CAMBRIDGE TEXTBOOKS IN LINGUISTICS

Principles of phonetics

John Laver

PRINCIPLES OF PHONETICS

JOHN LAVER

PROFESSOR OF PHONETICS
UNIVERSITY OF EDINBURGH

CAMBRIDGE
UNIVERSITY PRESS

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

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011–4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 1994
Reprinted 1995, 2002

Printed in the United Kingdom at the University Press, Cambridge

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

Library of Congress Cataloguing in Publication data
Laver, John
Principles of Phonetics / John Laver.
p. cm. – (Cambridge textbooks in linguistics)
Includes bibliographical references and index.
ISBN 0 521 45031 4 (hardback) ISBN 0 521 45655 X (paperback)
1. Phonetics. I. Title. II. Series.
P221.L293 1994
414–dc20 93-18183 CIP

ISBN 0 521 45031 4 hardback
ISBN 0 521 45655 X paperback

TAG

CONTENTS

TABLES

ACKNOWLEDGEMENTS

In writing this book, I have had the benefit of help, advice and critical comment from many friends and colleagues: Marie Alexander, Brian Annan, Ron Asher, Saiqa Asif, Janet Mackenzie Beck, André-Pierre Benguerel, Art Blokland, Gunilla Blom, Adam Brown, Fouzia Bukshaisha, Guy Carden, John Clark, Bruce Connell, Alan Cruttenden, Jonathan Dalby, Gerard Docherty, Hartvig Eckert, John Esling, Ahmed Ferhardi, Arne Foldvik, Helen Fraser, Lindsay Friedman, Osamu Fujimura, John Gilbert, Jonathan Harrington, Jimmy Harris, Tom Harris, Dick Hayward, Katrina Mickey Hayward, Eugénie Henderson, Jo-Ann Higgs, Steve Hiller, Sandy Hutcheson, Mervyn Jack, Carolyn Johnson, Yasuo Kato, John Kelly, Alan Kemp, Handan Kopkalli, Bob Ladd, Peter Ladefoged, Roger Lass, Timo Lauttamus, Nick Laver, Björn Lindblom, Jim Lubker, Sudaporn Luksaneeyanawin, Fergus McInnes, Angus McIntosh, Geoff Nathan, Paroo Nihalani, Francis Nolan, John Ohala, Sallyanne Palethorpe, Jeff Pittam, Peter Roach, Sheila Rowe, Edmund Rooney, Ann Sutherland, Robin Thelwall, Elizabeth Uldall, Jo Verhoeven, Nigel Vincent, Briony Williams, Ken Winch and Colin Yallop.

Jo-Ann Higgs and Gerard Docherty used early versions of the text as a support for their advanced undergraduate course in phonetic theory in the Department of Linguistics at the University of Edinburgh over a number of years, and their comments have been of special value to me. I have also used later drafts of the book in teaching this course myself, and in teaching one on phonetics for postgraduate students of the Department of Applied Linguistics at the University of Edinburgh, and I am very grateful to all the students who discussed their response to the text with me.

Eugénie Henderson and Peter Ladefoged both contributed editorial advice during the preparation of the book, and their guidance has been very welcome. It is a matter of extreme regret to me that Eugénie Henderson's death in July 1989 prevented her seeing the full outcome of her characteristically benevolent advice. Lindsay Friedman helped to bring the book

through the middle stages of its preparation, and her critique of early theoretical matters was invaluable. In the final stages, Marie Alexander worked very hard in helping to check the consistency of the text, the figures, the appendices and the language statistics. I very much appreciated her meticulous and knowledgeable help. Doreen Cairns typed early versions of the manuscript, and Jan Wallace helped more recently with references.

Frances MacCurtain of the Middlesex Hospital in London made the xeroradiographic photograph of my production of the neutral configuration of the vocal organs shown as the frontispiece, from which Janet Mackenzie Beck created the drawing reproduced on the page facing it. She also drew the schematized version reproduced as figure 5.12, which has been used as the basis for many of the articulatory diagrams throughout the book. Tom Harris took the fiberoptic telelaryngoscopic pictures reproduced in figure 7.4.

Nigel Vincent offered some particularly constructive criticisms as a reviewer of the manuscript for Cambridge University Press. Some of the suggestions made by several anonymous reviewers have also been incorporated. Penny Carter, Marion Smith and Judith Ayling have been the staunchest (and surely the most patient) of editors. Jenny Potts made a very significant contribution as copy-editor for the Press. Fiona Barr composed the indexes. I offer my warmest thanks to all these colleagues and friends. I am responsible for the unwitting or wilful mistakes that remain.

In a textbook of this sort I also owe a special debt to all those writers whose clarity of phonetic commentary has allowed me to draw examples from their work on the phonetics and phonology of individual languages. I am especially grateful to John Esling of the University of Victoria and Speech Technology Research Ltd of Victoria, Canada, and to Kay Elemetrics Corporation of New Jersey, USA, for their kind permission to cite phonetic examples from the twenty-five languages illustrated in their joint project on the *University of Victoria Phonetic Database* (Esling 1991). The language database was developed at the Phonetics Laboratory of the University of Victoria, and sound recordings are available on floppy disks for use with an IBM PC and Kay Elemetric Corporation's *Computerized Speech Lab*. For more information, contact Dr John Esling at the Phonetics Laboratory, Department of Linguistics, University of Victoria, PO Box 3045, Victoria, BC V8W 3P4, Canada, or Kay Elemetrics Corporation at 12 Maple Avenue, PO Box 2025, Pine Brook, NJ 07058-2025, USA.

I am grateful to Marconi Speech and Information Systems, our industrial partner in the IED/SERC (Alvey) Large Scale Demonstrator Project on Integrated Speech Technology, for consent to the publication of the

phonetic distance matrices in chapter 13. I thank the Alvey Directorate and the Information Engineering Directorate of the Department of Trade and Industry, and the Science and Engineering Research Council for their support of this project over the period 1984–91 (IED/SERC Grants GRD/29604, GRD/29611, GRD/29628, GRF/10309, GRF/10316, and GRF/70471). I also thank the Medical Research Council for its support of our work on vocal profiles (MRC Grant G987/1192/N: 1979–82), and our work on microperturbatory aspects of laryngeal function (MRC Grant 8207136N : 1982–5), parts of which are described in this book. Permission to adapt material from copyright publications for a number of figures and tables is gratefully acknowledged, as indicated in the list attached below.

Tolerance and support have been provided by many people in connection with the preparation of this book, not least by my colleagues in the Centre for Speech Technology Research in the University of Edinburgh, especially Mervyn Jack and Ron Asher. The highest price for a book, however, is always paid by an author's family. For their enduring love and support, I dedicate this book to my wife and colleague Sandy Hutcheson and to my children Nick, Michael, Claire and Matthew.

'IBM' and 'PC' are registered trademarks of International Business Machines Corporation. 'Computerized Speech Lab' is a registered trademark of Kay Elemetrics Corporation.

Permission to reproduce copyright material is gratefully acknowledged for the following figures, tables and appendices: Academic Press (fig. 15.8, table 6.3); Edward Arnold (figs. 15.14, 15.19); Professor T. Balasubramanian (figs. 8.3a, 8.3b); Basil Blackwell (figs. 3.3, 3.4, 6.2, 7.6); Professor L. Cagliari (fig. 10.16); Cambridge University Press (figs. 3.1, 4.1, 7.1, 7.2, 7.3, 10.9, 15.5, 15.6, 15.7, 15.11, 15.12, 15.16, 15.20, tables 12.1, 19.1 and the map of language families in appendix II); Chulalongkorn University Press (fig. 15.9, table 15.3); Croom Helm (fig. 3.2); Edinburgh University Press (figs. 10.3, 10.4, 10.6); Lawrence Erlbaum Associates (fig. 12.4); Professor G. Fant and Dr A. Kruckenberg (fig. 15.4); Foris (fig. 11.4); Julius Groos Verlag (figs. 10.18, 10.20); T. Harris (fig 7.4); Heffer (figs. 10.4, 10.5); Dr S. Hiller (fig. 15.18); the *International Journal of American Linguistics* (table 12.2); the *International Phonetic Association* (figs. 10.7, 10.14, 15.15a and b, appendix I); S. Karger (fig. 10.17, table 15.2); Longman (fig. 14.1); Marconi Speech and Information Systems (fig. 13.1); Dr F. McInnes (fig. 13.2); Penguin (fig. 15.13); Prentice-Hall (fig. 6.1, table 6.1); Springer-Verlag (figs. 4.3, 10.19, 15.2, 17.1); University of Chicago Press (fig. 12.2).

Introduction

This book is designed to equip the reader with a foundation for independent research in the phonetic study of speech. It makes no assumptions about prior knowledge of the subject, and is meant to be suitable as a comprehensive reference textbook on phonetics for graduates with no previous direct experience in the subject, who are following conversion courses to prepare themselves for research in phonetics and related subjects. It can also be used as a textbook supporting undergraduate courses on phonetics from elementary to advanced levels.

The book is thus addressed to readers who wish to pursue the subject from an initial to an advanced stage. I have tried to offer a general orientation, and a framework which provides a guide to navigation through some of the central territory of phonetics. Above all, I hope that the book may stimulate readers to reflect on the theoretical foundations of the subject.

Speech is our most human characteristic. It is the most highly skilled muscular activity that human beings ever achieve, requiring the precise and rapid co-ordination of more than eighty different muscles, many of them paired. Even the expertise of the concert pianist pales into relative insignificance beside the intricately co-ordinated muscular vocal skills exercised by a ten-year-old child talking to friends in the school playground. A pianist playing a rapid arpeggio makes about sixteen finger strokes per second, each the product of multiple motor commands to the muscles of the fingers, wrist and arm. Speech is both faster and more complicated (Boomer 1978). The process of speaking at a normal rate is achieved by means of some 1,400 motor commands per second to the muscles of the speech apparatus (Lenneberg 1967). As children, we take a number of years to acquire the skills of producing and perceiving speech, but once we have learned these abilities, only pathology or accident deprives us of their use.

Speech is the prime means of communication for virtually every social group, and the structure of society itself would be substantially different if we had failed to develop communication through speech (Bickerton 1990:

240). Furthermore, our social interaction through speech depends on much more than solely the linguistic nature of the spoken messages exchanged. The voice is the very emblem of the speaker, indelibly woven into the fabric of speech. In this sense, each of our utterances of spoken language carries not only its own message, but through accent, tone of voice and habitual voice quality it is at the same time an audible declaration of our membership of particular social and regional groups, of our individual physical and psychological identity, and of our momentary mood.

A comprehensive understanding of a phenomenon as complex and multistranded as speech necessarily has to draw on the resources of a large number of different disciplines. Given the prime communicative function of spoken language, one discipline that lies at the heart of any adequate study of speech is linguistics. But speech, as just indicated, is a carrier of more information than solely the meaningful patterns of individual utterances of spoken language. A vast amount of social and personal information about the speaker is carried as well. In this century, the horizons of communication through speech have expanded greatly, through telecommunications, broadcasting and computing. To reach a full understanding of the nature of communication through speech one would therefore have to appeal to concepts not only from linguistics, but also from sociology, anthropology, philosophy, psychology, anatomy, physiology, neurology, medicine, pathology, acoustics, physics, cybernetics, electronic engineering, computer science and artificial intelligence. The study of speech in this broad view thus covers a remarkably wide domain, embracing aspects of the social sciences, the life sciences, the physical sciences, the engineering sciences and the information sciences.

All the disciplines mentioned above take speech as part of their professional domain, though their principal focus is elsewhere. Phonetics is the discipline that takes speech as its central domain, and which stands at the intersection of all these disciplines. In the definition to be promoted in this book, phonetics is the scientific study of all aspects of speech. Phonetics and linguistics are seen here as sister sciences, together making up 'the linguistic sciences'. This allows one to say that phonetics is the scientific study of speech, and that linguistics is the scientific study of language. Their domains overlap in the area of spoken language, but each has legitimate concerns outside the professional remit of the other.

More traditional definitions of phonetics content themselves with a narrower range of territory, usually limiting the domain of phonetics to the study of those aspects of speech relevant to language, which locates phonetics within the encompassing discipline of linguistics. One candid objective of

be properly achieved by simultaneously absorbing the underlying structure of phonetic theory. For convenience of reference, the sets of symbols and labels relevant to each chapter are gathered together at the end of that chapter.

The sequence of presentation of the descriptive theory merits some comment. The book is divided into eight parts. Part I consists of three introductory chapters explaining general concepts used in the analysis of spoken language. Chapter 1 presents elementary ideas helpful in understanding the coded nature of spoken language and the different types of information available to the listener, and introduces a number of different levels of analysis. Chapter 2 offers a basic view of the relationship between phonetics and phonology, as the abstract set of sound-patterns that constitute the coded nature of linguistic communication through speech. Chapter 3 explores the distinction between an 'accent', a 'dialect' and a 'language', and touches on a number of linguistic disciplines to which phonetics makes a contribution, such as sociolinguistics, dialect geography and historical linguistics.

The two chapters of Part II then launch the more technical core of the book. Chapter 4 begins by considering the relationship between biological aspects of the speech apparatus and its phonetic use. It then explores different conceptual approaches to the question of how the stream of speech can be segmented for analysis, and proposes some fundamental units of phonetic analysis: features, segments, syllables, settings, utterances and speaking turns. The concept of the segment, as the phonetic manifestation of the linguistic units of consonants and vowels, is probably the one initially most familiar to readers of this book. Chapter 4 concludes with a preliminary identification of the six basic strands of speech production: initiation of an airstream, phonation, articulation, temporal, prosodic and metrical organization.

Chapter 5 is a pivotal chapter in the design of the book. Because chapters 6–17 look in turn at each strand of vocal performance in fairly extensive detail, it was thought helpful to the reader to have an initial summary overview of the whole architecture of the descriptive theory proposed in the book, before launching into the local detail of each of the following chapters. Chapter 5 is therefore designed to give the reader an architectural frame of reference within which to locate the remaining chapters of the book. By its nature it can only give a skeletal outline of a large and detailed area. But its summary form may also allow it to function as a review of the design of the descriptive theory to which the reader can return as desired. The principal innovations in the architecture of the descriptive theory

presented in this book are the concepts of 'aspect of articulation', 'inter-segmental co-ordination' and featural 'settings'.

Part III consists of two chapters explaining initiation and phonation. Chapter 6 on initiation describes the mechanisms for setting a flow of air in motion for the purposes of speech. Chapter 7 on phonation explains the ways that the larynx can impose a variety of patterns of vibration on the flow of air from the lungs, to provide a source of acoustic energy which the articulatory apparatus can then further modify to produce the segments of speech.

The four chapters of Part IV present an analysis of the segmental performance of the sounds representing consonants and vowels. Stop articulations are described in chapter 8, fricative articulations in chapter 9 and resonant articulations in chapter 10. Chapter 11 explains the articulatory basis of segments made with more than one major articulatory constriction.

Part V explores the relationships between segments next to or near to each other in the stream of speech. Chapter 12 looks at how the articulation of adjacent segments is co-ordinated, especially in the events at the junction of the two sounds. Chapter 13 explores in more detail than is usual in phonetic textbooks the topic of phonetic similarity between segments, invoking the concept of a 'setting' as a feature running through several segments to explain degrees of similarity, and using setting-analysis as a basis for constructing descriptive 'vocal profiles' of the voices of individual speakers.

Part VI presents a temporal, prosodic and metrical analysis of speech. Chapter 14 considers matters of the temporal duration of segments. Chapter 15 moves to the prosodic description of the use of pitch, in tonal and intonational functions of speech melody. It also discusses the control of loudness in speech. Chapter 16 integrates the temporal and prosodic material of the preceding two chapters, and presents a metrical analysis of speech where the focus is on stress and rhythm. Chapter 17 then presents a description of the way that speech can vary in its rate and continuity, as a final part of the temporal organization of speech.

Part VII examines the principles by which different types of transcription can be classified. Chapter 18 includes numerous examples of detailed phonetic transcriptions of utterances from native speakers of each of a number of languages from widely different parts of the world. These transcriptions were made jointly by the author and a colleague in the University of Edinburgh (Sandy Hutcheson), for the purpose of teaching postgraduate courses in the use of phonetic transcription for fieldwork. The transcriptions called potentially on the entire repertoire of general phonetic notation. The chapters leading up to chapter 18 explain the meanings and use of all the

transcriptional symbols available in this repertoire, and the relationships between them. If, when chapter 18 is reached, the transcriptions in this chapter are able to be interpreted by the reader with full and confident understanding, then the book will have succeeded in a significant part of its objective of providing a foundation for independent phonetic research. Learning the meaning and use of the phonetic symbols has an obvious practical value, but the most important consequence of acquiring a deep understanding of the notation is the accompanying insight that is gained into the underlying theoretical architecture that binds the concepts represented by the symbols into a coherent framework.

