a [w] in that it is an approximant, and it is like a [v] in that it is labio-dental. You can hear the contrast between [v], [ʋ] and [w] (illustrated from the African language Isoko) on Track 13.3. Speakers of Indian English, because they lack a /v/ vs /w/ distinction, will often produce [w] for [v], or a [ʋ] for [v].

Many of the native languages of India are not stress-timed, so Indian English is often spoken without the stress timing that is typical of so many varieties of English. Linked to this is the fact that the word stress often falls on the wrong syllable in Indian English.

13.4 An Overview of Some Common Phenomena Found in Accent Variation

It is clear from our brief description of the above varieties of spoken English that the speech of any speech community, and indeed of a given speaker within a community, is typically variable. It is such variation that can eventually lead to divergence between the speech of different communities over time. Several different factors are involved in such variation. A host of sociolinguistic factors is relevant. These may include the sex, age or social class of the speaker, the speaker's social aspirations, the structure of the society in which the speaker lives, and complex aspects of the social networks in which the speaker lives, involving such things as solidarity, group identity and individual identity. We have not examined these factors here, but we note that an understanding of them is vital if we are to gain a comprehensive understanding of variation in pronunciation in a given speech community.

Variation in pronunciation is also constrained by factors other than societal ones, such as the nature of the vocal tract, the relationship between the phonemes within a phoneme system, and factors to do with the relative perceptual salience of sounds, depending on what sounds they are preceded or followed by, where they occur within syllable structure, and whether they occur in stressed or unstressed syllables. The following is a summing up of the sorts of phenomena we have discussed in describing accent variation.

13.4.1 Vowel Phenomena

Diphthongal realizations of vowel phonemes may be triggered by an adjacent consonant, as in the [iːə] realization of /iː/ before 'dark l', or may occur 'spontaneously'; we have seen many examples of the latter in General American, New York City English, General Australian, London English, RP and Tyneside English. Similarly, RP used to have monophthongal realizations of the mid vowels /e/ and /o/, but these have undergone diphthongization in the history of RP.

Diphthongization of these mid vowels has also taken place in the history of GA, but not in SSE.

We have seen many examples of monophthongal realizations of diphthongs, as in [ɛ:] instead of [ɛə] and [ɔ:] instead of [ʊə] in RP.

As we have seen, phonemes may have allophones, but those allophones may gain phonemic status if a pattern of complementary distribution is disrupted, resulting in a newly emergent pattern of parallel distribution. Examples include the emergence of the /ə/-ending centring diphthong phonemes (/ɪə/, /ɛə/, /ʊə/), and the /ʊ/ vs /ʌ/ distinction (the FOOT/STRUT split) in RP and other accents of English.

Since articulations may shift, it is possible for the realizations of one phoneme to merge with those of another, resulting in the loss of a phonemic distinction. Examples are the mergers of the /k/ vs /x/ and /ʍ/ vs /w/ contrasts in many accents of English, resulting in words which were previously minimal pairs becoming homophones, as in and .

When the realization of one phoneme encroaches on the realization of another vowel phoneme, ‘evasive action’ may be taken, so that the phonemic contrast is sustained and merger is avoided. We have seen examples of this in both London English and Australian English.

Vowels frequently reduce to either a weak form (e.g. /iː/ → [i] and /uː/ → [u]), or to schwa, in unstressed syllables which, as we have seen, are perceptually less salient. This phenomenon occurs in almost every variety of English.

13.4.2 Consonantal Phenomena

A common phenomenon is intervocalic weakening, in which a consonant articulation becomes more vowel-like, in the sense of becoming voiced, or undergoing a diminution in degree or duration of stricture. Flapping in North American English, voicing of /t/ in Australian English and 'T-to-R' in Tyneside English are all examples of this phenomenon.

In coda position, consonants often undergo weakening in the form of reduction in degree of stricture, sometimes leading to complete elision. Examples are the erosion and eventual loss of coda [ɹ] in non-rhotic accents, and the reduction of voiceless stops to glottal stops, which occurs to some extent in all of the accents we have considered. Vocalization of coda /l/, which occurs in London English, is another such process.

