



THE PSYCHOLOGY OF LANGUAGE

FROM DATA TO THEORY

TREVOR HARLEY

THIRD EDITION

SAMPLE CHAPTER

THE PSYCHOLOGY OF LANGUAGE

The Psychology of Language

From Data to Theory

Third Edition

Trevor A. Harley

 Psychology Press
Taylor & Francis Group
HOVE AND NEW YORK

First published 2008 by Psychology Press, an imprint of Taylor & Francis
27 Church Road, Hove, East Sussex BN3 2FA

Simultaneously published in the USA and Canada by Psychology Press
270 Madison Avenue, New York, NY 10016

*Psychology Press is an imprint of the Taylor & Francis Group,
an informa business*

© 2008 Psychology Press

All rights reserved. No part of this book may be reprinted or reproduced or utilized in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

British Library Cataloguing in Publication Data

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

Library of Congress Cataloging-in-Publication Data

Harley, Trevor A.

The psychology of language : from data to theory / Trevor
Harley.—3rd ed.

p. cm.

Includes bibliographical references and indexes.

ISBN 978-1-84169-381-1—ISBN 978-1-84169-382-8

1. Psycholinguistics. I. Title.

BF455.H2713 2008

401'.9—dc22

2007021225

ISBN 978-1-84169-381-1 (Hb)

ISBN 978-1-84169-382-8 (Pb)

Typeset in 10/12pt Times by Graphicraft Limited, Hong Kong

Printed and bound in the UK by Ashford Colour Press Ltd, Gosport, Hampshire

*For Siobhan, without whom everything
would be impossible*

Contents

Preface	xi	Where did language come from?	51
How to use this book	xv	Do animals have language?	54
		The biological basis of language	67
		The cognitive basis of language:	
		The cognition hypothesis	79
		The social basis of language	82
		The language development of	
		visually and hearing-impaired	
		children	84
		What is the relation between	
		language and thought?	87
		Summary	99
		Some questions to think about	100
		Further reading	100
SECTION A: INTRODUCTION	1		
1. The study of language	3	4. Language development	103
Introduction	3	Introduction	103
What is language?	5	The driving forces of language	
The history and methods of		development	105
psycholinguistics	9	Do children learn any language	
Language and the brain	14	in the womb?	119
Themes and controversies in		Phonological development	120
modern psycholinguistics	19	Lexical and semantic development	125
Summary	24	Early syntactic development	136
Some questions to think about	25	Summary	150
Further reading	25	Some questions to think about	151
		Further reading	151
		5. Bilingualism and second language	
		acquisition	153
SECTION B: THE BIOLOGICAL		Introduction	153
AND DEVELOPMENTAL BASES		Bilingualism	153
OF LANGUAGE	49		
3. The foundations of language	51		
Introduction	51		

Second language acquisition	158	Some questions to think about	283
Summary	162	Further reading	283
Some questions to think about	162		
Further reading	163		
SECTION C: WORD RECOGNITION	165	SECTION D: MEANING AND USING LANGUAGE	285
6. Recognizing visual words	167	10. Understanding the structure of sentences	287
Introduction	167	Introduction	287
Basic methods and basic findings	168	Dealing with structural ambiguity	289
Meaning-based facilitation of visual word recognition	185	Early work on parsing	291
Morphology: Processing complex words	191	Processing structural ambiguity	298
Models of visual word recognition	192	Gaps, traces, and unbounded dependencies	313
Coping with lexical ambiguity	199	The neuroscience of parsing	315
Summary	207	Summary	319
Some questions to think about	208	Some questions to think about	320
Further reading	208	Further reading	320
7. Reading	209	11. Word meaning	321
Introduction	209	Introduction	321
A preliminary model of reading	211	Classic approaches to semantics	324
The processes of normal reading	212	Semantic networks	325
The neuropsychology of adult reading disorders: Acquired dyslexia	220	Semantic features	328
Models of word naming	227	Family resemblance and classification	335
Summary	239	Combining concepts	338
Some questions to think about	240	Processing figurative language	340
Further reading	240	The neuropsychology of semantics	342
8. Learning to read and spell	241	Connectionist approaches to semantics	353
Introduction	241	Summary	359
Normal reading development	241	Some questions to think about	360
Developmental dyslexia	249	Further reading	360
Summary	255	12. Comprehension	361
Some questions to think about	255	Introduction	361
Further reading	256	Memory for text and inferences	363
9. Understanding speech	257	Reference, co-reference, and ambiguity	373
Introduction	257	Models of text representation and processing	378
Recognizing speech	257	Individual differences in comprehension skills	388
Models of speech recognition	267	The neuropsychology of text and discourse processing	390
The neuropsychology of spoken word recognition	281	Summary	392
Summary	282		