Part VIII rehearses the main principles of phonetic description that have been presented, placing them in a broader context of other approaches to the description of speech. In addition, chapter 19 considers the question of empirical justification for the architectural design of a descriptive phonetic theory, by appeal to statistical facts about the relative frequency of occurrence of the different phenomena described.

Phonetic concepts presented in the book are illustrated from as wide a range of different languages as feasible, drawn from reliable sources in the literature of the subject and from personal experience. The linguistic examples are co-ordinated with the introduction of phonetic concepts and symbols, and illustrations in phonetic transcription are offered from over 500 different languages. There are several reasons for choosing to illustrate the book with copious examples drawn from this wealth of languages. A central part of the motivation in drawing on a wide range of linguistic material is to offer objective evidence to support the generality of the descriptive theory being put forward. A related objective is to give the reader an appreciation of the paradox that, underlying the apparently extraordinary diversity of the patterns of spoken language, there is a remarkable unity in the phonetic resources exploited by the languages of the world. The challenge of writing this book lay above all in the attempt to capture the range and elegance of this combination of diversity and unity.

Every effort has been made to account for all phonetic processes known to be exploited by spoken language. If any reader knows of phonetic phenomena in some language which are either not covered, or inadequately described, I would be most grateful to receive information about the nature of the phonetic process concerned, with phonetic transcriptions and glosses of illustrative words, for possible inclusion in a later edition.

Another motivation for frequently drawing on illustrations from real languages is that the theoretical discussion of phonetic principles can become very abstract without continual renewal of connection with the linguistic

material that the theory seeks to illuminate. A final part of the motivation is the belief that readers who experiment with their own ability to pronounce the illustrative material are more likely to reach an integrated understanding of the underpinning theoretical concepts. Readers are encouraged to use the linguistic examples for practical performance, especially (but not only) if the material can be practised with the help of an experienced phonetician.

The language material is normally presented in phonetic (rather than 'phonemic') transcription. Where no source is indicated for language examples, the illustration is offered from my own experience. The symbols used in the transcriptions of examples drawn from the work of other phoneticians and phonologists have been standardized wherever necessary to fit the conventions of this book. Where examples from a particular language have been borrowed from another writer, specific reference is always given, to give due acknowledgement and to allow consultation for further details.

Particular care has been taken to indicate, wherever feasible, the original source of technical concepts. I feel that it is important for students and professionals alike to appreciate the intellectual continuities in the broad landscape of the rich and well-populated literature of phonetics. Any reader fortunate enough to acquire the majority of the publications listed in the References to this book would possess the nucleus of a very good phonetics library. Recommended further reading is indicated at the end of chapters.

Wherever practicable, the language-family and the part of the world in which the language is spoken is given. Names of countries are as up-to-date as political developments at the time of going to press allowed. Each of the languages mentioned in the text is listed in appendix II, together with its language-family and geographical location. The affiliation identified in appendix II is normally the major language-group of which the language is a member; the language-family given in the body of the text usually indicates a lower-level grouping. Appendix II also serves as a Language Index, giving the page numbers of all mentions of the language concerned. The appendix is prefaced by a map of the world on which the locations of the major language-families are indicated (Crystal 1987: 294–5).

Examples are drawn from the language-families of every inhabited continent, but the languages of three particular areas tend to recur. One is Europe, with its relatively familiar (mostly Indo-European) languages. The other two are Africa and the Americas. The languages of Africa have received a good deal of attention from phoneticians and phonologists, and Africa (particularly the languages of the Niger-Congo family) is the area with which I am myself most familiar. The languages of North and South America and the Mesoamerican Indian languages of Central America have

received particular attention from the phoneticians and linguists trained in the American tradition, especially from the many linguists associated with the Summer Institute of Linguistics. Both Africa and the Americas are areas rich in phonetic processes unfamiliar to the languages of Europe. For comprehensive information on the language-families of the world, the reader might like to consult the publications listed in the section on Further Reading at the end of chapter 3 and at the beginning of appendix II.

The language examples offered in the first four chapters are deliberately taken mostly from English. This is because it is likely to be the native or second language of the large majority of readers of this book, and therefore the examples will perhaps have a greater immediacy for these readers than they might otherwise have. But it is also directly because the detailed discussion of matters of pronunciation of less familiar languages really requires a prior understanding of the descriptive articulatory concepts offered in the later chapters, in which the examples drawn from languages other than English are concentrated.

The English examples are normally taken from an accent of England called Received Pronunciation (RP). RP is a non-regional accent of England, and is probably most familiar to the readers of this book as the accent used by most announcers and newsreaders on national and international BBC broadcasting channels. If no further identification is given to the term 'English', it can be assumed that reference is being made to English pronounced with an RP accent. RP is discussed more fully in chapter 3.

Words in many languages are contrastively identified by different patterns of pitch, or 'tone'. Diacritic symbols marking such tonal characteristics are often included in the transcriptions of the linguistic examples, when known, for review purposes. They typically consist of acute or grave accent-diacritics, placed above the symbols representing vowels. They can perhaps be left unremarked during the first reading of the text until chapter 15, where tone is discussed in more detail.

Figures are numbered in sequence throughout each chapter, as are tables. Technical terms, on introduction and definition, and as topic identifiers in the recommendations for Further Reading at the ends of chapters, are printed in **bold-face**. The Subject Index prints the page numbers of these boldened items also in bold, for convenience of review.

1

The semiotic framework

The point of departure for this book is that speech is the most subtle and complex tool of communication that man has ever developed. The aim of this chapter is to discuss the nature of spoken communication, and to offer a broad framework within which the analysis of speech in general, and of spoken language in particular can be set. Some basic concepts are defined for use throughout the book.

One of the most important concepts to be introduced is the notion of the acts of spoken language as artefacts, used as elements of a coded system of **signs**. As a **code**, spoken communication can only be successful when used between people skilled in the production and interpretation of the relevant signs. Another very important concept discussed in this chapter is the many-layered nature of speech, with each layer carrying a different type of information. The framework proposed here for the analysis of these concepts comes originally from the discipline of semiotics.

1.1 Semiotics as the general theory of signs

Semiotics is the study of all aspects of sign systems used for communication. The term 'semiotic' was first used in ancient Greek medicine, to refer to the theory of medical symptoms used in the diagnosis and prognosis of disease (Morris 1946: 285). The Stoic philosophers then used the term to refer to the general theory of signs. The semiotic analysis of speech can involve a consideration not only of the communicative signs themselves, but also of the mechanisms by which the signs are produced by the **speaker** and perceived by the **listener**, using the auditory and visual channels of communication. Also of semiotic interest are the ways in which the manufacture and use of particular signs can carry information about the characteristics of their producers.

There is a continual flow of information between participants in any conversation. The first step in setting up a convenient semiotic framework for the discussion of how speech works is to distinguish between three major

types of information exchanged. In order, these will be called 'semantic', 'evidential' and 'regulative' information.

1.2 Semantic, evidential and regulative information in speech

The largest part of this book will be about the analysis of the aspects of speech that convey the first type of information mentioned above – **semantic information**. This is what most people think of as the direct 'meaning' of a spoken utterance. It is the propositional content of the communicative acts of conversation, and the more complex the proposition the more likely it is to rely on being communicated by spoken words. Simple semantic information can often be exchanged by other means, however. A message corresponding to the potential utterance 'Come here' can often be communicated perfectly unambiguously by a beckoning gesture, for instance, within the interpretive conventions of the particular culture for interpreting 'meaning'. More indirect aspects of meaning, such as **pragmatic meaning** (where an utterance such as *It's cold in here* is taken to constitute a request to shut the door rather than solely to offer a comment about the local temperature), will not be dealt with directly in this book. References to work on pragmatics can be found in the recommended publications for Further Reading at the end of the chapter.

Evidential information can be said to be conveyed by signs in speech which act as attributive **markers** (Laver and Trudgill 1979: 3). These are used by the listener as the basis on which to attribute personal characteristics to the speaker. The attributes of the speaker fall into three groups:

physical markers – those that indicate physical characteristics such as sex, age, physique and state of health;

social markers – those that indicate social characteristics such as regional affiliation, social and educational status, occupation and social role;

psychological markers – those that indicate psychological characteristics of personality and affective state or mood.

Markers of physical characteristics lie, for example, in a speaker's **voice quality**. Social markers often include such features as **accent** and choice of vocabulary. Psychological markers of personality and mood are often taken to reside in a speaker's (habitual and momentary) **tone of voice**. In the sense that evidential markers of this sort can be said to be **symptoms** of a speaker's individual characteristics, their semiotic status is close to the diagnostic role of signs in the medical origins of semiotic theory in ancient Greece.

The meaning of the term 'marker' as presented here is derived from the concept that Abercrombie (1967: 6) introduced to the semiotic study of speech under the name 'index'. Abercrombie drew this term, with its related adjective 'indexical', from the semiotic writings of Charles Peirce, the late-nineteenth-century American pragmaticist philosopher (Buchler 1940).

The markers that serve to identify an individual's membership of a given social group are of especial interest to the sector of linguistics usually called **sociolinguistics** or the **sociology of language**. Seminal work in this area was done by Labov (1966, 1972a, 1972b), and by Hymes (1974, 1977) and Trudgill (1974), to which interested readers are directed. Other useful sources on sociolinguistics are given in the list of Further Reading at the end of the chapter.

There is one particular sort of evidential information that it is helpful to label separately, which identifies the speaker's desire to retain or to yield the role of speaker in the time-course of a given conversation. This can be called **regulative information**, and it is used by the participants in a conversation to control the time-sharing of the interaction. In most conversations the participants take turns to occupy the role of speaker, and conventional mechanisms exist for managing smooth transitions at turn-boundaries. These are well understood by adult participants native to the culture. They are normally exercised outside conscious awareness, but the temporal structure of speaking-turns in conversation is a skilled product of co-operative work between the participants. The skill has to be learned by each speaker as part of his or her socialization into the use of spoken language. It is hence a skill whose details are specific to the language community concerned. Yielding the floor to the other speaker is managed in English by a variety of conventional signals. Some of them use the visual channel (head-movements and eye-contacts), and some of them use speech (mostly particular types of intonation and timing).

Having drawn this triple distinction between the semantic, evidential and regulative types of information exchanged in conversational interaction, we can now turn to a more detailed consideration of the different means available to speakers for transmitting this information. The chief distinction to be drawn here is that between vocal and non-vocal behaviour. By definition, vocal behaviour is audible, while non-vocal behaviour is visible.

1.3 Vocal and non-vocal behaviour

It will be clear from the discussion in the section above that the information imparted by the **vocal behaviour** of speech consists of more than the audible elements of spoken language. When we listen to a speaker, we

perceive not only semantic information conveyed by the words the speaker is producing, but also regulative information about the conversational structure of the utterance. We also perceive evidential information marking the speaker's identity and his or her affective state. In the definitions used in this book, the analysis of speech covers the study of this whole complex of vocal behaviour – the coded vocal material of spoken language, the vocal shaping of conversational structure, and the vocal marking of speaker-attributes.

Conversation itself consists of more than just vocal behaviour, however. In all normal circumstances of face-to-face conversation, there is a constant two-way traffic of information on the visual channel (particularly of evidential and regulative information), communicated by **non-vocal behaviour**. Gesture, posture, head-movements, body-movements, facial expression, gaze and eye-contact behaviour are continually in play, complementing and supporting speech activity. These non-vocal aspects are themselves elements of a coded system of communication, which users have to learn. A small part of non-vocal communication (notably some aspects of facial expression) is thought to be of universal human significance, but most non-vocal communication consists of behaviour that is particular to the culture of the speaker. This too is therefore a skill whose performance has to be learned, albeit a somewhat less complex skill than that involved in spoken language. The analysis of non-vocal material lies outside the study of speech as such, but it is integral to the study of the communicative context in which speech is almost always performed.

1.4 Verbal and non-verbal behaviour

The second distinction to be drawn between different means of communicating information is that between verbal and non-verbal means. These terms are interpreted in many different ways in books on speech and conversational interaction. In this book, **verbal** elements are defined as any aspects of communicative behaviour that serve to identify individual words as units of spoken language; **non-verbal** elements are then any features of communicative behaviour that serve functions other than that of verbal identification. **Vowels** and **consonants**, for example, serve the verbal function, as does **word-stress**. **Intonation**, as pitch-melody associated with units of spoken language usually longer than individual words, is a non-verbal element, as is the use of stress for purposes of emphasis (**emphatic stress**). Linguistic communication is thus served by both verbal and non-verbal devices. Verbal devices are solely linguistic, but non-verbal communication includes both linguistic and non-linguistic processes (Key 1980). One such non-linguistic process is communication through **tone of voice** (discussed

below in section 1.10) as a non-verbal vocal element. Examples of non-verbal non-vocal means of communication include gesture, posture, looking-behaviour and facial expression.

1.5 Signs and symbols

The semiotic framework used in this book, following the precedent set by Abercrombie (1967), is largely derived from some of the key concepts in the writings of Charles Peirce, the pragmaticist philosopher mentioned earlier. Peirce defined semiotics as 'the formal doctrine of signs ... [where] ... a sign is something which stands to somebody for something in some respect or capacity' (Hartshorne and Weiss 1931–5, vol. II: 227–8). A three-way relationship is implied between a person, an entity and the sign used to refer to the entity. The referential link between the entity and the sign can be either non-arbitrary or arbitrary. The link between wind-direction and a weathercock is non-arbitrary, for example, and many signs which are used to communicate evidential information are of this non-arbitrary kind. The habitual pitch of a person's voice, for instance, is often fairly direct evidence of the overall size of the vocal folds in the speaker's larynx.

More tenuously, a non-arbitrary link is often claimed to exist between certain **onomatopoeic signs** used in language, such as the words *clink* and *clank*, and the entities to which they refer. But such onomatopoeic signs are a tiny minority of the signs used for linguistic purposes. Linguistic signs virtually all show a culturally determined, arbitrary link between the sign and the entity (or **referent**) to which reference is made. Linguistic signs of this latter sort are usually called **symbolic signs**. A **symbol** is a sign whose reference is arbitrary, and governed only by social convention. The meaning of a symbolic sign can thus only be appreciated by someone who has learned the appropriate referential convention. There is no element of necessity in the referential link between the linguistic sign *cat* and the feline animal which is its conventional referent in English, for example. The arbitrary convention linking such symbolic signs with their referents is a convention to which all users of a given language tacitly subscribe.

Words are symbolic signs, and the semiotic status of these linguistic elements is relatively clearcut. But the semiotic status of the individual consonants and vowels which make up the verbal means for differentiating one word from another is less obvious. It is helpful to draw a sharp theoretical distinction between an entity such as an individual consonant, and the actual event of pronunciation. We can call the act of pronunciation a **speech-sound**, for convenience. The semiotic position to be adopted here is that it is the speech-sound that 'stands for' the consonant, and that a conso-

is one where the relevant distinctive pattern is fully available to the listener in the actual pronunciation of the material concerned.

Pattern as such is an abstract concept. To exist in the real world, a given pattern has to have a physical **embodiment** (or **realization**, or **manifestation**). Terms often associated with pattern and embodiment are 'form' and 'substance' respectively (Abercrombie 1967). **Form** is to do with the identity of the pattern as a representative of a linguistic unit, and **substance** is to do with the medium in which linguistic patterns are embodied. In this connection, **phonology** is often said to be the study of the form of spoken language, and **phonetics** to be the study of the substance of spoken language. It would be rather more accurate, in terms of some of the attitudes espoused by this book, to say that phonology and phonetics both concern themselves with the form and substance of spoken language, though they vary in their primary focus. Phonology has a primary interest in the form of spoken language, while phonetics has a primary interest in the substance of spoken language. This book, however, goes further than this. It presents phonetics as a subject that also takes a legitimate interest in facets of the substance of speech other than solely those that embody the formal symbolic patterns of spoken language. Two of these facets have already been mentioned. One (the evidential aspect of speech) concerns the way that evidence about the identity of the speaker is carried in speech. Another (the regulative aspect of speech) concerns the way that participants in a conversation co-operate with each other in the management of taking turns to occupy the role of speaker.