This phenomenon, in which /p/, /t/ and /k/ may be realized as [pΦ]/[pf], [tθ]/[ts] and [kx] respectively, appears to be connected with strong aspiration. It has been attested in New York City, Liverpool and London.

The principal point to be borne in mind about such processes is that they are rarely limited to specific accents of English. Nor are they limited to English. Because they arise from the nature of the human vocal tract, human perceptual capacities and the structure of human language phonologies, they are widely attested across the world's languages.

Notes

1 There is often quite marked nasalization of the vowel in cases like these, where a nasal follows an open vowel.

2 Some Tyneside speakers do in fact have a contrast, between [ʊ] and an unrounded, sometimes centralized, version of [ʊ], which we might transcribe as [ɨ].

3 It has often been noted that auxiliary verbs which end with a tensing coda consonant, such as and , nevertheless do not undergo tensing.

Exercises

References for the exercises are given at the end of the chapter.

1 London English

Listen to Track 13.4. This is a sound file from the IViE corpus (Intonational Variation in English: see Grabe, Nolan and Farrar 1998 for details and further references), which contains recordings of UK speakers from London, Cambridge, Bradford, Leeds, Liverpool, Newcastle, Cardiff, Northern Ireland (Belfast) and the Republic of Ireland (Dublin). You can access the corpus at the following website: www.phon.ox.ac.uk/files/apps/IViE. Alternatively, enter 'IViE' on a search engine. The speaker is reading the first two paragraphs of the following passage:

> Once upon a time, there was a girl called Cinderella, but everyone called her Cinders. Cinders lived with her mother and two stepsisters called Lilly and Rosa. Lilly and Rosa were very unfriendly and they were lazy girls. They spent all their time buying new clothes and going to parties. Poor Cinders had to wear all their old hand-me-downs! And she had to do the cleaning!
>
> One day, a royal messenger came to announce a ball. The ball would be held at the Royal Palace, in honour of the queen's only son, Prince William. Lilly and Rosa thought this was divine. Prince William was gorgeous, and he was looking for a bride! They dreamed of wedding bells!
>
> When the evening of the ball arrived, Cinders had to help her sisters get ready. They were in a bad mood. They'd wanted to buy some new gowns, but their mother said that they had enough gowns. So they started shouting at Cinders. 'Find my jewels!' yelled one. 'Find my hat!' howled the other. They wanted hairbrushes, hairpins and hairspray.

(a) Is the speaker consistently non-rhotic?

(b) In words which would have word-final schwa in

RP (such as and), what vowels do you find with this speaker?

(c) Is there any evidence of 'TH-fronting' (the articulation of the dental fricative phonemes /θ/ and /ð/ as labio-dental fricatives [f] and [v])?

(d) This speaker exhibits /l/ vocalization. We have said that this occurs in the rhyme of the syllable in London English, but not in the coda. Is this true for this speaker?

(e) In what contexts does this speaker produce glottal stop realizations of the /t/ phoneme?

2 Tyneside English (Newcastle)

Listen to Track 13.5. This is also an IViE file. The speaker is reading all three paragraphs of the Cinderella passage in exercise 1.

(a) North of England accents are said to lack the FOOT/STRUT split (i.e. the phonemic distinction between /ʊ/ and /ʌ/). Is this the case for this speaker? Provide examples of relevant words on the sound track.

(b) In words written with <-er(s)> at the end, such as and , what vowel does this speaker utter in that final unstressed syllable? Is this what you would expect from a Geordie speaker?

(c) Like many speakers of Tyneside English, the speaker has two realizations of the /aɪ/ diphthong. Can you transcribe them and give examples of words on the recording which contain them?

(d) In words such as , does the speaker have a monophthongal or a diphthongal realization of the vowel?

(e) Geordie speakers are said to have 'clear' (palatalized) [lʲ] in all syllabic contexts. Is this true for this speaker?

(f) Cite examples of glottaling in the speech of this speaker. The speaker realizes /d/ as a glottal stop in one word: can you identify which word this is?

(g) We said that the centring diphthongs /ɪə/ and /ʊə/ are often realized as /ɪɐ/ and /ʊɐ/ in Tyneside English. Can you hear any examples of this on the recording?