Some questions to think about	393	15. The structure of the language system	463
Further reading	393	Introduction	463
SECTION E: PRODUCTION AND OTHER ASPECTS OF LANGUAGE	395	What are the modules of language?	464
13. Language production	397	How many lexicons are there?	465
Introduction	397	Language and short-term memory	471
Speech errors and what they tell us	399	Summary	476
Syntactic planning	404	Some questions to think about	477
Lexicalization	412	Further reading	477
Phonological encoding	428	16. New directions	479
The analysis of hesitations	432	Introduction	479
The neuropsychology of speech production	435	Themes in psycholinguistics revisited	479
Writing and aphasia	447	Some growth areas?	482
Summary	448	Conclusion	483
Some questions to think about	450	Appendix: Connectionism	485
Further reading	450	Interactive activation models	485
14. How do we use language?	453	Back-propagation	487
Introduction	453	Further reading	489
Making inferences in conversation	454	Glossary	491
The structure of conversation	458	Example of sentence analysis	500
Collaboration in dialog	459	References	501
Sound and vision	460	Author index	575
Summary	462	Subject index	591
Some questions to think about	462		
Further reading	462		

Preface to the Third Edition

I would like to thank Psychology Press for providing the welcome opportunity for me to write the third edition of my text. If back in 1994 when I finished writing the first edition someone had told me that more than 10 years later I'd be writing the third, I would have been extremely happy.

As I remarked in the preface to the first edition, although language might not be all that makes us human, it is hard to imagine being human without it. Given the importance of language in our behavior, it is perhaps surprising that until not so long ago, relatively scant attention was been paid to it in undergraduate courses. Often at best it was studied as part of a general course on cognitive psychology. That situation has changed. Furthermore, the research field of psycholinguistics is blossoming, as evinced by the growth in the number of papers on the subject, and indeed, in the number of journals dedicated to it. With this growth and this level of interest, it is perhaps surprising that there are still relatively few textbooks devoted to psycholinguistics. I hope this book fills this gap. It is aimed at intermediate and advanced level undergraduates, although new postgraduates might also find it useful, and I would be delighted if it found other readers.

I have tried to make as many of the citations as possible refer towards easily obtainable material. I have therefore avoided citing unpublished papers, doctoral theses, and conference papers. New papers are coming out all the time, and if

I were going to make this book completely up to date, I would never stop. Therefore I called a halt at material published in early 2006. Of course, given current publication lags, much of this material would actually have been written some years before, and the current state of people's thinking and work, as discussed in conferences and seminars, might be very different from the positions attributed in this book. This outcome is most unfortunate, but unavoidable.

It is now impossible to appreciate psycholinguistics without some understanding of connectionism. Unfortunately, this is a topic that many people find difficult. The formal details of connectionist models are in an Appendix: I hope this does not mean that it will not be read. I toyed with a structure where the technical details were given when the class of model was first introduced, but a more general treatment seemed more appropriate.

I have been very gratified with the positive feedback I have received on the first two editions of this book, and the number of suggestions and comments I have received. I have tried to take these into account in this revision. "Taking into account" does not mean "agreeing to everything"; after consideration, there are some suggestions that I decided against implementing. One of these was numbering sections, which I find esthetically displeasing. I take the general point that there are many cross-references, but I take this as a

positive point: we should try to foster as many connections between parts of the subject as possible. I have also found that it is impossible to please everybody. Indeed, people's suggestions are often contradictory. In such cases I can only rely on my own inclinations. One example of this is the material on the Sapir-Whorf hypothesis: some people have suggested deleting it, whereas others have found it to be one of the most useful sections, and wanted more information on it. Almost everyone wanted more material on their particular specialty, and probably everyone is going to be disappointed. The book is quite long enough as it is.

It is difficult to believe that it is 6 years since I finished writing the second edition. Much has happened in psychology since then—particularly the widespread use of brain imaging. I have tried to incorporate this research where appropriate.

Whereas the main theme of the first edition was interaction versus independence, and that of the second, broadly speaking, was connectionism, the main theme that has emerged in this third edition is that of the role of language in a wider context. Are language processes specific to language, and do we have innate linguistic knowledge and dedicated regions of the brain doing language processing? Or are they a reflection of more general cognitive processes? In a way this theme is a synthesis of modularity and connectionism.