1.8 Code and medium

Formal aspects of speech make up the code of spoken language. A speaker **encodes** a message into its particular linguistic form by creating appropriate patterns through his or her manipulation of the **medium** of speech (Abercrombie 1967: 1). Since the code must be embodied in a physical medium, all instances of coded linguistic messages are necessarily transmitted as artefacts created out of the substance of that medium. The listener **decodes** the message partly by attending to the distinctive patterns available in the artefacts of the substance of the utterance, and partly through predictive knowledge of probable grammatical sequences and likely choices of vocabulary.

The events of speech are artefacts of the spoken medium, just as the events of writing are artefacts of the written medium. Because these events are artefacts, they not only embody patterns, they also carry with them clear evidential information about their manufacture. This information indicates both the type of apparatus used to create the artefacts, and the personal style of

the artisan who made them. Evidence of the nature of the vocal apparatus lies in the characteristic voice quality of the speaker, and evidence of personal style lies for example in the details of the speaker's habitual pronunciation. When we perceive a spoken message, therefore, we are able not only to decode the linguistic form of the message, but may also be able to identify the speaker, through his or her voice and accent. The symbolic information communicating the linguistic message and the evidential information marking the identity of the speaker are hence two sides of the same coin.

1.9 Communicative and informative behaviour

The term 'communicative' has been used so far without definition. It may now be helpful to make the meaning of this term more explicit. An important semiotic distinction can be drawn between 'communicative' and 'informative' aspects of signals used in speech. A signal is **communicative** if 'it is intended by the sender to make the receiver aware of something of which he was not previously aware'. A signal is **informative** if, regardless of the intentions of the sender, 'it makes the receiver aware of something of which he was not previously aware' (Lyons 1977: 33). Linguistic activity is thus both communicative and informative, under these definitions.

1.10 Linguistic, paralinguistic and extralinguistic behaviour

It is convenient to draw a second set of distinctions between different means for conveying information in speech, differentiating between linguistic, paralinguistic and extralinguistic behaviour. All three types of behaviour are informative, in the terms introduced immediately above, but only linguistic and paralinguistic behaviour is coded and communicative.

The term **linguistic** is familiar from everyday use as an adjective associated with **language**. The commonest form of linguistic behaviour is communicative behaviour which uses the dual-level code of spoken language made up of the phonological and grammatical units discussed above. Other forms of linguistic behaviour are found in writing systems (and sign language), but it is worth pointing out that literate communication is still a minority activity in the language communities of the world.

Paralinguistic behaviour in speech is communicative behaviour that is non-linguistic and non-verbal, but which is nevertheless coded, and which is designed to achieve two goals of conversational interaction. These are the communication of the speaker's current affective, attitudinal or emotional state (such as anger, sadness, excitement, disappointment, happiness, cordiality etc.), and the regulation of the time-sharing of the conversation.

The adjective 'paralinguistic' derives from the term 'paralanguage', which

was first suggested by the American linguist Archibald Hill. Paralinguistic behaviour includes communication by tone of voice. As the label 'paralinguistic' suggests, it is in some respects similar to linguistic behaviour, and in some other respects dissimilar. It is similar to linguistic behaviour in that paralinguistic signs form a coded type of communication whose meaning is not necessarily obvious to all human perceivers on some universal basis. It is particular to the culture of the speaker, and its conventional interpretation must be learned. For example, the interactional meaning of the use of a falsetto tone of voice by an adult male speaker is not necessarily the same from one culture to another. One paralinguistic use of falsetto in English by male speakers is as a mocking device, where mimicking another participant's utterance with a falsetto phonation counts as an accusation of either effeminacy or whining complaint. In Tzeltal, a Mayan language of Mexico, however, the use of falsetto in greeting someone is a marker of honorific respect (Brown and Levinson 1978: 272). In Shona, a language of Zimbabwe, falsetto is used to mock someone thought to be boasting (B. Annan, personal communication).

Sequential structure is crucial to the way that language works, with grammatical units being given their distinctive shape by differing linear combinations of phonological units. Paralinguistic communication differs from linguistic behaviour in this respect, in that no meaningfulness is imparted by sequence, and there is therefore no possibility of meaningful structure. The choice from a conventional repertoire of a given paralinguistic tone of voice, for instance, involves no consideration of what the immediately earlier choice was, nor of what the next one might be, in terms of any 'tone of voice' structure. The only way in which sequential relationships carry any significance in paralinguistic communication is in judgements about the relative degree of manifestation of a feature. A judgement by the listener that a speaker has started to speak more slowly, for example, clearly depends on the speaker's earlier rate of articulation. But such a relationship is merely sequential and not structural, in the technical sense of 'structure' to be described in the next chapter.

Paralinguistic behaviour is also different from the linguistic behaviour of spoken language in that non-vocal elements play a very large part. Gesture, posture, body-movement, facial expressions and looking-behaviour all form part of the elaborate, coded, paralinguistic system of communication together with the vocal strand of tone of voice. Facial expression is the strand amongst these paralinguistic resources that is most similar to that of tone of voice.

Extralinguistic behaviour in speech is then the residue of the speech signal after analysis of all coded linguistic and paralinguistic aspects is complete. Non-coded, extralinguistic aspects of speech are often rich in evidential

information about the identity of the speaker, particularly with respect to habitual factors such as the speaker's voice quality, and overall range of pitch and loudness (Laver 1980). Extralinguistic behaviour is thus informative, in the terms introduced earlier, but not communicative.

One might think that conclusions about social characteristics are more likely to be drawn from linguistic and paralinguistic sources, since by definition social characteristics are learned, not innate. Similarly, conclusions about psychological states and attributes might be thought to depend on linguistic and paralinguistic information. Equally, conclusions about physical attributes might be assumed to rely mostly on organic, extralinguistic information. The attribution of social, psychological and physical characteristics from speech cannot, however, be correlated directly with linguistic, paralinguistic and extralinguistic information respectively. In drawing conclusions about the personality of a speaker, for example, a listener may impute psychological attributes to the speaker (sometimes mistakenly) from evidence taken from any or all of the three sources of information (Sapir 1927). A listener might conclude, for instance, that a loud, resonant voice is correlated with an authoritative personality. But if that voice were solely the outcome of a large and powerful physique, then the listener would erroneously be drawing psychological conclusions from extralinguistic sources. The cognitive processes of attribution by listeners of social, psychological and physical characteristics to speakers from the rich variety of markers in their speech remains a fertile potential area for research (Laver and Trudgill 1979).

1.11 The domain of phonetics

Many traditional approaches to phonetics take the domain of the subject to be the description of spoken language. As evident from the Introduction, and from the comments offered in this chapter, this book urges a wider semiotic perspective, in the belief that the proper domain of phonetics as a discipline is the study of all aspects of speech. We shall return in chapter 4 to a more detailed consideration of the specific objectives of the discipline of phonetics.

Further reading

Sebeok (1991) is a readable and wide-ranging book about semiotics in general. Hartshorne and Weiss (1931–5) is a substantial edition of the collected writings of the pragmaticist philosopher Charles Sanders Peirce. The most accessible general introduction to Peirce's philosophy is Feibleman (1970). Burks (1949) continues some of Peirce's semiotic themes in his article on iconic, indexical and symbolic signs.

Useful textbooks and edited collections of articles about **sociolinguistics** are Chaika (1989), Fasold (1984, 1989), Fishman (1968), Gumperz (1983), Gumperz and Hymes (1986), Hudson (1980), Hymes (1974), Milroy (1980, 1987), Romaine (1982), Saville-Troika (1989), Trudgill (1978) and Wardhaugh (1986). Ager (1990) is a commentary on contemporary sociolinguistic variation in French in both France and Canada. Barbour and Stevenson (1990) outline sociolinguistic variation in German. Cheshire (1991) considers sociolinguistic matters in English in different countries in the world. For more references to sociolinguistic variation in English around the world, further recommended reading is suggested in chapter 3.

Topics in **general linguistics** are covered in the Cambridge University Press series of Textbooks in Linguistics, and in its other series on linguistics, in such publications as Allwood, Anderson and Dahl (1977) on logic in linguistics; Bauer (1983) on English word formation; Brown and Yule (1983) on discourse analysis; Bynon (1977) on historical linguistics; Comrie (1976) on aspect and (1985) on tense, and Palmer (1986) on mood and modality; Corbett (1991) on grammatical gender; Croft (1990) and Shopen (1985) on the typology of languages, and on typical grammatical features; Elliot (1981) on child language; Huddleston (1984, 1988) on the grammar of English; Hurford and Heasley (1983), Kempson (1977) and Lyons (1977) on semantics, and Cruse (1986) on lexical semantics; Levinson (1983) on pragmatics; Matthews (1974, 1981, 1991) on morphology and syntax respectively and Radford (1981, 1988) on transformational syntax. Newmeyer (1988, vol. I) is a review of most of these foundational areas of linguistic theory. Lyons (1968) is a comprehensive account of linguistic theory for students pursuing a specialized interest in the discipline; Lyons (1991) is a volume of essays on linguistic theory; and Lyons (1981) offers an overall view of linguistics for the educated non-specialist reader. Other recommended publications on **pragmatics** are Austin (1962), Blakemore (1992), Brown and Levinson (1978), Searle (1969) and Sperber and Wilson (1986).

The organization of behaviour in face-to-face **conversational interaction** is discussed in Atkinson and Heritage (1984), Kendon (1981), Laver (1976), Laver and Hutcheson (1972), Nofsinger (1991) Scherer and Ekman (1982) and Siegman and Feldstein (1978). A more detailed discussion of non-vocal behaviour and its relationship to vocal facets of conversation can be found in Buck (1984), Chaika (1989: 75–97), Coulmas (1981), Hinde (1972, 1974), Kendon (1981), Key (1980), Laver and Hutcheson (1972), Poyatos (1983) and Weitz (1974). Sacks, Schegloff and Jefferson (1974) is one of the foundational publications on strategies for conversational turn-taking.

Pioneering publications in **paralinguistics** were Abercrombie (1967, 1968),

But the senses of touch, pressure, muscle-tension and joint-position are all also relevant to considerations of how speakers control and monitor the actions of their vocal apparatus in the production of speech.

The auditory system is subject, as are all sensory systems, to psychophysical ranges and limits of sensitivity, and these are discussed at more length in chapter 15. The term **perceptual** will be used in the remainder of this book, if offered without further qualification, to refer specifically to auditory perception.

There are four perceptual domains available to the human auditory system which are exploited in listening to speech. These are the domains of perceptual **quality**, **duration**, **pitch** and **loudness**. The only way that sounds can differ audibly from each other is in terms of these four perceptual attributes. Chapters 6–13 offer a description of the ways that a speaker can control the production of sounds which differ in their perceptual quality. Chapters 14 and 17 describe the ways in which the units of speech can differ in terms of their **temporal** characteristics (duration, rate and continuity). Chapter 15 describes the **prosodic** attributes of speech (pitch and loudness). Chapter 16 then explores the way that facets of all four domains – quality, duration, pitch and loudness – are integrated to form the **metrical** structure of speech, in terms of its stress and rhythm.

2.3 The organic level

The next level of analysis to be discussed is the **organic level**. Speakers differ organically from each other in anatomical factors such as the dimensions, mass and geometry of their vocal organs. They will differ in such details as the overall length of the vocal tract; the volume and shape of the pharynx, mouth and nasal cavities; the nature of the dentition, the size and shape of the lips, tongue and lower jaw; the three-dimensional geometry of the structures in the larynx; and the volume and power of the respiratory system. People differ in the organic basis of their vocal apparatus no less than they do in the details of their facial appearance. That this should be so is less surprising than it might seem at first, given the anatomical connection that necessarily exists between the two.

More minor physiological factors may also play a role at the organic level, when the degree of muscular tonus changes for reasons outside the voluntary control of the speaker (such as excitement or fatigue), or when endocrinal or hormonal states have an effect on the detailed quality of the speaker's vocal performance. A speaker may produce utterances which differ to some small degree in these more minor physiological factors, while still being identical in linguistic and paralinguistic respects. Utterances from

a single speaker may also differ on a more striking organic basis when his or her state of health varies. A cold in the head, for instance, can often change a person's voice rather dramatically.

Two different speakers will almost always differ on the organic level, even though they may be judged to be producing linguistically and paralinguistically identical utterances. Only in the case of identical twins is it at all likely that the organic anatomical make-up of any two speakers could be more closely similar.

The extremely low probability of two speakers being organically the same, and therefore having the potential to produce very close copies of each other's utterances, has implications for the nature of imitative mimicry. Since mimics and impostors cannot change the organic basis of their vocal apparatus, even though they can to some extent compensate for organic differences by muscular adjustments, they are obliged to select individual features for imitation from the complex of features that characterize their target speaker. Mimicry is therefore necessarily a stereotyping process, not one of exact copying. The process of imitation, and by extension the process of a child's initial acquisition of patterns of speech, therefore depend on activities at the next level of analysis, the 'phonetic' level, which concerns voluntary, learnable movements of the vocal apparatus.

2.4 The phonetic level

If the organic foundation of a speaker's vocal apparatus endows him or her with quasi-permanent voice characteristics, the voluntary use the speaker can learn to make of the apparatus is much more plastic. The term **phonetic** will be used in this book to refer to any learnable aspect of use of the vocal apparatus. One could alternatively construe this as the aspect of speech under 'potential muscular control', but it is probably more helpful to insist on the 'learnability' of phonetic activity. This has the advantage of emphasizing the fact that the individual speaker acquires phonetic behaviour in a social context, guided by norms of phonetic similarity to the speech behaviour of the community in which the child grows up.

It is one of the most basic assumptions of phonetic theory that two organically different speakers should be able to produce phonetically identical utterances. This amounts to claiming that two spoken events can be phonetically identical but nevertheless sound acoustically different. The basis for maintaining this is that the phonetic level of description is abstract, not concrete. An assertion about **phonetic sameness** between two sounds is an assertion of comparability of particular abstract features in the sounds, rather than a claim about complete acoustic identity. To say that two sounds are

phonetically equivalent rests on an idealizing assumption that organic differences between speakers can be ignored in evaluating phonetic quality, as if both speakers could be held to be producing their performance on the same notional vocal apparatus. Phonetic sameness, which has to be conceded as a possibility before descriptive phonetic theory can work at all, is thus not a simple concept.

A speech event capable of displaying **phonetic equivalence** between speakers will be called a **phone**. The adjective from 'phone' will be taken in this book to be 'phonetic'. The notion of phonetic equivalence is one extreme of the scale of **phonetic similarity**. This concept of phonetic similarity is explored in detail in chapter 13.

Phonetic description is said to be based on the assumption that the process of description does not require knowledge about the formal, linguistic value that the event being described might have as a coded, communicative element in some particular language. In this sense, phonetic description of a given stretch of speech is held to be independent of the phonological description of the language involved. This forms the basis on which descriptive phonetic theory can be regarded as a general theory capable of application to the sounds of any language in the world.

It is not completely true, however, that the categories of phonetic description have evolved entirely independently of experience of typical correspondences between phonological units and their phonetic manifestations in languages in general. The structure of general phonetic theory is inevitably coloured to a certain extent by general phonological considerations. This point is explored further in chapter 4, in the discussion of the concept of phonetic units as basic descriptive categories.

2.4.1 *Phonetic notation*

One of the chief resources of phonetics is **phonetic notation**, as a set of written symbols used for transcribing the phones of actual pronunciation. In order to distinguish **phonetic transcription** from orthographic and other symbols, a widely followed convention will be adopted in this book of enclosing phonetic symbols in **square brackets**. The orthographic representation *seen* will be transcribed phonetically as [sin], for example. The detailed meaning of all the different phonetic symbols used will be introduced more fully in appropriate chapters. When an understanding of the phonetic meaning of a symbol is necessary for an appreciation of the theoretical point under discussion in this chapter, a brief description will be offered. Chapter 18 provides a general discussion of different types of transcription, and suggests principles for their classification.

2.5 The phonological level

The next level of analysis to be considered is the **phonological level**. Phonology is a very large and active subject with a very substantial literature, some of which is recommended in the section on Further Reading at the end of the chapter.