3 Standard Scottish English (Glasgow)

Listen to [Track 13.6](). This is a recording from the PAC project (La Phonologie de l'anglais contemporain; see Carr, Durand and Pukli 2004 for details). The PAC website is at: http://www.projet-pac.net/

The speaker is reading the following PAC word list:

1. pit
2. pet
3. pat
4. pot
5. put
6. putt
7. sea
8. say
9. sigh

10. sue
11. stir
12. steer
13. stairs
14. err
15. far
16. war
17. more
18. purr
19. moor
20. feel
21. fill
22. fell
23. fall
24. full
25. fool
26. fail
27. foal
28. file
29. foul
30. foil
31. furl
32. bird
33. bard
34. beard
35. bared
36. board
37. barred

38. bored
39. bode
40. bowed
41. bead
42. bid
43. bed
44. bad
45. bard
46. pant
47. plant
48. master
49. afterwards
50. ants
51. aunts
52. dance
53. farther
54. father
55. row
56. rose
57. rows
58. pore
59. poor
60. pour
61. paw
62. paws
63. pause
64. pose
65. wait

66. weight
67. side
68. sighed
69. agreed
70. greed
71. brood
72. brewed
73. fir
74. fair
75. fur
76. four
77. fore
78. for
79. nose
80. knows
81. cot
82. caught
83. meat
84. meet
85. mate
86. naught
87. knot
88. doll
89. dole
90. fierce
91. bird
92. scarce
93. pert

94. start
95. horse
96. hoarse
97. word
98. gourd
99. short
100. sport
101. next
102. vexed
103. leopard
104. shepherd
105. here
106. there
107. weary
108. spirit
109. marry
110. Mary
111. merry
112. sorry
113. story
114. hurry
115. jury
116. bury
117. berry
118. heaven
119. leaven
120. earth
121. berth

122. cook
123. soot
124. look
125. room
126. pearl
127. peril

(a) SSE is said to be rhotic, but there are speakers who are variably non-rhotic. Is this speaker consistently rhotic?

(b) SSE is said not to have a contrast between long /uː/ and short /ʊ/, of the sort found in RP. Instead, SSE is said to have a single /u/ phoneme, realized as the high central rounded vowel [ʉ]. Is this true for the speaker on the recording?

(c) Some SSE speakers have only a two-way contrast in words such as and , with and having /ʌ/, having /ɛ/. Others have a three-way contrast, with /ɪ/ in , /ʌ/ in and /ɛ/ in . Does this speaker have a two-way contrast in these ‘pre-r’ positions, or a three-way contrast?

(d) SSE is said to possess a contrast between short /o/, as in , and short /ɔ/, as in . It is said not to have a contrast, of the sort found in RP, between words of the lexical set LOT (which have a short /ɔ:/ in RP) and words of the sets THOUGHT and NORTH (with a long /ɔ:/ in RP). Does this speaker exhibit any length difference between pairs such as / and / ?

(e) The speaker is typically SSE in that she has two realizations of the /ai/ diphthong: [ɐe] and [ʌi] Can you identify words containing these?

(f) SSE speakers are said to have monophthongal /e/ and /o/, distinct from the RP diphthongs /eɪ/ and /oʊ/. Is this true for this speaker?

(g) SSE speakers are said to have 'dark l' ([ɫ]) in all syllabic positions. Is this true for this speaker?

(h) The speaker has both the approximant [ɹ] and the tap [ɾ]. Identify words containing these.

4 West Texan English (Lubbock)

Listen to Track 13.7. This too is a PAC recording; the speaker is reading out the following written passage:

> I saw Hoyt on the news the other day. I can't be sure whether I got the facts right. But I think he wants Lubbock farmers to plant soy, rye, and maybe rice for the new crop. I tuned in to the farm report last night. They say cotton prices will continue to rise. It sure is good that rot didn't set in from all the rain. I tell you, the morning dew wasn't much help either. I still think we will get back all our investment this year. Getting in the cotton crop this year was a challenge. I had my oldest boy at work with us since he isn't in school anymore. He likes to work on the farm. But I feel he should work in a business like oil, cattle ranching, or maybe nothing to do with agriculture. He just turned nineteen. You know, he doesn't play with toys like tin soldiers or his bow and arrow. He's almost a man, and quite a fine one at that. There's this story about my boy