The structure of the book has settled down—whereas the second edition involved major surgery on the first, all I have done in the third edition is to split the lengthy chapter on reading into adult and child reading, and added a chapter about how we use language.

Students often find the study of the psychology of language rather dry and technical, and many find it difficult. In addition to making this edition as comprehensible as possible, I have also tried to make it fun. I have also tried to emphasize applications of research.

The American Psychological Association now recommends the use of the word “participants” instead of “subjects.” I have followed this recommendation

There is a web site associated with this book. It contains links to other pages, details of

important recent work, and a “hot link” to contact me. It is to be found at: <http://www.dundee.ac.uk/psychology/language>. I still welcome any corrections, suggestions for the next edition, or discussion on any topic. My email address is now: t.a.harley@dundee.ac.uk. Suggestions on topics I have omitted or under-represented would be particularly welcome. The hardest bit of writing this book has been deciding what to leave out. I am sure that people running other courses will cover some material in much more detail than has been possible to provide here. However, I would be interested to hear of any major differences of emphasis. If the new edition is as successful as the second, I will be looking forward (in a strange sort of way) to producing the fourth edition in 5 years time.

I would like to thank all those who have made suggestions about both editions, particularly Jeanette Altarriba, Gerry Altmann, Elizabeth Bates, Helen Bown, Paul Bloom, Peer Broeder, Gordon Brown, Hugh Buckingham, Lynne Duncan, the Dundee Psycholinguistics Discussion Group, Andy Ellis, Gerry Griffin, Zenzi Griffin, Annette de Groot, Francois Grosjean, Evan Heit, Laorag Hunter, Lesley Jessiman, Barbara Kaup, Alan Kennedy, Kathryn Kohnert, Annukka Lindell, Nick Lund, Siobhan MacAndrew, Nadine Martin, Randi Martin, Elizabeth Maylor, Don Mitchell, Wayne Murray, Lyndsey Nickels, Jane Oakhill, Pádraig O'Seaghdha, Shirley-Anne Paul, Martin Pickering, Julian Pine, Ursula Pool, Eleanor Saffran, Lynn Santelmann, Marcus Taft, Jeremy Tree, Roger van Gompel, Carel van Wijk, Beth Wilson, Alan Wilkes, Suzanne Zeedyk, and Pienie Zwitserlood. I would also like to thank several anonymous reviewers for their comments; hopefully you know who you are. Numerous people pointed out minor errors and asked questions: I thank them all. George Dunbar created the sound spectrogram for Figure 2.1 using MacSpeechLab. Lila Gleitman gave me the very first line; thanks! Katie Edwards, Pam Miller, and Denise Jackson helped me to obtain a great deal of material, often at very short notice. This book would be much worse without the help of all these people. I am of course responsible for any errors or omissions that remain. If there is any one else I have

forgotten, please accept my apologies. Many people have suggested things that I have thought about and decided not to implement, and many people have suggested things (more connectionism, less connectionism, leave that in, take that out, move that bit there, leave it there) that are the opposite of what others have suggested.

I would also like to thank Psychology Press for all their help and enthusiasm for this project, particularly Lucy Kennedy, Tara Stebnicky, and Mandy Collison. Holly Loftus made up the ancillary materials: Shirley-Anne Paul helped me to check them. Finally, I would like to thank Brian Butterworth, who supervised my PhD. He probably doesn't realize how much I appreciated his help; without him, this book might never have existed.

Finally, perhaps I should state my bias about some of the more controversial points of psycholinguistics: I think language processing is

massively interactive, I think connectionist modeling has contributed enormously to our understanding and is the most profitable direction to go in the near future, and I think that the study of the neuropsychology of language is fundamental to our understanding. Writing this edition has fostered and reinforced these beliefs. I realize that many will disagree with me, and have tried to be as fair as possible. I hope that any bias there is in this book will appear to be the consequence of the consideration of evidence rather than of prejudice.

*Trevor A. Harley
School of Psychology
University of Dundee
Dundee DD1 4HN
Scotland*

March 2007

How to Use This Book

This book is intended to be a stand-alone introduction to the psychology of language. It is my hope that anyone could pick it up and finish reading it with a rich understanding of how humans use language. Nevertheless, it would probably be advantageous to have some knowledge of basic cognitive psychology. (Some suggestions for books to read are given in the “Further reading” section at the end of Chapter 1.) For example, you should be aware that psychologists have distinguished between short-term memory (which has limited capacity and can store material for only short durations) and long-term memory (which is virtually unlimited). I have tried to assume that the reader has no knowledge of linguistics, although I hope that most readers will be familiar with such concepts as nouns and verbs. The psychology of language is quite a technical area full of rather daunting terminology. I have italicized technical terms when they first appear, and short definitions are given in boxes on the pages concerned. There is also a glossary near the end of the book with short definitions of the technical terms.