We have seen that the function of phonology is to relate the phonetic events of speech to **grammatical units** operating at the morphological, lexical, syntactic and semantic levels of language. Phonology is intimately connected with the phonetic study of speech – indeed it is not unreasonable to suggest that neither good phonology nor good phonetics is feasible without an adequate understanding of the other. For more comprehensive accounts of phonological theory as such, the references cited immediately above should be consulted. The central concern of this book is rather to give an account of a descriptive phonetic theory which will be of use in relating phonetic material to the study of phonology. The exposition given in this chapter on the relation between phonetics and phonology is therefore limited to an outline of the basic concepts necessary to understand the phonological roles of speech-sounds discussed in the later chapters on general phonetic description.

At the phonological level of analysis, two utterances are held to be different if the phonetic differences between them serve to identify the two utterances as representing different grammatical units of a given language. When two utterances are phonologically identical except for one distinctive difference, the two grammatical units involved are said to form a **minimal pair** of contrasting forms.

Limiting exemplification in the remainder of this chapter to grammatical units at the word level (**lexical units**), and to phonological units at the level of consonants and vowels, we can say, for instance, that a minimally different pair of words in English is *pan* and *tan*. The relevant phonetic difference that distinguishes them is that the consonant at the beginning of the first word is pronounced with the lips momentarily closing off the escape of the breath from the mouth, while the consonant at the beginning of the second word is pronounced with the tip or blade of the tongue sealing off the escape of the airflow. The two initial consonants, when they serve to distinguish different words in this way, are said to have a **contrastive function**, and to be in **opposition** to each other in the phonological code of which they are elements. The process of substituting one element for another to establish that an opposition exists in the language under examination is sometimes called a **commutation test**, and the elements which display such an opposition are said to have a **commutative relationship** with each other.

In order to be quite explicit about the level of analysis concerned, whenever phonological elements with a contrastive grammatical function are being discussed, their transcription will be enclosed in **slant brackets**. This convention is used very widely by linguists and phoneticians. The minimal lexical pair in English mentioned above, *pan* and *tan*, would be transcribed phonologically as /pan/ and /tan/ respectively. This means that comment on contrastive phonological elements involved, through the use of slant brackets, can be distinguished from the provision of detailed information on their pronunciation, for which square phonetic brackets are used.

Using the type of brackets as a signal of the level of analysis concerned allows not only a very desirable explicitness, but also provides a means for making concise statements about the phonetic manifestations of phonological elements. It is possible, for instance, to say that the word *pan* /pan/ in a given accent of English is pronounced as [p^{h}an], where the phonetic meaning of [^{h}] is that the onset of voicing for the vowel after [p] is audibly delayed for a fraction of a second. (Voicing, which is the buzzing sound which can be heard when the vocal folds in the larynx are made to vibrate, is fully explained in chapter 7.) A yet more compact version of the statement about the phonetic manifestation of the phonological elements would be the formula:

/pan/ ⇒ [p^{h}an]

where the arrow '⇒' (or its equivalent '→') stands for the relation 'is pronounced as', or 'is manifested phonetically as', or 'is realized phonetically as'. The formulaic device of the arrow will be used in the rest of this book in this realizational sense.

As with phonetic symbols, the meaning of phonological symbols will be explained in this chapter only if they are crucial to an understanding of the discussion. Otherwise they are merely illustrated here, and explained in detail in later chapters. Before going further, it is necessary to discuss three major facets of phonological organization. Appeal will be made to the notions of 'system', 'structure' and 'context'.

2.5.1 *Phonological system*

We have seen that there is a contrastive opposition in English between the initial consonants of the words *pan* /pan/ and *tan* /tan/. In fact, the list of consonants which are in potential opposition at this initial place in the set of such words in English is larger than merely /p/ and /t/. Other minimal pairs are provided by the availability, for example, of /k/ in *can*, /b/ in *ban*, /d/ in *Dan*, /m/ in *man*, /n/ in *Nan*, /f/ in *fan*, /v/ in *van*, /ð/ in *than* and

to illustrate their comments (*ibid.*). The transcription is enclosed here in square phonetic brackets because the exact phonological status of their transcription is unspecified.

Three-consonant syllable onsets in Serbo-Croatian

[smraad] 'filth'	[ʒdrijebac] 'stallion'
[svjedok] 'witness'	[ʒglob] 'joint'
[ʃtrik] 'rope'	[htjeti] 'to want'
[ʃkrina] 'box'	[gdje] 'where'
[zgneetʃiti] 'to crush'	[zdvojiti] 'to get together'

The classic examples of languages with very long clusters of consonants as permissible syllable-onsets are the languages of the Caucasus. In Georgian, a South Caucasian (or Kartvelian) language, syllable-initial clusters of consonants can range from two to six, as in the following examples (Catford 1977b: 292):

Multiple consonant clusters as syllable-onsets in Georgian

[prt͡s'kvna] 'to peel' [t͡s'vrtna] 'to train' [brt͡s'q'inva] 'to shine'

The sequence 't͡s'' in these Georgian examples functions as a single consonant in that language, so the word meaning 'to peel' has an initial consonant cluster of six consonants, the word for 'to train' has five, and the word for 'to shine' has four. The diacritic sign ['] means that the sound immediately preceding it is an 'ejective' sound; this is described in detail in chapter 6.

A vowel can be regarded as the **nucleus** of a phonological syllable. Every syllable in every language must obligatorily have a nucleus (though we shall see that in some circumstances in some languages the nuclear place in a given syllable may be filled by what some readers may prefer to call a consonant). However, even though every syllable must have a nucleus, it is not obligatory in every language that the syllable must have a consonantal onset. If the onset of a syllable is regarded as the initial place in the structure of the syllable, potentially fillable by one or more consonants, it is also possible for this place to be left empty (i.e. to be filled by zero consonants). Thus in English (RP), the syllable structure of a word such as 'end' /end/ can be represented as VCC (or ØVCC). Other examples of words with syllable structures showing zero-consonant onsets are:

Zero-consonant onsets to words in English

/ant/ 'ant'	/ilz/ 'eels'
/ist/ 'east'	/ɒks/ 'ox'

Some languages, such as Quileute, Coeur d'Alène and Puget Salish, which are all Amerindian languages, do not allow this vowel-onset structure, and

word-initial syllables all have a consonantal onset. The majority of the languages of the world, however, do permit zero-consonant onsets to word-initial syllables, as in the following illustrations from Hawaiian and Finnish (Sloat, Taylor and Hoard 1978: 64):

Zero-consonant onsets to words in Hawaiian
[ahi] 'fire' [ola] 'life'

Zero-consonant onsets to words in Finnish
[on] 'is' [iso] 'large'

Hawaiian, and Finnish differ from English, however, in that while their syllables may begin with either a consonant or a vowel, they cannot begin with more than one consonant.

English, as we have seen, allows open syllables and closed syllables. The number of final consonants in closed syllables in English can range from one to four consonants, as in 'sick' /sɪk/ (CVC), 'six' /sɪks/ (CVCC), 'sixth' /sɪksθ/ (CVCCC) and 'sixths' /sɪksθs/ (CVCCCC). In many West African languages, however, such as Etsako (Laver 1969: 49) and Urhobo (Kelly 1969: 157), only open syllables are phonologically permitted, as in the following instances:

Open syllables in Etsako
[uta] 'to say' [una] 'to run'
[usa] 'to shoot' [ozi] 'crab'

Open syllables in Urhobo
[se] 'read!' [so] 'sing!'
[si] 'write!' [hɔ] 'wash!'

In Etsako, the words are made up in these cases of two open syllables of structure V + CV. When words originally made up of closed syllables are borrowed into Etsako from another language, as in the case of a **loanword** borrowed from English with a structure -VC such as *bread* (CCVC), Etsako speakers regularize the syllable-structure pattern to conform to the rules of Etsako phonology. They insert an extra vowel in the appropriate places, thereby creating a CV+CV+CV syllable structure, to give the form /burɛdi/ (Laver 1969: 55). Firth (1948) refers to this process as **naturalization**, and gives the example of the English loanword *screwdriver* being naturalized in Hausa, a language of Northern Nigeria and sub-Saharan Africa, as [sukuru direba], in which all the syllables are made into open syllables. (The symbol 'r' in the above loanword examples in Etsako and Hausa is a general symbol for an 'r-sound', with no specific implication as to the exact phonetic real-

ization). A similar example comes from Turkish, where the English loanword *sport* is naturalized as [siˈpor] (Waterson 1956), because of sequential consonantal constraints in Turkish not allowing [s] and [p] to occur next to each other at the beginning of a syllable.

2.5.2.1 *Definition of a minimal pair of words*

Now that the concept of phonological structure has been established, the notion of a **minimal pair** of words can be amplified into a definition:

> *A minimal pair of words consists of two words of identical structure in a given accent of a language, which differ in the systemic choice made at only one place in that structure.*

While /pan/ and /tan/, both of CVC structure, form a minimal pair in English, /pan/ and /tap/ do not, though both are also of CVC structure, because they differ in more than one structural place. The words /span/ and /tan/ do not form a minimal pair either, because their structures (CCVC and CVC) are not identical.

2.5.3 *Phonological context*

When considering the characteristics of individual phonological units such as consonants and vowels, it is important to know in which contexts they can be found. One aspect of **context** is the structural position in which the unit can occur. This can be referred to as the unit's **structural context**. In English, for instance, /h/ can occur in syllable-initial position, as in *hat* /hat/, or in *perhaps* /pəhaps/ (structurally CV + CVCC), but nowhere else. The consonant /ŋ/, usually spelt *ng*, as in *hang* /haŋ/ and *hanging* /haŋɪŋ/ (structurally CVC + VC), can occur in syllable-final position, but nowhere else. Unlike either /h/ or /ŋ/, /p/ can occur in either syllable-initial or syllable-final position, as in *pip* /pɪp/, and so can /t/, as in *tot* /tɒt/, and /k/, as in *cook* /kʊk/.

Another aspect of context is the identity of elements adjacent to the individual phonological unit such as a given consonant or vowel. This can be referred to as the unit's **environment**, or **environmental context**. In terms of preceding environmental context of this sort (**left-context**), in single words in English /ŋ/ can only occur after vowels of relatively short duration. There are many examples of words like *sing* /sɪŋ/ and *long* /lɒŋ/. But there are no words in English which have relatively long vowels preceding /ŋ/, which would have to be spelt *seeng* or *lawng* or *voong*, for instance. Similarly, at the beginning of a word, /l/ can only occur either by itself, as in *long* /lɒŋ/, or after /p, b, k, g, f, s/, as in *played* /pleɪd/, *blade* /bleɪd/, *clay* /kleɪ/, *glade*

/gleɪd/, *flayed* /fleɪd/ or *slay* /sleɪ/. There are a few words used by English speakers which break this rule, such as *shlep* and *zloty*, but these are loan-words from other languages, and are signalled as such by having contextual shapes that are not a native part of the phonology of English.

In terms of succeeding environmental context (or **right-context**), in some accents of English /r/ is not able to be followed by another consonant within the same syllable, whereas in others (the majority) this is permitted. For example, the word *port* is pronounced /pɔt/ in RP and many accents of England, but /port/ in many American accents.

The importance of taking phonological context into account lies not least in the influence it has on the detailed phonetic manifestation of the phonological unit concerned. The pronunciation of a given consonant or vowel can be sensitive to either or both structural and environmental context. In English, /p/ has a number of different pronunciations depending on these two factors. As mentioned earlier, in initial position in a word like *pan* /pan/ ⇒ [p^han], there is a moment of relative silence after the lips open for the release of the consonant, when only the outflowing breath is audible (signalled by the [^{h}] in the phonetic transcription), before the pronunciation of the vowel can be heard. We shall use the phonetic term **aspiration** to refer to this detail of pronunciation, and the [p] can therefore be said to be **aspirated**. In *span* /span/, pronounced typically as [span], there is no such momentary silence, and the onset of audible voicing for the vowel is simultaneous with the opening of the lips. The pronunciation of /p/ is therefore said to be **unaspirated** in English in this position. Aspirated and unaspirated sounds are discussed at greater length in chapter 12.

Different adjacent elements have their effect also. The pronunciation of /p/ in *span* /span/ ⇒ [span] is different from that in *spoon* /spun/ ⇒ [sp^{w}un], with the lips being held momentarily in a more rounded position for /p/ in *spoon* [sp^{w}un] than for /p/ in *span* [span], (with the rounded position of the lips being signalled by the diacritic [^{w}] in the phonetic transcription). This is because in the pronunciation of the /p/ of *spoon* the muscles that control the position of the lips are being made to anticipate the rounded posture needed for the pronunciation of the following /u/ vowel. We shall use the phonetic term **lip-rounded**, for the moment, to refer to this detail of pronunciation. The precise area and shape of the opening of the lips varies with that of the sound whose anticipatory influence is being exerted on the sound in question, but this fine degree of detail is not usually annotated by using different transcriptional diacritics. Lip-rounding (under the more technical label 'labialization') is discussed at greater length in chapter 11.

2.5.4 *Phonological distribution*

An individual consonant or an individual vowel in the phonology of a given language can occur in different locations. This contextual concept has been expressed above in terms of the possibilities of location in different syllable structures and different environmental contexts. For any given consonant or vowel, the range of possible locations is constrained by the phonology of the language. This range of permitted locations constitutes the **phonological distribution** of that consonant or vowel. It is knowledge of constraints on such distributions that enables a speaker to decide that *zloty* cannot be a native word in English, because the sequence /zl-/ breaks the rules about contextual occurrences in syllable-initial structural position of the /z/ and /l/ consonants. The distributional range of contextual occurrences of a given consonant or vowel in an accent is said to constitute its **phonotactic range**.

The distributions of two consonants (or two vowels) are often compared, for purposes of describing the phonological function of the segments in question. When two consonants or two vowels share the same distribution exactly (that is, when they can be said to have exactly the same phonotactic range), they are said to be in **parallel distribution**, or to be **distributionally equivalent** (Lyons 1968: 70). When they share no locations, they are said to be in **complementary distribution**.

Consonants which are in parallel distribution are in contrastive opposition to each other, since by definition they are competing for occupancy of the same places in structure. Consonants which are in complementary distribution are not in contrastive opposition to each other, since by definition they are never in competition for occupancy of the same places in structure. There are also, however, many instances in languages where the consonants show an **overlapping distribution**, where some contexts are shared and some are not (Lyons 1968: 71). All the statements in this paragraph hold equally true for relationships between vowels.

In English, almost no two consonants (or two vowels) show precisely the same phonotactic range, and the condition of parallel distribution can therefore normally be achieved only on the more restricted consideration of overlapping distribution. Even /p/ and /b/ differ in their phonotactic range in English, though they are otherwise rather similarly distributed, in that /b/ cannot occur in syllable-final position after /m/, though /p/ can, as in *lamp* /lamp/ and *jump* /ʤʌmp/.

2.5.5 *The phoneme as a contrastive and distributed phonological unit made up of phonetically similar allophones*

The discussion above has made appeal to the notion of a given consonant (or vowel) participating in two axes of interaction with other con-

2.5.5.1 *Definitions of the phoneme and the allophone*

It is now possible to introduce two very widely used phonological concepts, that of the **phoneme** and the **allophone**. The notion of the phoneme has its roots in the alphabetic tradition of writing (Jones 1957), and there are almost as many definitions of the phoneme as there have been phoneticians and phonologists. It is a very convenient concept, but definitions of the phoneme should be understood as exemplary rather than theoretically strict.

In this book, a definition of the phoneme can be approached by invoking the phonological ideas introduced above of contrastive function, system, structure, context, parallel distribution, complementary distribution and overlapping distribution. The paradigmatic and syntagmatic axes of interaction described immediately above furnish the basis for the definition, which is presented here in two parts. The first part of the definition focuses on the issue of contrastiveness between one phoneme and another, and the second part on the issue of the contextually-distributed nature of the members of a phoneme conceived as a set of sounds.

The first part of the definition of a **phoneme** is as follows:

> *Two speech sounds are said to be manifestations of different phonemes in a given accent of a language when they act as the basis of a contrastive opposition that distinguishes a pair of words of identical phonological structure, differing in the systemic choice made at a single place in that structure.*

In order to satisfy these conditions, the two speech sounds must thus show parallel distribution over the context in question, in that both sounds must have the potential of occupying the single place in structure concerned. Phonemes can be transcribed using symbols enclosed in slant brackets, as in the transcription of the English word *cat* /kat/, and such **phonemic transcription** will be treated as an instance of a **phonological transcription**, as introduced earlier.