that I still do like to tell. It’s a tale that really shows how he turned everything all around. My son Roy was about ten years old at the time. His friend Tom came over to play. They decided to steal my neighbor’s dune buggy and to go for a joy ride in it. They went back behind this cotton field onto a horse trail. Well, that dune buggy filled up, with first one kid and then another. Tom was at the wheel driving this contraption. They didn’t drive too far when they hit a big brass nail, from the railroad or something. The tire went flat, they hit a holding pen and then they all fell head first into ashes and dried mud when the buggy crashed. Those kids didn’t have good sense at all about what they were doing. My neighbor Ken had to sell the dune buggy at a loss after that. I know I sent Roy to bed without dinner, and that was the least of his punishments. I told my son he had a duty to pay back the damages. He toiled all summer in Ken’s tack room so he could do right by him. My boy gave up swimming at the community pool that whole hot summer. He used the cash to pay back Ken. Did Roy pull something like that again? No he did not. It’s been nine years since that happened. He learned to choose his friends and his actions more carefully. I never used a lash or hit him, but he learned to toe the line and not to steal or lie ever again. And the law left him alone because he paid Ken back. He’s not a bad boy, my Roy, and he didn’t fail to make good choices other times. Just once he thought he’d like to take out a hot rod one day and got lucky that his mistake didn’t ruin his life. Roy has now become a good role model for Beau, my middle son, and for Luke, the youngest.

(a) Is the speaker rhotic or non-rhotic? Cite examples from the recording.

(b) This speaker has variable ‘smoothing’ (monophthongization) of the /aɪ/ diphthong (in words of the lexical set PRICE), resulting in an [a:] pronunciation. Give examples of the variability.

(c) The speaker also has variable ‘smoothing’ of the /ɔɪ/ diphthong (in words of the lexical set CHOICE), resulting in an [ɔ:] pronunciation. Give examples.

(d) Can you hear any examples of Southern Breaking (diphthongization) in the vowels /æ/, /ɛ/ and /ɪ/?

(e) Many varieties of Texan English are said to exhibit (variably) neutralization of the /ɪ/ vs /ɛ/ distinction before coda /n/. Is this true for this speaker?

5 Australian English

Listen to Track 13.8. This is a PAC recording. The speaker is reading the same word list as the SSE speaker in exercise 3.

(a) Australian English is said to be non-rhotic. Is this true for this speaker?

(b) Australian English is said to have ‘dark l’ in all syllabic positions. Is this true for this speaker?

(c) Australian English is said to have Flapping for intervocalic /t/. Is this present on the recording? Is it variable?

(d) In our discussion of Australian English, we said that the vowel phoneme system was the same as that of RP, but that there were vowel shift phenomena. Are any of these present on the recording?

6 Indian English (Mumbai/Bombay)

Listen to Track 13.9. This is a PAC recording. See word list 2 on the PAC website.

(a) The speaker is variably non-rhotic. Give examples which demonstrate both rhotic and non-rhotic pronunciations.

(b) Cite examples of TH-stopping on the recording.

(c) Is there aspiration of the word-initial /p/ phonemes in words 1-6?

(d) Does the speaker have ‘clear’ (palatalized) realizations of the /l/ phoneme?

(e) How does the speaker realize the /r/ phoneme?

7 Comparison of the accents in exercises 1–6 with RP and GA

Listen to Tracks 13.10 and 13.11. These are PAC recordings of an RP and a GA speaker. They are reading the following passage from the PAC project:

Christmas interview of a television evangelist

If television evangelists are anything like the rest of us, all they really want to do in Christmas week is snap at their families, criticize their friends and make their neighbours’ children cry by glaring at them over the garden fence. Yet society expects them to be as jovial and beaming as they are for the other fifty-one weeks of the year. If anything, more so.

Take the Reverend Peter ‘Pete’ Smith, the ‘TV vicar’ who sends out press releases in which he describes himself as ‘the man who has captured the spirit of the age'. Before our 9 a.m. meeting at his 'media office' on Crawshaw Avenue, South London, he faced, he says, a real dilemma. Should he make an effort 'to behave like a Christian' – throw his door open, offer me a cup of tea – or should he just play it cool, study his fingernails in a manner that showed bored indifference and get rid of me as quickly as possible? In the end, he did neither.