Connectionist modeling is now central to modern cognitive psychology. Unfortunately, it is also a topic that most people find extremely difficult to follow. It is impossible to understand the *details* of connectionism without some mathematical sophistication. I have provided an Appendix that covers the basics of connectionism in more mathematical detail than is generally necessary to

understand the main text. However, the general principles of connectionism can probably be appreciated without this extra depth, although it is probably a good idea to look at the Appendix.

In my opinion and experience, the material in some chapters is more difficult than others. I do not think there is anything much that can be done about this, except to persevere. Sometimes comprehension might be assisted by later material, and sometimes a number of readings might be necessary to comprehend the material fully. Fortunately the study of the psychology of language gives us clues about how to facilitate understanding. Chapters 7 and 11 will be particularly useful in this respect. It should also be remembered that in some areas researchers do not agree on the conclusions, or on what should be the appropriate method to investigate a problem. Therefore it is sometimes difficult to say what the “right answer,” or the correct explanation of a phenomenon, might be. In this respect the psychology of language is still a very young subject.

The book is divided into sections, each covering an important aspect of language. Section A is an introduction. It describes what language is, and provides essential background for describing language. It should not be skipped. Section B is about the biological basis of language, the relationship of language to other cognitive processes, and language development. Section C is about how we recognize words. Section D is about comprehension:

how we understand sentences and discourse. Section E is about language production, and also about how language interacts with memory. It also examines the grand design or architecture of the language system. The section concludes with a brief look at some possible new directions in the psychology of language.

Each chapter begins with an introduction outlining what the chapter is about and the main problems faced in each area. Each introduction ends with a summary of what you should know by the end of the chapter. Each chapter concludes with a list of bullet points that gives a one-sentence summary of each section in that chapter. This is followed by questions that you can think about, either to test your understanding of the material, or to go beyond what is covered, usually with an emphasis on applying the material. If you want to follow up a topic in more detail than is covered in the text (which I think is quite richly referenced, and should be the first place to look), then there are suggestions for further reading at the very end of each chapter.

One way of reading this book is like a novel: start here and go to the end. Section A should certainly be read before the others because it introduces many important terms, without which later going would be very difficult. However, after that, other orders are possible. I have tried to make each chapter as self-contained as possible, so there is no reason why the chapters cannot be read in a different order. Similarly, you might choose to omit some chapters altogether. In each case you might find you have to refer to the glossary more often than if you just begin at the beginning. Unless you are interested in just a few topics, however, I advise reading the whole book through at least once.

Each chapter looks at a major chunk of the study of the psychology of language.

OVERVIEW OF THIS BOOK

Chapter 1 tells you about the *subject of the psychology of language*. It covers its history and methods. Chapter 2 provides some *important*

background on language, telling you how we can describe sounds and the structure of sentences. In essence it is a primer on phonology and syntax.

Chapter 3 is about *how language is related to biological and cognitive processes*. It looks at the extent to which language depends on the presence and operation of certain biological, cognitive, and social precursors in order to be able to develop normally. We will also examine whether animals use language, or whether they can be taught to do so. This will also help to clarify what we mean by language. We will look at how language is founded in the brain, and how damage to the brain can lead to distinct types of impairment in language. We will examine in detail the more general role of language, by examining the relation between language and thought. We will also look at what can be learned from language acquisition in exceptional circumstances, including the effects of linguistic deprivation.

Chapter 4 examines *how children acquire language*, and how language develops throughout childhood. Chapter 5 examines *how bilingual children learn to use two languages*.

We will then look at what appear to be the simplest or lowest level processes and work towards more complex ones. Hence we will first examine how we recognize and understand single words. Although these chapters are largely about recognizing words in isolation, in the sense that in most of the experiments we discuss only one word is present at a time, the influence of the context in which they are found is an important consideration, and we will look at this as well.

Chapter 6 addresses *how we recognize words* and how we access their meanings. Although the emphasis is on visually presented word recognition, many of the findings described in this chapter are applicable to recognizing spoken words as well. Chapter 7 examines *how we read and pronounce words*, and looks at disorders of reading (the dyslexias). Chapter 8 looks at how we learn to read and spell, how reading should best be taught, and at the difficulties some children face in learning to read—developmental dyslexia. Chapter 9 looks at the speech system and *how we process speech and identify spoken words*.