This first part of the definition is idealistic in that it makes appeal to the notion of minimal pairs of words of identical structure, whereas in any accent of a given language such minimal pairs may be only sparsely available, through the accidental circumstances of the lexical evolution of the language. English is relatively rich in such minimal pairs, as it happens. Where minimal pairs of words are not available, then the weaker condition of overlapping (rather than fully parallel) distribution over the relevant subpart of the words in question can apply.

The second part of the definition of a **phoneme** is as follows:

> *Speech sounds regularly occurring in a number of different structures and contexts may be classified as members of a given phoneme if their occurrences are in complementary distribution, and if they display sufficient phonetic similarity to make it plausible to class them together as members of a common set.*

The members of a given phoneme are called **allophones**. Allophones can be transcribed by using phonetic symbols enclosed in square brackets, as in the pronunciation of the English word *peep* as [pʰipˀ]. The [ʰ] element of this transcription means that, after the tiny but audible explosion when the lips separate at the end of the pronunciation of the first [p], there is a perceptible delay of about a twentieth of a second before the buzzing sound technically called **voicing** begins for the following vowel; this is the same process as exemplified by [ʰ] in the transcription of *pan* [pʰ an] discussed earlier. The [ˀ] element of the transcription means that the second [p] sound typically lacks this audible explosion, when it is in final position in an utterance. We shall use the phonetic term **unreleased** to refer to this detail of pronunciation. Unlike the unreleased pronunciation of [pˀ], the pronunciation of [pʰ] is **released**, as well as being **aspirated**. The concept of release of particular types of sounds is discussed at greater length in chapter 12. In this example, the two [p] sounds are in complementary distribution, in that under the rules governing the phonetics of English [pʰ] is appropriate for the pronunciation of the /p/ phoneme as a single consonant at the beginning of a syllable, while [pˀ] is appropriate for a pronunciation of the /p/ phoneme in utterance-final position.

It is important to note that the concept of an allophone is itself an abstract concept, and is not to be equated directly with that of a phone, which is a single differentiable phonetic event. The phonetic manifestation of a given allophone may vary slightly, on a random basis. On each occasion that a phonetic difference is perceptible, a different phone is involved. To take the example given immediately above, the [pʰ] element in the allophonic transcription [pʰipˀ] can on different occasions represent several slightly different phones, where the duration of delay in the onset of voicing is fractionally but audibly different.

A more extended example of allophones in a Southern British English accent which satisfy the conditions of complementary distribution and phonetic similarity mentioned in the definition of the phoneme, and which are therefore candidates for being grouped into a single phoneme, can be found in the following illustration. This shows different phonetic realizations of the phoneme /p/ (coincidentally always spelled 'p' in the orthography of English) in words such as:

peat /pit/ ⇒ [p^{h}it], where the [p^{h}] is aspirated, but where the lips are not rounded, because the lip position for the following vowel is unrounded;

pot /pɒt/ ⇒ [p^{hw}ɒt], where the [p^{hw}] is aspirated and lip-rounded;

spot /spɒt/ ⇒ [sp^{w}ɒt], where the [p^{w}] is lip-rounded but unaspirated;

pop /pɒp/ ⇒ [p^{hw}ɒ$p^{\neg}$], where the lips can in some circumstances at the end of an utterance remain closed rather than opening to release the compressed air built up during the final [p], so that the [$p^{\neg}$] is unreleased;

port /pɔt/ ⇒ [p^{hw}ɔt], where the [p^{hw}] is aspirated and lip-rounded (with a degree of rounding of the lips which is slightly more extreme for *port* than for *pot*, because the lip position for the following vowel has a smaller, more rounded opening in the case of *port*).

In the above cases, the pronunciations of /p/ show a strong degree of phonetic similarity to each other, in that they all involve the lips momentarily closing completely, with a simultaneous absence of voicing. They also show a variety of phonetic differences from each other (differences of aspiration versus non-aspiration, release or non-release, and differences in the presence and degree of lip-rounding).

These phonetic differences can all be described in the accent concerned as tied either to particular structural contexts or to statable environmental contexts. For example, in terms of responsiveness to particular structural constraints, the aspirated version of /p/ occurs in these English examples when it is a single consonant onset at the beginning of a word, and the pronunciation of /p/ is (optionally) unreleased when it is in a final position in the utterance. In terms of being subject to the influence of environmental context, the pronunciations of /p/ in the above examples show lip-rounding in anticipation of vowels whose performance requires a lip-rounded position (in *pot*, *spot*, *pop* and *port*). Equally, lip-rounding is absent from the pronunciations of /p/ when the following vowel shows no lip-rounding (in *peat*).

The different pronunciations discussed above can therefore be regarded as satisfying the requirements for allophonic grouping into a single phoneme. These requirements are, in summary, that the different pronunciations are complementary in their distribution, that they show close phonetic similarity, and that their phonetic differences can be attributed either to their structural position, or to interaction with their environmental context.

Daniel Jones gives many examples of allophonic groupings into phonemes

in his book on the phoneme (Jones 1950), and in many articles (e.g. Jones 1944a, 1957). The following illustrations are taken from Jones (in W.E. Jones and Laver 1973: 170):

> If one isolates the sounds of **g** in the words **goose** and **geese**, one hears them to be different, and one can feel that they have different tongue articulations: the second has a fronter articulation than the first. But from the point of view of the structure of the language the two sounds count as if they were one and the same. In the terminology I find it convenient to use they are 'members of a single phoneme'. One of the sounds is the variety appropriate in English to a following [u:], while the other is conditioned to the following [i:].

Jones gives another example, from Italian, which shows that the allophones of a given phoneme can show tangible differences in some of the phonetic features involved, so long as a substantial degree of phonetic similarity exists in the remaining features. He also shows that the classification of particular allophones into a given phoneme is specific to the given language only:

> In Italian and some other languages our English **ng**-sound occurs, but only in specific phonetic contexts, namely before [k] and [g]; it is used in these contexts to the exclusion of [n]. Thus the **n**'s in the Italian words **banca** and **lungo** are pronounced [ŋ], and an Italian never uses an ordinary [n] in such a situation. He is not like a Russian who uses [n] before [k] and [g] as well as in other positions. The result is that an Italian whose attention has not been called to the fact is unaware that the sound of **n** before [k] and [g] is in any way different from that of any other **n**. To him these nasal consonants are one and the same, and for all linguistic purposes they count as if they were one and the same – the sounds [n] and [ŋ] are 'members of the same phoneme' in Italian. (They are not so in English or German). (in Jones and Laver 1973: 171)

It has to be acknowledged that the part of the definition of the phoneme which relates to phonetic similarity is clearly more convenient than rigorous, and gives the individual analyst more latitude. But it is an important practical ingredient in the definition, none the less. It is the case, for instance, that the sounds which typically represent consonants are in complementary distribution with those normally representing vowels, in that they seldom compete for the same place in the structure of a syllable. If the only relevant criterion were one of complementary distribution, then it would be theoretically possible to collapse the manifestations of consonants and vowels together into single phonemes (and some phonologists have indeed made such proposals for some languages – Kuipers (1960), for instance, suggests that Kabardian, a language of the Caucasus, can be analysed on this basis

as if it had no independent vowel-phoneme at all). Most phoneticians would resist such a solution, however, not least because it contravenes a reasonable application of this second criterion of phonetic similarity. The definition is also convenient rather than rigorous in that the idea of 'phonetic similarity' is itself partly a matter of judgement on the part of the analyst as to what model of descriptive phonetic theory should be invoked, and therefore as to what should count as phonetically similar.

2.5.6 *Formulaic phonological rule notation*

The passage quoted above from Jones (1944a), on the status of [ŋ] in Italian as an allophone of /n/ in the context preceding /k/ or /g/, affords the opportunity to introduce a useful expansion of the formulaic device ' ⇒ ', explained in section 2.5 as meaning 'is realized phonetically as'. It is often convenient to be able to express contextually or structurally determined allophonic variation in a given language or accent as a formulaic rule, as introduced into linguistics by generative phonology. More extensive descriptions of the conventions of such rules are available in Clark and Yallop (1990: 156–69), Kaye (1989: 32–5) and Sommerstein (1977: 114–42). Formulaic rules of this sort give a compact way of summarizing the conditions under which phonological processes such as realization rules operate. Using A, B, C, D as notional symbols, the most general form for a realization rule is:

$$A \Rightarrow B / C \_\_ D$$

The single slash symbol '/' means 'in the context of'. The symbol '__' is called an **environment bar**, and is a device for helping to identify the structural or contextual location of the item on which the rule is to operate. In conjunction with a symbol placed to the left of the environment bar, left-context can be specified, and right-context is indicated by a symbol placed to the right. The rule as written above is thus interpreted to mean 'A is realized as B when A occurs in a sequence with C as its left-context and D as its right-context'. As well as this joint condition of both left- and right-contexts being prescribed, this formalism allows the specification of a more limited contextual condition, where the elements that trigger the realizational rule are simply the occurrence of C as left-context, with D null (i.e. zero), or the occurrence of D as right-context, with C null.

The entities in the formalism represented here by A, B, C and D may be phonemes, phonetic events, groups of phonological or phonetic features etc., but for our immediate purposes we shall limit them to phonemic and phonetic symbols. A version of the above rule might then read:

$$/X/ \Rightarrow [y] / /C/ \_\_ /D/$$

sonant system of English' and 'the vowel system of English' (or of Zulu, or of any given language). In the polysystemic approach, one has to speak in a more restricted and specific compass, of a given system operating at a particular place in the structure of a given type of larger linguistic unit such as the syllable, or the word. The monosystemic approach lends itself to a panoramic view of the sound-system of a language at large, and its transcriptional devices tend to be alphabetic in nature. The polysystemic approach perhaps favours a more microscopic attention to phonological and phonetic detail, and its transcriptional formulae are seldom devised with easy legibility in mind – that is not their prime purpose, which is rather to act as a powerful tool for professional phonologists.

A polysystemic approach of this sort has been associated with the work of J.R. Firth and his associates in the School of Oriental and African Studies in the University of London, in the approach usually called 'Prosodic Analysis' (though the term 'prosodic' is used in that approach in a somewhat different sense than in this book, mostly to refer to what we shall be calling 'features'). The single most influential paper in this tradition was probably Firth's own 'Sounds and prosodies' (Firth 1948), reprinted in Firth's collected papers (Firth 1957). There is only a limited opportunity in the present book to refer to the work of this school in any detail, but the serious student of phonetics and phonology is urged to become familiar with these publications, which explore a very wide range of oriental and African languages. Although most of the scholars actively subscribing to the Prosodic Analysis approach have now retired, the contributions achieved under the aegis of this tradition still reflect some of the theoretically most sophisticated ideas in the history of the British school of phonetics and phonology. In particular, their rejection of the validity of a monosystemic approach to phonology, in favour of a polysystemic perspective, is increasingly finding echoes in the work of modern phonologists.

The comments that follow in later chapters will mostly be based on the monosystemic approach, explicitly relying on the concepts of the phoneme and the allophone. It will sometimes be convenient, nevertheless, to consider the systems operating at different structural places separately from each other. To the extent that separate sub-systems are under consideration at any point, we will have moved away from a fully monosystemic approach towards a polysystemic approach.

2.6 Language-system and language-behaviour

It may be helpful to conclude this chapter on the relationship between phonetics and phonology with a brief mention of a basic but some-

what problematic distinction that readers are likely to encounter in their wider reading of textbooks on general linguistic theory. Different linguists use different terms, but the concept is long-established in modern linguistics. It is characterized by Lyons (1981: 10) as a distinction between a **language-system** and **language-behaviour**: 'A language-system is a social phenomenon, or institution, which of itself is purely abstract, in that it has no physical existence, but which is actualized on particular occasions in the language-behaviour of individual members of the language-community.' The distinction was originally introduced by de Saussure (1916 (1959)), in his differentiation of *langue* and *parole*, and was reinforced by Chomsky (1965) in a broadly comparable distinction between 'competence' versus 'performance'. De Saussure, in the fragmentary notes on his work collected by his editors Bally, Sechehaye and Riedlinger (de Saussure 1916 (1966)), gives only discursive descriptions of the concepts of *langue* and *parole*, but the essence of his concept of *langue* refers to the collective language-system shared by a community of speakers, and *parole* to concrete utterances produced in the act of speaking by an individual in actual situations (Crystal 1991: 194, 251). *Langue* is a social fact, while *parole* is an individual act, said de Saussure (1916 (1966): 13).

Chomsky (1965) called the tacit operational knowledge of the language-system possessed by typical speaker/listeners their linguistic **competence**, and called the exploitation of such knowledge by speakers and listeners in actual language-behaviour their linguistic **performance**. The description of linguistic competence, which many theoretical linguists take to be their primary task, concentrates on a highly idealized perspective of language. Chomsky and Halle (1968: 3), for example, describe the distinction between competence and performance in the following terms:

> The performance of the speaker or hearer is a complex matter that involves many factors. One fundamental factor involved in the speaker-hearer's performance is his knowledge of the grammar that determines an intrinsic connection of sound and meaning for each sentence. We refer to this knowledge – for the most part, obviously, unconscious knowledge – as the speaker–hearer's 'competence'. Competence, in this sense, is not to be confused with performance. Performance, that is, what the speaker–hearer actually does, is based not only on his knowledge of the language, but on many other factors as well – factors such as memory restrictions, inattention, distraction, nonlinguistic knowledge and beliefs, and so on. We may, if we like, think of the study of competence as the study of the potential performance of an idealized speaker–hearer who is unaffected by such grammatically irrelevant factors.

Chomsky (1965: 3) offered a more outright indication of the way that this distinction focuses the objectives of linguistic study: 'Linguistic theory is concerned primarily with an ideal speaker–listener, in a completely homogeneous speech community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance.' An important difference between de Saussure's *langue* and Chomsky's 'competence' is thus that while *langue* as a social concept emphasizes the institutional character of a language-system as a property of the whole language-community, 'competence' as a cognitively oriented concept refers not so much to the speaker-general language-system as such, but to the 'typical speaker's knowledge of the language-system' (Lyons 1981: 10). An intrinsic part of competence, for Chomsky and for the whole generativist tradition of linguistics, is that it reflects the fact that the generative 'capacity to produce and understand syntactically well-formed sentences is a central part – indeed, the central part – of a speaker's linguistic competence' (Lyons 1981: 234). The notion of competence in this respect is especially valuable in that the rule-governed ability which is the basis for producing and understanding a virtually infinite potential number of sentences begins to explain the 'creative' nature of language.

The distinction between competence and performance has been widely adopted by generative linguistics, though scholars such as Hymes (1971a) have argued for an enrichment of the concept of competence to include not only narrowly linguistic but a wider range of knowledge about other sociocultural communicative systems relevant to the use of language in different circumstances and for different purposes, in the notion of **communicative competence**. Others, such as Linell (1979: 17–26), have contested the psychological validity of the distinction.

The general concepts of language-system and language-behaviour are clearly useful for the discipline of phonetics. Phonetics has a part to play in the study of spoken language in both perspectives. It may be relevant, however, to signal some reservations about the full utility of these and related concepts for the discipline of phonetics. First, all three pairs of distinctions (language-system and language-behaviour, *langue* and *parole*, competence and performance) incorporate an assumption that it is sufficient to frame the concepts in terms of the typical speaker/listener, and to seek to characterize the linguistic attributes of the language-community. But phonetics, while being centrally interested in speaker-general phenomena in the same way, is also necessarily interested in the individual speaker/listener. Some

very important applications of phonetics, such as speech pathology and speech technology, have an absolute need to be able to relate the speech patterns of the individual, however idiosyncratic, to the generalized patterns that characterize their sociolinguistic community. This has implications for the need to study and define the language-system that can be taken to underlie the utterances of that individual speaker's language-behaviour, even if he or she is unique in using the systematic patterns concerned. From both a social and a theoretical point of view, it is in any case arguable whether the number of speakers of a given language is at all a relevant measure of the degree of interest that linguists should properly take in the properties of that language. In a discussion of how to choose the world's major languages for a descriptive survey, Comrie (1987: ix–x) offers a salutary comment on this point: 'When the world's linguists learned in 1970 that the last speaker of Kamassian, a Uralic language originally spoken in Siberia, had kept her language alive for decades in her prayers – God being the only other speaker of her language – they may well have wondered whether, for this person, *the* world's major language was not Kamassian.' General phonetic theory should have the capacity, so long as deviations from standard anatomy are not in question, of describing the articulatory strategies of production of any type of speech sound used by any speaker for communication through spoken language, not only by those who conform more closely to the norms of their sociolinguistic group.