'As a matter of fact, John', he says in a loud Estuary English twang, 'St Francis said, "At all times preach the gospel and speak whenever you have to." ' But hey, he didn't mean "Be on your best behaviour and be happy all the time." I could have been extra-polite to you, but the real me would have come out as I was talking. You cannot disguise what you are.'

'And what are you then, Pete?'

'Well, I'm a Christian, John. I've been one since I was 14. And I know for sure that Christianity will be judged more on who you are rather than what you have to say about it. Many church leaders don't appear to understand this. They think we can only be really Christian when we are ramming the doctrine of the Creation down people's throats. But if you try to force-feed people they get sick of it and think you're a pain. It's seen as the job of a Christian leader to wear a dog-collar and dress in purple and always be talking about the real meaning of the New Testament. In reality, that turns people right off!'

In many ways, 'Pete' Smith looks exactly how you'd expect a high-profile, born-again Christian to look: tall, handsome, clean-cut and evenly sun-tanned. He has those scarily white teeth that TV evangelists tend to have, and he doesn't wear a dogcollar. In fact, when doing his various religious programmes on Sunday mornings, he has been known to wear a black leather jacket instead, in casual mode. Today, the look is more business-like: metal-rimmed glasses, a grey suit, a blue open-neck shirt, and fashionable black shoes with large buckles. Smith is 44 but he looks a mere 24.

During the whole interview, there wasn't any talk of the poor or the needy but only of his forthcoming trip to China in February and the masses waiting for his message there. I ventured a few questions relating to the charity trust he founded some ten years ago and which, it is generally agreed, employs eight hundred staff and runs schools, hospitals and hostels around the world. And what about the gambling organization he has been willing to advise? Is that a temporary activity or might it be true that he has accepted to be paid to sit on its Board of Directors? Which side is religion on these days? Does money matter? It was as if I had launched a few missiles in his direction. He just sighed in answer: 'I'm only human, John. God knows I do my best and often fail, But it's no skin off my nose if our enemies sneer at some of the good work we do. Truth will out.'

(a) Listen to the realizations of the /r/ phoneme and the /l/ phoneme in all of the accents in exercises 1–7. What differences do you find?

(b) Listen to the vowel in words of the lexical set LOT (e.g.). What differences do you find? Many American speakers lack the rounded vowel [ɒ] in words of the lexical set LOT, in words such as and . Instead, they are said to have an unrounded [ɑ] vowel, so that words such as sound like to British hearers. Is this true for the GA speaker?

(c) Which accents exhibit 'happY Tensing'? Recall that this is the short, tense [i] pronunciation in the final, unstressed, syllable of words such as , , .

(d) Listen to words which have word-final schwa in RP (e.g.). What vowels do you find in such words among the various speakers?

(e) Among the non-rhotic accents, RP is said to have the centring diphthongs /ɪə/, /ɛə/ and /ʊə/, in words of the lexical sets NEAR, SQUARE and CURE. What realizations of these phonemes do you find among the non-rhotic speakers?

References

The references to the IViE and PAC projects are:

Carr, P., Durand, J. and Pukli, M. (2004). 'The PAC project: principles and methods'. 36: 24–35. A pdf version of this article can be found under 'publications' on the PAC website: http://www.projet-pac.net/

Grabe, E., Nolan, F. and Farrar, K. (1998). 'IViE: a comparative transcription system for intonational variation in

English'.

. Sydney, Australia.

For more on the PAC project, see:

Przewozny, A., Moore, S. and Turcsan, G. (eds.) (2012).

,

CLLE-ERSS, Université Toulouse II-Le Mirail.

Suggested Further Reading

For a more detailed account of articulatory phonetics, see D. Abercrombie (1967) , Edinburgh: Edinburgh University Press, and J. C. Catford (1988) , Oxford: Clarendon Press. An introduction to phonetics which concentrates mostly on English is J. D. O'Connor (1973) , London: Penguin. For an introduction to phonetics geared towards English, and with a good introductory coverage of acoustic phonetics, see P. Ladefoged (2010) (sixth edition, with Keith Johnson), New York: Harcourt Brace Jovanovich. For standard descriptions of the RP accent, see A. C. Gimson (1993) , London: Arnold (third edition, ed. by A. Cruttenden).