We then move on to how words are ordered to form sentences. Chapter 10 looks at *how we make use of word order information in understanding sentences*. These are issues to do with syntax and parsing. Chapter 11 examines *how we represent the meaning of words*. Chapter 12 examines *how we comprehend and represent beyond the sentence level*; these are the larger units of discourse or text. In particular, how do we integrate new information with old to create a coherent representation? How do we store what we have heard and read?

In Chapter 13 we consider the process in reverse, and *examine language production and its disorders*. By this stage we will have an

understanding of the processes involved in understanding language, and these processes must be looked at in a wider context. Chapter 14 looks at *how we use language*. It examines how we control conversation, and how speakers and listeners cooperate to make dialogue efficient. It also examines how listeners draw inferences about what the speaker means from apparent violations of the normal expectation of how to run a good conversation.

In Chapter 15 we look at *the structure of the language system as a whole*, and the relation between the parts. Finally, Chapter 16 looks at *some possible new directions* in psycholinguistics.

7

Reading

INTRODUCTION

In Chapter 6 we looked at how we recognize words; this chapter is about how we read them. How do we gain access to the sounds and meanings of words? We also examine the effects of brain damage on reading (giving rise to acquired dyslexia), and show how reading disorders can be related to a model of reading. The next chapter looks at how children learn to read.

Reading aloud and reading to oneself are clearly different, but related, tasks. When we read aloud (or name words), we must retrieve the sounds of words. When we read to ourselves, we read to obtain the meaning, but most of us, most of the time, experience the sounds of the words as “inner speech.” Is it possible to go to the meaning of a word when reading without also accessing its sounds? By the end of this chapter you should:

- Know how different languages translate words into sounds, and understand the alphabetic principle.
- Understand the motivation for the dual-route model of reading, and know about its strengths and weaknesses.

- Appreciate how different types of dyslexia relate to the dual-route model, and also the problems they pose for it.
- Know about connectionist models of reading and how they account for dyslexia.

The writing system

The basic unit of written language is the letter. The name **grapheme** is given to the letter or combination of letters that represents a phoneme. For example, the word “ghost” contains five letters and four graphemes (“gh,” “o,” “s,” and “t”), representing four phonemes. There is much more variability in the structure of written languages than there is in spoken languages. Whereas all spoken languages utilize a basic distinction between consonants and vowels, there is no such common thread to the world’s written languages.

KEY TERM

Grapheme: a unit of written language that corresponds to a *phoneme* (e.g., “steak” contains four graphemes: s t ea k).

The sorts of written language most familiar to speakers of English and other European languages are *alphabetic scripts*. English uses an alphabetic script. In alphabetic scripts, the basic unit represented by a grapheme is essentially a phoneme. However, the nature of this correspondence can vary. In transparent languages such as Serbo-Croat and Italian there is a one-to-one grapheme–phoneme correspondence, so that every grapheme is realized by only one phoneme and every phoneme is realized by only one grapheme. In languages such as English this relation can be one-to-many in both directions. A phoneme can be realized by different graphemes (e.g., compare “to,” “too,” “two,” and “threw”), and a grapheme can be realized by many different phonemes (e.g., the letter “a” in the words “fate,” “pat,” and “father”). Some languages lie between these extremes. In French, correspondences between graphemes and phonemes are quite regular, but a phoneme may have different graphemic realizations (e.g., the graphemes “o,” “au,” “eau,” “aux,” and “eaux” all represent the same sounds). In *consonantal scripts*, such as Hebrew and Arabic, not all sounds are represented, as vowels are not written down at all. In *syllabic scripts* (such as Cherokee and the Japanese script kana), the written units represent syllables. Finally, some languages do not represent any sounds. In ideographic languages (sometimes also called *logographic languages*), such as Chinese and the Japanese script kanji, each symbol is equivalent to a morpheme.

One consequence of this variation in writing systems is that there must be differences in processing between readers of different languages.



There is much more variability in the structure of written languages than there is in spoken languages. In consonantal scripts, such as Hebrew (above) and Arabic, not all sounds are represented.

Hence this chapter should be read with the caution in mind that some conclusions may be true of English and many other writing systems, but not necessarily all of them.