Second, to address the specifics of the competence/performance dichotomy, if competence reflects the structure and content of a mental grammar and the operational knowledge necessary to generate sentences by means of this grammar, then any overall model of performance would need to incorporate a component representing this competence, as the quotation above from Chomsky and Halle (1968: 3) indicates. One view of the relationship between competence and performance is thus that the former is included within the latter. The highly abstract view of linguistic competence as taken by the generativist tradition has had the result of promoting a focused interest in a narrow definition of grammar, to the considerable benefit of formal aspects of linguistic study. But such a promotion of interest in competence at the expense of the wider concept of performance has had the inevitable disadvantage of holding performance-related issues other than those of competence somewhat at arm's-length. It remains true, nevertheless, as Chomsky (1965: 15) comments, that although 'there has been a fair amount of criticism of work in generative grammar on the grounds that it slights study of performance in favor of study of underlying competence', the study of competence itself 'can provide some insight into performance'.

Third, the true complexity of an adequately comprehensive model of the performance of utterances of spoken language is astonishing, and phonetics and linguistics are still very far from a full understanding of the cognitive and motor processes involved. The spoken performance of even the most simple sentence necessarily exploits the full depth of a speaker's operational knowledge of neurolinguistic and neuromuscular strategies of planning and executing ideational, pragmatic, semantic, lexical, syntactic, morphological, phonological and motor programs for the production of the utterance involved. It also exploits highly complex perceptual and neurolinguistic strategies of self-monitoring for the orthodoxy of what was actually said against what the speaker intended to say, and of planning and executing appropriate corrective material (Fromkin 1980; Laver 1991). The listener's perceptual and neurolinguistic performance in decoding the linguistic meaning of the acoustic material produced by the speaker is no less complex (Caplan 1987; Flores d'Arcais 1988).

The term 'neurolinguistic' is used here to emphasize the eventual need, at some stage in the modelling of how speakers and listeners process the production and perception of spoken language, to address the question of the real strategies and mechanisms used by the human neurological apparatus in achieving these operations. The more familiar term for the study of such processes is **psycholinguistics** (Garman 1990), reflecting what is sometimes a more abstract cognitive approach to the problem.

Psycholinguistic research in cognitive aspects of spoken language has tended to be atomistic, with individual researchers specializing in a given area, such as syntax, morphology, phonology, phonetics etc. Each of these areas can be hypothesized to be an (internally complex) module in the overall cognitive process which begins with the generation of some initial idea and ends in its eventual articulation as an audible utterance. There is now some evidence, however, that cognitive research into speech production and perception is beginning to address the challenge of modelling the possible links between the different modules, fashioning the beginnings of an understanding of the complete cognitive chain from ideation to articulation (Garrett 1988; Laver 1989).

Within the study of spoken language, phonetics as a discipline needs to be able to describe the phonetic substance of any communicative aspect of speech, as well as accounting for the formal aspects of the phonological level of description. It needs to be able to describe what real speakers actually do on real occasions of utterance, as well as to identify the social value of that performance as a manifestation of a sociolinguistic code. Phonetics thus has the responsibility of providing a well-founded description of the substance

3

Accent, dialect and language

The purpose of this third preliminary chapter is to introduce some further basic concepts, this time to do with the variations of accent, dialect and language between speakers, and of these aspects of language across space and time. The discussion will include issues in dialectology, sociolinguistics, linguistic geography and historical linguistics. It is necessary to say from the outset that this chapter thus touches on disciplines to which phonetics makes a contribution, but which in themselves are not traditionally central to phonetics as such. It is also the case that the contribution of phonetics to these disciplines has so far not been as extensive as it might profitably be, and there is considerable room for further phonetic research in these areas. The treatment of these neighbouring disciplines is inevitably presented in rather general terms, but readers are likely to encounter in their wider reading in linguistics all the speech-related technical concepts introduced here.

An important initial distinction needs to be introduced between 'accent' and 'dialect'. The technical meanings of both these terms differ from their everyday meanings. The technical meaning of the term **accent** is simply **manner of pronunciation**. In this sense, everyone speaks with an accent. Technically, it is not possible to speak aloud without speaking with an accent. The notion of accent is a phonological and phonetic concept, with some implications for the lexical level of analysis as well.

The technical meaning of the term **dialect** covers the types and meanings of words available and the range of grammatical patterns into which they can be combined. **Dialects** are discernibly different to the extent that they involve different morphological, syntactic, lexical and semantic inventories and patterns. A dialect can be expressed in either spoken or written form. In spoken form, a given dialect can often be associated with more than one accent. Many writers draw no specific distinction between the concepts of dialect and accent, but this book will maintain the distinction where feasible.

A **language** is the entity made up of a group of related dialects and their associated accents. From a technical point of view, it is not only impossible

to speak a given language without exhibiting some particular accent, it is also impossible to speak (or write) without exemplifying a choice of dialect, in terms of the vocabulary used and the sequences in which the words are combined. This book exemplifies a written dialect that is very largely common to educated use throughout the English-speaking world. To that extent, it could be referred to as **Standard English**. It is standard in the sense that it is a dialect understood and used in a relatively standard way by a very large number of people. The term 'standard' is here being employed in a **descriptive**, not a prescriptive mode. A **prescriptive attitude** would flavour the definition if 'standard' were to be interpreted to mean 'considered to represent correct and socially acceptable usage for educated purposes'. Such views about 'correctness' and 'social acceptability' would be undesirably self-regulating, in that it would be the self-selected people who use the dialect who would be ruling on its acceptability. Prescriptivist attitudes about 'correctness' have no place in a scientific approach to the description of speech that seeks to maintain objectivity.

3.1 The Standard English dialect, and non-regional and regional accents

An implication might be drawn from the comments above that Standard English is invariant over the English-speaking world. There are of course a number of detailed differences of usage between different national groups. British and American versions of Standard English, for example, will in fact be treated later as different dialects, for the sake of illustrating dialectal differences with familiar material. But such differences as exist between the various national forms of Standard English (American, Australian, British, Canadian etc.) are nevertheless small compared with the differences between these and other dialects of English. For the moment, then, it will be useful to maintain the convenient fiction that Standard English is a single dialect.

Since Standard English is by definition a dialect, unlike other dialects only in its unusually wide geographical scope, it follows that Standard English can be spoken with a variety of accents. Accents always mark the geographical origin of the speaker. Some accents mark regional origin very locally, and some mark only the fact that the speaker is an American, or is Australian, Canadian, English, Irish, Scots, Welsh etc. To the extent that a particular accent marks a speaker as having a very specific regional origin, it can be called a **regional** or **local** accent. Highly localized regional accents of this sort are also often called **broad accents**. To the extent that the speaker's local origin can be marked only within national boundaries, the accent can be said to be a **non-regional accent**.

A given speaker may be able to speak with two or more accents, with one accent often being broader than the other(s). Furthermore, a given speaker may well be able to use more than one dialect. A speaker native to Scotland will normally speak either the Standard English dialect with a regional Scottish accent, or a Scottish dialect of English with either the same or a broader Scottish accent, as the social occasion requires. Equally, a speaker native to Kentucky might speak Standard English with a regional Kentucky accent while working as, say, a receptionist for a multinational corporation, but speak a Kentucky dialect with a broader Kentucky accent when visiting relatives in his or her home town.

Received Pronunciation (RP) is an example of a non-regional accent of British English, which has been described in detail by very many authors, amongst whom Jones (1962) and Gimson (1962, 1989) are outstanding. The pronouncing dictionary specifying the details of RP which has been regarded for very many years as the most authoritative source is the one first published by Daniel Jones in 1917. This has been updated in many successive editions. Gimson, as one of Jones' senior colleagues, and the holder of the Chair of Phonetics at University College London immediately after Jones, was responsible for the 13th and 14th editions (Gimson 1967, 1977). This 14th edition was in turn the basis for a further revision by Ramsaran (1988). Another excellent and up-to-date pronouncing dictionary which describes RP (as well as a number of other accents of English) is Wells (1990). Wells maintains a historical continuity with Daniel Jones, being Gimson's successor in the Chair of Phonetics at University College London.

The term 'Received', in its Victorian sense of being 'received in polite society', gives a historical clue to the origins of the RP accent. According to Abercrombie (1965), RP developed as an accent of the English public schools (i.e. private schools, in the paradoxical usage of the English), and 'is maintained, and transmitted from generation to generation, mainly by people educated at [these] public schools' (Abercrombie 1965: 12). Wells (1982, vol. I: 10) summarizes these characteristics of RP by stating that:

> In England ... there are some speakers who do not have a local accent. One can tell from their speech that they are British (and very probably English) but nothing else ... It is characteristic of the upper class and (to an extent) of the upper-middle class. An Old Etonian sounds much the same whether he grew up in Cornwall or Northumberland.

Although the function of this accent as a marker of socioeconomic status is now a good deal weaker than previously, RP is still perhaps better regarded as marking social rather than narrowly regional aspects of identity

(Ramsaran 1990). The accent is non-regional in that the public schools themselves are regionally distributed throughout England (though it is true to say, demographically, that the population of RP speakers is denser in the south of England than anywhere else).

Petyt (1980: 29) suggests a descriptive link between the regional nature of accents and dialects and the socioeconomic status of the speakers concerned, such that the less regional variation is involved, the higher the status of the speaker. Crystal (1988: 63) comments that RP, like other accents, is subject to change with time:

> Early BBC recordings show the remarkable extent to which RP has altered over just a few decades, and they make the point that no accent is immune to change, not even the 'best'. In addition, RP is no longer as widely used as it was even 50 years ago. Only about 3 per cent of British people speak it in a pure form now. Most other educated people have developed an accent which is a mixture of RP and various regional characteristics – 'modified' RP, some call it, or perhaps we should talk about modified RP's, as in each case the kind of modification stems from a person's regional background, and this varies greatly.

Crystal also points out that the Victorian stigmatization of regional accents by educated British society is now fading, and that 'several contemporary politicians make a virtue out of their regional background, and the BBC employs several announcers with regionally modified accents. Nor is it uncommon, these days, to find educated people expressing hostility towards RP, both within and outside Britain, because of its traditional association with conservative values' (Crystal, *ibid.*). There are of course numerous other accents of British English, and a variety of different regional dialects within Britain. References on this topic are given in the recommended Further Reading at the end of the chapter.

Another example of a non-regional accent is the one that is often called **General American**. This name, according to Van Riper (1973: 232), was first promoted by Krapp (1925) and Kenyon (1930). Krapp characterized General American in the following terms: 'One may say that in America three main types of speech have come to be recognized, a New England local type, a Southern local type, and a general or Western speech covering the rest of the country, and all speakers in the South and New England at the moments when their speech is not local in character' (Krapp 1925, vol. I: 35, cited in Van Riper 1973).

Wells (1982, vol. I: 10) characterizes General American in the following terms, though with a reservation about the complete truth about the standard nature of the accent:

> In the United States, it is true not just of a small minority, but of the majority, that their accent reveals little or nothing of their geographical origins. They are the speakers of General American ... This is a convenient name for the range of United States accents that have neither an eastern nor a southern colouring; dialectologically, though, it is of questionable status.

The claim to generality of the term 'General' American is thus somewhat too sweeping (McDavid and McDavid 1956), but the term is now thoroughly established. In the *Random House Dictionary of the English Language* (1966), the phrase 'General American Speech' is defined as 'a pronunciation of American English showing few regional peculiarities: most U.S. radio and television announcers use General American Speech'. Bowen (1975) gives an outline of the pronunciations associated with General American. Kurath and McDavid (1961) offer an account of the accents associated with the dialects spoken along the eastern seaboard of the United States that are described in Kurath (1949). Kenyon and Knott (1944) cover the pronunciation of accents spoken in the Mid-West and Far West of the States in their pronouncing dictionary of American English. A general treatment of accents of English in both American and British English (RP) is given by Kurath (1964).

A large survey of Australian accents was conducted by Mitchell and Delbridge (1965). A commentator on the unusual uniformity of Australian English accents is Bernard (1969).

Canadian English is discussed by Avis (1986). Wells (1982) is an excellent and very detailed three-volume work on the accents of English round the world. Bailey and Görlach (1982) offer an edited collection of chapters on varieties of English dialects and accents in many different parts of the world, including Africa, Asia, the Pacific and the Caribbean as well as Britain, Ireland, Canada and the United States.

This book is centrally concerned with matters relating to pronunciation, rather than directly with higher-order linguistic phenomena at the level of dialects. But it is important to appreciate the distinction between dialect and accent, so in order to elaborate the distinction it may be helpful now to offer a few examples of specific differences between dialects, then of differences between accents, and then of differences within the accent of a single speaker speaking on different occasions. All the differences discussed below in this chapter fall within a definition of linguistic behaviour – differences due to paralinguistic changes of tone of voice signalling attitudinal and emotional information about the speaker are not considered here.

3.3 Differences of accent between speakers

Considerations of differences between accents invoke the concepts of system and structure introduced in chapter 2. It will be recalled that these concepts deal with the stock of phonological units used to distinguish different words in the language – in particular, with their number (the concept of 'system') and the permissible ways the phonological units can be arranged into sequences (the concept of 'structure'). We shall deal here with systems and structures involving phonological units only at the level of consonants and vowels – in other words the phoneme. It should be noted, however, that there are other types of (non-segmental) phonological units, to do with prosodic and temporal features such as intonation, stress and rhythm, which would have to be included in an overall scheme.

At the segmental level of consonants and vowels, the overall inventory of consonant phonemes of a given accent makes up its **consonant system**, and the vowel phonemes its **vowel system**. (It should be remembered, nevertheless, that while this monosystemic approach is convenient for summary presentation, it brings with it the concomitant disadvantage that it obliges us to collapse all the contextually specific sub-systems of such units into a global single system, with a corresponding loss of polysystemic detail). RP has 24 consonants in its consonant system, and 20 vowels in its vowel system. General American, which is less homogeneous than RP, has 24 consonants and 15 or 16 vowels, depending on which sub-type of the General American accent is being considered (Kreidler 1989). Most Scots accents have 25 or 26 consonants and 13 or 14 vowels. Figure 3.1 lists these consonant and vowel systems for the three accents, with illustrative words. Differences between accents which involve the number and type of word-differentiating phonemic distinctions available to the accents are called **systemic differences** (Laver and Trudgill 1979: 16).

Examples of systemic differences can be seen in a comparison between RP and typical accents of Scotland. RP speakers differentiate between *Sam* and *psalm* by having two vowel phonemes available in their vowel system, which yield the pronunciations /sam/ and /sɑm/. Most Scots have only one vowel at this place in their vowel system, so that in their accents the two words have the same pronunciation – that is, they are **homophones**. Similarly, RP distinguishes between *tot* and *taught*, and between *pull* and *pool*, whereas most Scots have the same vowel in each pair of words, such that *tot* and *taught* are homophones, as are *pull* and *pool*. On the other hand most Scots have two vowel phonemes in their system to distinguish words such as *tide* and *tied* as /tʌɪd/ and /taed/, whereas most RP speakers have only one, which can be transcribed as /taɪd/. In the case of such RP speakers, there is therefore

Consonant systems

Keyword	Scottish English	RP	General American
pea	p	p	p
tea	t	t	t
lock	k	k	k
loch	x		
bee	b	b	b
dye	d	d	d
guy	g	g	g
me	m	m	m
knee	n	n	n
song	ŋ	ŋ	ŋ
thin	θ	θ	θ
then	ð	ð	ð
fan	f	f	f
van	v	v	v
see	s	s	s
zoom	z	z	z
she	ʃ	ʃ	ʃ
beige	ʒ	ʒ	ʒ
each	ʧ	ʧ	ʧ
edge	ʤ	ʤ	ʤ
hat	h	h	h
lay	l	l	l
ray	r	r	r
yes	j	j	j
witch	w	w	w
which	ʍ		

Vowel systems

Keyword	Scottish English	RP	General American
bead	i	i	i
bid	ɪ	ɪ	ɪ
bed	ɛ	ɛ	ɛ
Sam	a	a	a
psalm		ɑ	ɑ
cot	ɔ	ɒ	
caught		ɔ	ɔ
pull	u	ʊ	ʊ
pool		u	u
mud	ʌ	ʌ	ʌ
bird	(ɪr)	ɜ	ɚ
word	(ʌr)		
heard	(ɛr)		
bay	e	eɪ	eɪ
side	ʌɪ	aɪ	aɪ
sighed	ae		
boy	ɔe	ɒɪ	ɔɪ
cow	ʌu	ɑʊ	aʊ
go	o	oʊ	oʊ
beer	(ir)	ɪə	(ir)
bare	(er)	ɛə	(ɛr)
poor	(ur)	ʊə	(ur)

taxi	e	ɪ	i
axis	ɪ		ɪ
taxes		ə	ə
china	ʌ		

Figure 3.1 Consonant and vowel systems of Received Pronunciation, Scots and General American English (after Giegerich 1992)

no distinction between these words, which are homophones with the same pronunciation /taɪd/. A minority of RP speakers do maintain a distinction between the two words, but it is a distinction based on vowel quantity (length) rather than vowel quality. The same vowel quality [aɪ] is used in both words, but the vowel in *tide* is pronounced with a shorter duration, and the one in *tied* with a longer duration.