For an introduction to English phonetics and phonology which covers, in much greater detail, some of what we have covered here, and more, see H. Giegerich (1992) , Cambridge: Cambridge University Press.

For an introduction to phonological theory, see any of the following: P. Carr and J.-P. Montreuil (2012) (second edition), London: Palgrave Macmillan; J. Durand (1990) , London: Longman; F. Katamba (1988) , London: Longman; C. Gussenhoven and H. Jacobs (2005) (second edition), London: Arnold; R. Lass (1984) , Cambridge: Cambridge University Press; A. Spencer (1996) , Oxford: Blackwell; I. Roca and W. Johnson (1999) , Oxford: Blackwell.

Students may proceed from one of these textbooks to more advanced treatments of phonological theory, such as J. Goldsmith (1989) , Oxford: Blackwell; M. Kenstowicz (1994) , Oxford: Blackwell; and R. Kager (1999) , Cambridge: Cambridge University Press. For an approach to English phonetics and phonology from the viewpoint of the theory known as government phonology, with good coverage of varieties of British and American English, see J. Harris (1994) , Oxford: Blackwell.

The account of English word stress given here owes a great deal to the work of Lionel Guierre and his followers, especially A. Dechamps, J.-L. Duchet, M. Fournier and M. O'Neill (2004) , Paris: Ophrys. For detailed coverage of English word stress, see E. Fudge (1984) , London: Allen and Unwin.

For a general introduction to the study of intonation, see A. Cruttenden (1986) , Cambridge: Cambridge University Press. See too D. Crystal (1969) , Cambridge: Cambridge University Press. An excellent recent textbook coverage of English intonation is J. C. Wells (2006)

, Cambridge: Cambridge University Press. The accompanying CD is most useful.

A useful book on British accents and dialects is A. Hughes and P. Trudgill (1987) (second edition), London: Arnold. For an extensive description of the phonetics and phonology of a very wide range of English accents worldwide, see J. C. Wells's (1982) three-volume work , Cambridge: Cambridge University Press. I have adopted Wells's three-way distinction (derived from Trubetzkoy) between systemic, realizational and lexical distributional differences between accents. For further textbook discussion of accent variation based on this tripartite distinction, see the relevant parts of Giegerich (1992; see above) and, of course, Wells (1982). I have followed Giegerich (1992) in comparing and contrasting SSE with RP and GA (for the simple reason that SSE is the medium in which I teach English phonetics and phonology). The reader should consult Giegerich (1992) for similar sorts of discussion. For an excellent introduction to phonetics and phonology in general, and varieties of English, see B. Collins and I. M. Mees (2008) (second edition), London: Routledge. As with the Wells book on intonation, the accompanying CD is very useful. For further reading on specific varieties of English, see J. Durand (2004) 'English in early 21st century Scotland: a phonological perspective', 36:87–105; A. Przewozny (2004) 'Variation in Australian English', 36:74–86; and D. Watt and L. Milroy (1999) 'Patterns of variation and change in three

Newcastle vowels: is this dialect leveling?' in G. Docherty and P. Foulkes (eds.),

, London: Arnold, pp. 25–46.

It is as well to point out to any reader who wishes to pursue the further reading suggested here that the choice of symbols used to represent the vowel phonemes of various accents of English will almost certainly vary from one author to another. This is inevitable, since there is a necessary degree of arbitrariness built into such choices. However, the reader should not find it too demanding to work out the correspondences between the symbols. The main point, for non-native speakers, is that much of the literature, and also the pronouncing dictionaries, use the symbol /e/ for words such as 'dress' in RP and GA, whereas we have used the phonetically more accurate epsilon symbol 'ɛ', in both phonemic and phonetic representations. We have used that same symbol for the RP centring diphthong in words of the sort 'square': /ɛə/. This helps us show that the current monophthongal realization in RP is the long, low-mid vowel [ɛː], as in [skwɛː] ('square').

Index