Scots accents have at least one additional consonant in their consonant system compared with RP, in that where RP has only the consonant /k/, Scots has /k/ and /x/. For instance, *lock* and *loch* are pronounced identically in RP as /lɒk/, with a final /k/. All Scots accents pronounce *lock* with a final /k/, where the back of the tongue rises up to contact the underneath of the soft palate, completely blocking off the flow of air from the mouth for a moment. But the word *loch* is pronounced by Scots with a final /x/, for which the back of the tongue is brought close to the soft palate, though without completely closing off the outflow of air from the mouth, which creates a soft hissing/hushing sound technically called a 'fricative' (described in more detail in chapter 9).

Many Scots accents also have an additional consonant /ʍ/ in their systems (sometimes phonemically transcribed /hw/), which is used to distinguish the consonant at the beginning of *whine* /ʍaɪn/ from the one at the beginning of *wine* /waɪn/. The majority of RP speakers (not all) make no such difference, using only /w/. The phonetic difference between /ʍ/ and /w/, which are articulatorily identical, is that the former is voiceless and the latter voiced.

A **structural difference** between accents is a matter of the different rules governing the permissible sequences of phonemes in the phonemic make-up of word-shapes (Laver and Trudgill 1979: 18). It was mentioned earlier that a major structural constraint in RP is that /r/ is never pronounced before another consonant, whereas in very many other accents of English around the world, /r/ can be pronounced in this position. The type of accent where /r/ can be pronounced before consonants is called a **rhotic accent**, and the accents which do not allow this are called **non-rhotic accents**. Rhotic accents include all accents of Scotland, and all of the accents of the Mid-West and West of the United States. In addition to RP, many accents of England, the accents of Australia, New Zealand and South Africa, and many accents of the east and south of the United States are non-rhotic. On this, and related topics, see particularly Wells (1982).

The principal phonological differences between accents typically concern systemic and structural differences of the sort exemplified above. There is, however, a further type of difference of phonological relevance which deserves mention, though it is normally of a minor kind. This type of differ-

ence concerns the distribution of phonemic resources over the vocabulary of the dialect, in terms of which phonemes are selected in which words. Differences between two accents which are reflected in such variations of distribution are therefore called **selectional differences**. (These are also sometimes called 'distributional' differences, but to avoid confusion this term will not be used in this sense in this book.) A **selectional difference** is only relevant when the two accents in question share the same phonemic sub-system, but exploit the use of these phonemes differently in given words. A selectional difference can exist between two accents which are otherwise systemically and structurally identical, in the detail of the way that these phonological resources are lexically distributed across small sub-sets of the vocabulary (Abercrombie 1977: 22). An example of a selectional difference concerning vowels, taken from RP and some otherwise comparable accent, is the different distribution of the /a/ and /ɑ/ phonemes found in words like *pat*, *past* and *photograph*. The two accents could both distinguish between *pat* /pat/ and *past* /pɑst/ in the same way, but one might use the /a/ vowel of *pat* in *photograph* /ˈfoʊtəgraf/ where the other uses the vowel /ɑ/ of *past* in *photograph* /ˈfoʊtəgrɑf/.

An example of a selectional difference concerning consonants, taken from RP and General American, is the differing lexical distribution of the /h/ phoneme. General American has preserved the older (seventeenth-century) pronunciation /ɜb/ (or /ɜrb/) of the word *herb* without an /h/, whereas RP invariably uses the newer form /hɜb/. Another consonantal example is the RP pronunciation of *suggest* /səˈʤɛst/ where some American accents have /səgˈʤɛst/.

All the systemic, structural and selectional differences between accents that have been discussed so far have been differences at the phonological level, concerning the number, sequence and lexical distribution of phonemes characterizing a given accent. The final difference between accents to be discussed is at the phonetic level, and is to do with how speakers of the given accents actually pronounce the phonemes of those accents. Since this is a matter of detailed phonetic realizations, such a difference between accents can be called a **realizational difference**. It was said above that many Scots accents have only one vowel phoneme, which we shall transcribe /a/, in the area of the vowel system where RP has two – /a/ and /ɑ/ – in words such as *Sam* and *psalm*. The detailed phonetic realization of the Scots /a/ phoneme varies widely between accents: if the phonetic realizations of RP /a/ and /ɑ/ ([a] and [ɑ] respectively) were to be regarded as the end-points of a phonetic scale of vowel quality, then the qualities found as realizations of the Scots /a/ phoneme cover the whole of this scale. Where speakers of one Scots

accent pronounce both *Sam* and *psalm* as [sam], speakers of another may say the pair as [sɑm], those of a third may use a pronunciation of a vowel mid-way between the two, and those of a fourth may choose a quality nearer one of the end-points.

Realizational differences occur not only between accents, but also between speakers of a single accent who may otherwise share systemic, structural and selectional characteristics. A consonantal example of realizational differences within one accent of English can be found in the rather variable pronunciation of /s/ by different speakers of RP. Some speakers make this sound very high-pitched and piercing, even to the point of producing a momentary soft whistle. Others round their lips. Some (a small minority) raise their bottom lip slightly. Yet others curl the tip of the tongue slightly up and back, which has the effect of making the hissing sound less loud, and the apparent pitch of the hissing sound lower.

Some speakers have one or a small number of phonetic realizations that are sufficiently unlike those of the majority of their accent-group that they strike the listener as notably idiosyncratic, and strongly mark the identity of the speaker as an individual within the group. In extreme cases, such idiosyncrasy is perceived as a **speech defect**. But as a more general comment, it will be apparent on reflection that the phonetic realizations of every phoneme of every speaker have the potential of being slightly different from those of many other speakers even of the same general accent. A lifetime of settling to a habitual mode of speaking has given the personal accent of every speaker an individualizing realizational flavour, within the overall systemic, structural and selectional conventions of his or her own accent-group.

The overall range of phonetic variation between speakers within the accent of a single social or geographical group must not be so extreme as to threaten the classification of the speech patterns of the group as a single group-accent. As a set of working criteria, an accent can be regarded as a unified entity if its speakers share a relevant social or geographical attribute, and if the speakers of the accent successfully maintain a uniform set of systemic and structural characteristics, despite a certain amount of limited realizational variation between speakers. In addition, selectional variation must be very strongly limited. This matter is further discussed in section 3.5 below.

3.4 Differences of style and free variation within the accent of a single speaker

The speech of a single speaker on different occasions and with different conversational partners, within the confines of what would qualify as a single group-accent, may nevertheless differ in two further ways. The

increasing informality, it does seem to be the case that associated changes in pitch and loudness behaviour also occur, with pitch levels and ranges rising and loudness levels and ranges dropping (as a very broad generalization). Speech style is still a very under-researched area, however, and the ways in which pitch-patterns may change with increasing informality is not yet well documented for any language.

A final point to be made about the variations in the speech of a single speaker which are attributable to the style of speech used in different circumstances is that the phonological rules underlying casual, informal speech are often different from those applicable to formal speech. One example is the set of phonotactic rules which govern permissible sequences of consonants and vowels in syllable structure. The logical extreme of the reduction process mentioned above is deletion of individual segments from the pattern for individual words and phrases pronounced in isolation. This may give rise to syllable structures which are actually more rather than less articulatorily complex. Couper-Kuhlen (1986: 17) cites examples of this tendency from Bell and Hooper (1978: 18), commenting that '[In] casual speech even the phonotactic rules of a language may be violated: e.g. *potato* /pteɪtə/, *tonight* /tnaɪt/, *fatality* /ftælɪtɪ/', so that the syllable-initial consonantal sequences /pt-, tn-, ft-/ can be found in informal speech in English but not in formal speech. More strictly, one might say that the phonotactic rules governing informal speech are different from those of formal speech, rather than that informal speech 'violates' the rules of formal speech style – since informal speech is no less a legitimate expression of spoken language than is formal speech.

The second type of variation to be discussed in this section is **free variation**. This is where a speaker speaking in a single accent is free to choose between two or more forms of a particular word, without stylistic implications. One instance from RP is the free choice by a speaker on different occasions between the pronunciations /ivəˈluʃən/ and /ɛvəˈluʃən/ for the word 'evolution'. This is an example where the choice lies between two of the vowel phonemes available in the speaker's vowel system. Free variation also occurs on a microscopic scale at the phonetic level within the speech of every individual speaker, where the exact quality of the realization of any given phoneme in a given context displays small phonetic differences (on an apparently random basis) on different occasions of utterance (Gimson 1962: 47).

3.5 Accent as a marker of the speaker's group-membership and individuality

The accent of a given speaker can be viewed in (at least) two perspectives. It can be seen as a **group-marker** of the speaker's membership

of a variety of social groupings, and as an **individuating marker** uniquely identifying the speaker against the mass of other members of the wider group. A term used for an accent which characterizes the speech of a whole social group when speaking in a given style is **sociolect**. Within such a sociolect, the uniquely idiosyncratic accent of a given speaker speaking in a given style is called the speaker's **idiolect**. The term **lect** is then a neutral term for an accent without specific implications for its sociolectal or idiolectal status.

Comment on a sociolect considers the speaker as a member of a particular social group. Comment on an idiolect focuses on the speaker as an individual within his or her social group. A speaker-identifying idiolect typically relies for its individuating power on fine details of phonetic realizations. A group-identifying sociolect is usually chiefly signalled by more systemic, structural and selectional considerations.

Accent is a rich source of inference for the listener about the social attributes of the speaker (though it should be remembered that these judgements by the listener are not always reliably accurate). The syllable structures of the accent, together with the consonant and vowel systems, the lexical distributions of the phonemes and their phonetic realizations, are taken to mark the speaker's membership of a particular sociolinguistic community. As we saw earlier in this chapter, this marking can vary in its degree of geographical specificity, depending on whether the sociolinguistic community is non-regional (as in the case of the General American and Received Pronunciation accents), or regional (as in the case of Texan or Cockney accents).

Accent is also taken to mark a range of social attributes other than merely that of geographical origin. Wells (1970: 248), for instance, comments that 'Accent is still significantly linked to social class in England, and often constitutes an important index of class affiliation.' One particular facet of phonetic performance that serves this function in the accents of England is the detailed quality of the pronunciation of vowels. We shall see in chapter 10 that the pronunciation of vowel quality is an articulatory, auditory and acoustic continuum, and this makes it available to act as a marker of fine gradations of social status (Giles, Scherer and Taylor 1979: 361; Laver and Trudgill 1979).

The work of Labov (1966, 1972a, 1972b) has been very influential in the sociolinguistic study of speech, notably in understanding the role of accent as a sociolinguistic marker. In particular, Labov promoted the understanding that speakers' accents are not fixed and unchanging, but show many linguistic variables which can be manipulated (consciously or unconsciously) as

an adaptive response to the changing social situation in which the speakers find themselves. The changing social factors include such details as the relative social status of the speaker and listener, the familiarity of their acquaintance with each other, whether both participants are of the same sex and age, whether the conversation is within the hearing of bystanders, and the nature and purpose of the meeting itself. One of the responses by the speaker to these changing circumstances can be the degree of informality of the speech style adopted.

An example of sociolinguistic variability within an accent is discussed in the following comments:

> Work in the field of sociolinguistics has shown that social groups may differ not simply in terms of their phonological systems or what pronunciation they use but in *how often* they use certain pronunciations. In the pioneering work in this field, Labov (1966) demonstrated that in New York City speech the pronunciation of /r/ in words such as *farm* (where the *r* occurs before another consonant) and *far* (where it occurs word-finally) is a *linguistic variable*. While very few New York City speakers actually pronounce the /r/ in these positions on every occasion, most pronounce it on some occasions ... In New York City the frequency with which /r/ in *far*, *farm*, etc. is pronounced correlates very clearly with the age and social class of the speaker, and with the social context in which he is speaking ... In the New York City study, counts of how many /r/s were actually pronounced, and how many could have been but were not, showed that middle-class speakers, on average, pronounced a higher percentage of /r/s than working-class speakers, that younger speakers had a higher percentage than older speakers, and that formal styles of speech produced more /r/s than informal styles. Interestingly enough, studies of English towns where /r/ is also a variable, such as Reading, show exactly the reverse pattern (see Trudgill 1975) – a good illustration of the arbitrary nature of linguistic markers of social categories. (Laver and Trudgill 1979: 19–20)

Many comparable sociolinguistic studies of individual cities in the English-speaking world have been stimulated by Labov's (1966) work on the sociolinguistic patterns of lower east side New York speech. A number of publications describing sociolinguistic work on communities in North America, Great Britain and Autralia are given in the section on Further Reading below.

3.6 Dialectology and dialect geography

The customary name for the discipline that studies the phonetic and phonological patterns of accents and dialects of languages, and their

geographical distribution in space and time, is **dialectology** (Chambers and Trudgill 1980). It should be said that the general use of this term is not consistent with the practice advocated in this book, where 'dialect' is reserved for description at a higher linguistic level than pronunciation. But 'dialectology' is so well established as a label that it would be counter-productive to insist on a division between 'dialectology' and 'accentology' or the like. Furthermore, the established scope of the term 'dialectology' encompasses more than matters of accent alone, in that it also deals with dialectal patterns of lexical, syntactic and morphological usage. Walters (1988) offers a compact and valuable account of the history and the methods of both the traditional and the more modern sociolinguistically and quantitatively inclined versions of dialectology in North America and Europe.

A very substantial body of work in dialectology has been invested in the mapping of the geographical distribution of different dialectal forms. This part of the subject is therefore sometimes called **dialect geography**, or **regional dialectology**. The traditional practice in dialect geography was for fieldworkers to conduct face-to-face interviews with speakers they had selected as authentic representatives of the sociolinguistic community they were investigating, and to transcribe the utterances the informants produced for the dialectal forms for objects or activities listed in a standard questionnaire or wordlist, in a phonetic notation whose conventions varied somewhat from one project to another. Cross-project comparability was hence somewhat impeded. The process of gathering adequate data for large-scale projects was typically very expensive in both time and effort, and organizing and preparing the data for publication correspondingly difficult. It is perhaps not then surprising that few large-scale projects have succeeded in reaching final publication.

Carver (1987), Chaika (1989) and Walters (1988) consider the history of dialect surveys in North America. A survey of the dialects of the United States and Canada began in 1931 that sought to register both phonetic and linguistic data of dialect variation (Atwood 1963). The project was entitled the Linguistic Atlas of the United States and Canada (LAUSC), and it accomplished a very significant amount of dialectological work. But it was eventually abandoned in 1949, as simply not feasible. McDavid (1958) described the very ambitious objectives of that project in the following terms: 'the American Atlas seeks to record data illustrating the social differences, the dimension of time, and the process of language and dialect mixture that has been going on everywhere since the New World was settled' (cited by Chaika 1989: 226). Walters (1988: 121) describes the history of the

LAUSC project and related dialectological projects in North America as follows:

> In the United States and Canada, dialect geography is usually associated with the *Linguistic atlas of the United States and Canada (LAUSC)*, an as yet unrealized effort to produce a linguistic atlas for each of the regions of the two countries. The project was begun in the late 1920's in New England, yielding the *Linguistic atlas of New England* (*LANE*) (Kurath *et al.* 1939–43) and the *Handbook of the linguistic geography of New England* (Kurath *et al.* 1939). None of the other atlas projects has fared so well, with only the *Linguistic Atlas of the Upper Midwest* (*LAUM*) (Allen 1973–6) and a few fascicles of the *Linguistic atlas of the Middle and South Atlantic States* (*LAMSAS*) (R. McDavid *et al.* 1980–) reaching publication ... The single exception is the *Linguistic atlas of the Gulf States* (*LAGS*) (Pederson 1969, 1971, 1974, 1976; Pederson, McDaniel and Bassett 1984).

A current project (*DARE*, the *Dictionary of American Regional English*) is an on-going attempt to map the lexical forms of the dialects across the whole of the United States, comparing these forms with those recorded in the survey begun in 1931 (Cassidy 1985). Kurath (1972) gives a brief overall picture of the dialect geography of the whole of the United States.

A major project on lexical and phonetic aspects of dialects and accents in Scotland and nearby areas is the Linguistic Survey of Scotland, whose original design was outlined by McIntosh (1952), which resulted in the *Linguistic Atlas of Scotland* (Mather and Speitel 1975). As an illustration of such work, figure 3.2 is a reproduction of their map showing the distribution in different parts of Scotland, Northern Ireland, Orkney and Shetland of the words for 'splinter' in a number of different dialects (*skelf*, *stab*, *splice*, *stob*, *splinter*, *spell*, *spale*, *skelve*, *spelk*, *spilk* and *skelb*) (Mather and Speitel 1975: 33). The Linguistic Survey of Scotland includes an important Gaelic section, whose continuing work is described in Jackson (1958) and Gillies (1988). Wagner (1964–9) constitutes a monumental survey of the dialects of Irish. A comparable major work on the dialects of England is the *Linguistic Atlas of England*, published by Orton (1962), Orton and Barry (1969–71), Orton and Halliday (1962–3), Orton and Tilling (1969–71) and Orton and Wakelin (1967–8), and by Orton, Sanderson and Widdowson (1978). A dialect survey of Wales is reported by Thomas (1973).

One of the tools of the dialectologist is the isogloss. An **isogloss** is a line on a map drawn round all areas displaying the same pronunciation for a given consonant, vowel or word. A strong implication underlies the use of such a definite boundary – namely, that the phonetic and phonological characteristics of lects on one side of the isogloss are categorically different from

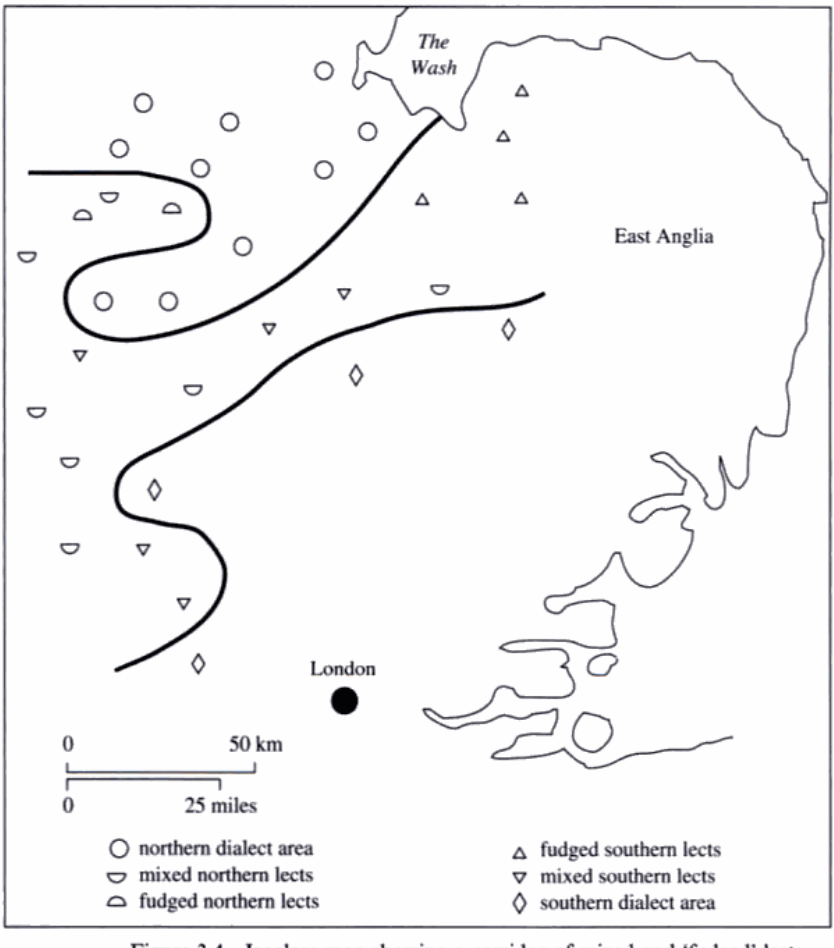

Figure 3.4 Isogloss map showing a corridor of mixed and 'fudged' lects with differing patterns of usage of [ʊ], [ʌ] and [ɤ] in words such as *dust* and *push* (adapted from Trudgill 1983:49–50)

applicable – speakers may vary in how often they choose particular pronunciations within their range of possibilities. The second is that such geographical accent-boundaries are more often gradual than discrete. Figure 3.4 is taken from Trudgill (1983: 49–50), where he comments on previous work by Chambers and Trudgill (1980), to demonstrate that between the northern and southern zones there is actually a corridor of mixed and 'fudged' lects, which forms a transitional zone. He characterizes this 'corridor of variability' as containing the following mixed lects:

(i) mixed northern lects: /ʌ/ and /ʊ/ alternate, but /ʊ/ predominates;
(ii) fudged northern lects: /ʊ/ alternates with an intermediate vowel [ɤ];
(iii) fudged southern lects: /ʌ/ alternates with an intermediate vowel [ɤ];
(iv) mixed southern lects: /ʌ/ and /ʊ/ alternate, but /ʌ/ predominates. (Trudgill 1983: 49)

The symbol [ɤ] that Trudgill refers to as an 'intermediate vowel' represents a pronunciation whose tongue position is close to that of [ʊ], but with a lip position like that of [ʌ]. The detailed qualities of these sounds are discussed in chapter 10.

3.7 Language differences between speakers

The dialects of a given language can vary in their similarity to each other. Conventionally, the limit that is set to the maximum degree of difference between two dialects is one of **mutual intelligibility**. When two forms of language are mutually unintelligible, they are usually considered to be dialects of two different languages. Any two dialects of English and German, for example, though showing vestigial similarities that reflect their common genetic status as closely related languages in the Indo-European family of languages, are unintelligible to listeners of the other language. But there are many instances where the application of the criterion of mutual intelligibility fails to yield a satisfactory classification of the two language forms concerned. Norwegian and Danish are regarded as different languages, and yet their speakers can hold relatively successful conversations despite the grammatical and phonetic differences that exist. The factor that is decisive in Norwegians and Danes preferring to treat them as different languages rather than different dialects of the same language is less linguistic than political, with differences of language being identified with national differences.

There are other problems with the criterion of mutual intelligibility. One is that intelligibility is a multidimensional scale, when considered within and across the full range of linguistic levels. Setting a limit within such a scale for what counted as a dialect-differentiating versus a language-differentiating degree of intelligibility would be quite arbitrary. Another problem is that matters of intelligibility are themselves subject to the judgemental attitudes of the listeners involved. There are numerous cases where the listeners of language form A find the speakers of language form B quite intelligible, but where the listeners of B claim not to be able to understand the speakers of

A. Once again, factors of social and political attitude dominate the question. It is therefore probably more practical to consider the differentiation of dialects and languages to be primarily a sociopolitical issue. One can perhaps accept the broad proposition that, in general, languages differ linguistically from one another in the same ways that dialects do, even if usually to a more extreme degree.

The concept of intelligibility itself is not a little problematic. A listener who understands French and Italian may be able to make partial sense of utterances heard from speakers of Spanish, because of the general relatedness of the three languages as members of the Romance language-family. But it would not be helpful to suggest that such a listener 'understands Spanish'. The field of cross-linguistic intelligibility offers interesting research possibilities.

3.7.1 *First, second and foreign languages*

Very many speakers, perhaps the majority of speakers in the world, understand at least one language other than their own. A speaker who understands only one language can be described as a **monoglot**, and one who understands two or several a **polyglot**. However, one could be basically a monoglot with only a smattering of knowledge of words and phrases in another language, without qualifying as a polyglot. To clarify the different status of a speaker's experiences of other languages, it is helpful to differentiate between a first language, a second language and a foreign language. A **first language** is a speaker's **native language**, or **mother tongue**, whose learning normally begins in the speaker's earliest experience of language-acquisition as a very small child. A **second language** is any other language that the speaker learns to control, at any time, to a level of near native-like proficiency. Few speakers ever learn to control a second language to the point of being indistinguishable, in every linguistic and phonetic respect, from a native speaker of that language. If they do succeed in attaining this level of expertise, then they can properly be termed **bilingual**. True bilingualism in this strict definition is fairly rare, and most speakers of languages other than their native language show signs of their non-native status in their linguistic and phonetic performance. Any language spoken by a speaker to less than second-language level can be called a **foreign language**, and the state of being able to control more than one language (to any reasonable degree) can be referred to as **multilingualism**. In the literature of applied linguistics, the term bilingualism is often used in this looser sense of multilingualism.

Learning a language other than one's own native language is always a process in which the patterns of the first language interfere with the learning

of the foreign language. At any given stage of acquisition of a foreign language, the linguistic competence underlying the utterances performed by the learner will show some amalgam of the patterns of the two languages. Selinker (1972) was amongst the first to develop the concept of an **interlanguage**, to account for the strategies deployed by learners of a foreign language at any given stage in their progressive mastery of that language.

An area where multilingualism is a prime consideration is the teaching of languages. This is a domain where acronyms abound. In the case of English, one encounters acronyms such as **ELT** (English Language Teaching), which is a relatively comprehensive term for teaching English to learners of all types; **TESL** (Teaching English as a Second Language), where the learners addressed are often immigrants to an English-speaking culture; **TEFL** (Teaching English as a Foreign Language), where the learners are normally neither native speakers nor immigrants; and **TESOL** (Teaching English to Speakers of Other Languages), which is a slightly more neutral term encompassing both TESL (mostly) and TEFL, but avoiding some of the potentially prejudicial implications of labels such as 'second language' and 'foreign language'. Accounts of these areas of English-language teaching can be found in Stern (1983).

The very brief sketch above of terms applicable to speakers who understand more than one language does not do justice to the diversity of the topic of multilingualism. In the broad definition as a person's ability to control two or more languages, even to an unequal degree, multilingualism (or 'bilingualism') is currently an intensive research area in applied linguistics (Spolsky 1988).

The deviations in performance of a foreign language by a speaker from native-speaker-like norms of pronunciation are almost all directly attributable to the influence of differences between the phonetic and phonological patterns of the foreign language and the speaker's own native language. The concepts introduced in section 3.3 of structural, systemic and realizational differences between two accents of a language can equally be applied to this situation of speakers trying to pronounce an accent of a language other than their own. When French students beginning to learn English pronounce the phrase *this red cart* as [zis ʁɛd̪ kɑʁt̪], normally pronounced by native speakers of RP as [ðɪs ɹɛd kɑt], the replacement by the learners of the RP [ð] by [z] in *this* reflects the systemic fact that while the consonant system of RP distinguishes two phonemes /ð/ and /z/, that of French uses only /z/. Similarly, where the RP vowel system distinguishes between /ɪ/ and /i/ (as in *hid* /hɪd/ and *heed* /hid/ respectively), French uses only /i/. The substitution of the RP [ɹ] by [ʁ] has a realizational explanation, in that while the consonant systems

of both RP and French include an /r/ phoneme, its realization in RP is (in this syllable-initial position) [ɹ], with the tip of the tongue slightly curled, while that in French is [ʁ], which is made by bringing the back of the tongue near to the end of the soft palate at the back of the mouth. The French learners' use of [d̪] for [d], and [t̪] for [t], also corresponds to realizational differences between the two accents, in that [d̪] and [t̪] in French are made partly by sealing off the flow of air through the mouth by the tip of the tongue against the upper front teeth, while [d] and [t] in English are characteristically made by the seal being made a little further back, on the ridge just above and behind the teeth.

The French insertion of [ʁ] before the consonant /t/ at the end of 'cart', as a pronunciation of an /r/ phoneme thought by the French learners to be implied by the spelling, is a structural matter. Where the RP accent of English is non-rhotic (i.e. excluding the pronunciation of /r/ before consonants, as described in section 3.3), all accents of French are rhotic.

Comparisons in such terms of the phonetic and phonological patterns of the target accent and that of the learner allow the teacher to predict typical errors likely to be made by beginners. They also enable the teacher then to plan appropriate strategies of pronunciation teaching, and to prioritize them in importance and sequence of learning. Two excellent books that have been recently published on learning English pronunciation based on this general approach are Gimson (1989, updated by Ramsaran) and Roach (1991).

3.7.2 *Lingua francas, pidgins and creoles*

Foreign languages are most often used for communication with speakers for whom the language concerned is their native language. A special case is where the language used is the native language of neither the speaker nor the listener, but where it may be the only medium of linguistic communication in common between the two, often for trade purposes. Where this practice is widespread in a given area, the status of the foreign language as an instrument of general communication is recognized by calling it a **lingua franca**.

The orginal 'lingua franca' (literally 'language of the Franks', the Arabic term of the day for all Europeans) came into being in the Eastern Mediterranean at the time of the Crusades about nine centuries ago, and evolved as a composite of Italian, Provençal, French, Spanish and Portuguese (Holm 1989: 607). Lingua francas in use today (or 'lingue franche', to recognize the phrase's Italian origin) notably include English and Mandarin Chinese. Other large-scale lingua francas are Malay in Malaysia and Indonesia, Swahili in East Africa, Hausa in West Africa,

The definitions of a pidgin language quoted above from Holm and Hall specified that a pidgin is native to no speaker. When a pidgin is acquired as the first, native language of a group of speakers (as happened historically to generations of speakers under the social and geographical displacement that accompanied slavery), it is said to constitute a **creole language**. Creole languages typically become more linguistically elaborated than their antecedent pidgins, and become autonomous languages in their own right. The evolution of creole languages, in the process called **creolization**, has claimed a good deal of recent research attention as an arena of creative language acquisition claimed to reveal universal aspects of the language faculty (Hymes 1971b; Romaine 1988). The potential for sociolinguistic research in this topic is also very substantial.

A large number of languages have contributed to the basis of modern creole languages via their parent pidgins, including Arabic, Dutch, English, French, Portuguese, Spanish and Swahili. Hancock (1971: 507–23), reprinted as an Appendix by Romaine (1988: 315–25), lists more than 200 pidgin and creole languages around the world. Well-known examples of creole languages based on English are Krio in Sierra Leone and Jamaican Creole in the Caribbean.

Holm (1989: 475) gives the following illustration of a less-known creole language from the Western Caribbean, Miskito Coast Creole English (the Miskito Coast runs from the northeastern coast of Honduras to the eastern coast of Nicaragua), where Creole English acts as a lingua franca for the various ethnic groups of the population:

Miskito Coast Creole English text

wen i pik it op naw i no kom we a de.
when he starts up (drinking) now he doesn't come where I am
i tel mi lay
he tells me lies

wen i kom naw. da iyvnin i sey, 'mama,' i sey, 'a
when he comes now in the evening he says mama he says I
did tayad an
ANT tired and

neva kom'. a sey, 'yu dam lay. siy yu ay? yu mi
didn't come I say you damn liar see your eye you ANT
dringkin; das wai
drinking that's why

yu no mi wahn kom ya.'
you NEG ANT want to come here

Underlying the apparent diversity shown by the thousands of mutually incomprehensible languages of the world, there is a remarkable, elegant and principled unity in the way that these languages exploit the phonetic resources of speech. It is these principles that Professor Laver sets out to describe in this major new textbook. Assuming no previous knowledge of the subject, it is designed for readers who wish to pursue the study of phonetics from an initial to an advanced stage, equipping them with the necessary foundations for independent research. The classificatory model proposed unifies the description of linguistic, paralinguistic and certain extralinguistic aspects of speech production. The book moves from a presentation of general concepts to a total of eleven chapters on phonetic classification, and it includes discussion of other issues such as the relationship between phonetics and phonology, the nature of accent, dialect and language, and the description of voice quality and tone of voice. Every descriptive category is illustrated by words in phonetic transcription, from over 500 of the world's languages.

Principles of phonetics will be required reading and an invaluable resource for all serious students and scholars of speech and language.

CAMBRIDGE
UNIVERSITY PRESS
www.cambridge.org

ISBN 0-521-45655-X

9 780521 456555