Unlike speech, reading and writing are a relatively recent development. Writing emerged independently in Sumer and Mesoamerica, and perhaps also in Egypt and China. The first writing system was the cuneiform script printed on clay in Sumer, which appeared just before 3000 bc. The emergence of the alphabetic script can be traced to ancient Greece in about 1000 bc. The development of the one-to-many correspondence in English orthography primarily arose between the fifteenth and eighteenth centuries as a consequence of the development of the printing press and the activities of spelling “reformers” who tried to make the Latin and Greek origins of words more apparent in their spellings (see Ellis, 1993, for more detail). Therefore it is perhaps not surprising that reading is actually quite a complex

Types of written languages		
	<u>Examples</u>	<u>Features</u>
Alphabetic script	English Other European languages	The basic unit represented by a grapheme is essentially a phoneme.
Consonantal script	Hebrew Arabic	Not all sounds are represented, as vowels are not written down.
Syllabic script	Cherokee Japanese kana	Written units represent syllables.
Logographic/ ideographic script	Chinese Japanese kanji	Each symbol represents a whole word.

cognitive task. There is a wide variation in reading abilities, and many different types of reading disorder arise as a consequence of brain damage.

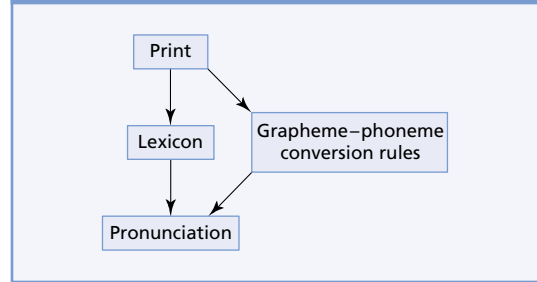
A PRELIMINARY MODEL OF READING

Introspection can provide us with a preliminary model of reading. Consider how we might name or pronounce the word “beef.” Words like this are said to have a regular *spelling-to-sound correspondence*. That is, the graphemes map onto phonemes in a totally regular way; you need no special knowledge about the word to know how to pronounce it. If you had never seen the word “beef” before, you could still pronounce it correctly. Some other examples of regular word pronunciations include “hint” and “rave.” In these words, there are alternative pronunciations (as in “pint” and “have”), but “hint” and “rave” are pronounced in accordance with the most common pronunciations. These are all *regular words*, because all the graphemes have the standard pronunciation.

Not all words are regular, however. Some are irregular or exception words. Consider the word “steak.” This has an *irregular* spelling-to-sound (or grapheme-to-phoneme) correspondence: the grapheme “ea” is not pronounced in the usual way, as in “streak,” “sneak,” “speak,” “leak,” and “beak.” Other exceptions to a rule include “have” (an exception to the rule that leads to the regular pronunciations “gave,” “rave,” “save” and so forth) and “vase” (in British English, an exception to the rule that leads to the regular pronunciations “base,” “case,” and so forth). English has many irregular words. Some words are extremely irregular, containing unusual patterns of letters that have no close neighbors, such as “island,” “aisle,” “ghost,” and “yacht.” These words are sometimes called lexical hermits.

Finally, we can pronounce strings of letters such as “nate,” “smeak,” “fot,” and “datch,” even though we have never seen them before. These letter strings are all *pronounceable nonwords* or *pseudowords*. Therefore, even though they are novel, we can still pronounce them, and we all tend to agree on how they should be pronounced. If you hear nonwords like these, you can spell

FIGURE 7.1



The simplified version of the dual-route model of reading.

them correctly; you assemble their pronunciations from their constituent graphemes. (Of course, not all nonwords are pronounceable—e.g., “xzhgh.”)

Our ability to read nonwords on the one hand and irregular words on the other suggests the possibility of a *dual-route model* of naming. We can assemble pronunciations for words or nonwords we have never seen before, yet also pronounce correctly irregular words that must need information specific to those words (that is, *lexical information*). The classic dual-route model (see Figure 7.1) has two routes for turning words into sounds. There is a *direct* access or *lexical route*, which is needed for irregular words. This must at least in some way involve a direct link between print and sound. That is, the lexical route takes us directly to a word’s entry in the lexicon and we are then able to retrieve the sound of a word. There is also a *grapheme-to-phoneme conversion* (GPC) route (also called the *indirect* or *non-lexical* or **sublexical** route), which is used for reading nonwords. This route carries out what is called *phonological recoding*. It does not involve lexical access at all. The non-lexical route was first proposed in the early 1970s (e.g., Gough, 1972; Rubenstein, Lewis, & Rubenstein, 1971). Another important justification for a grapheme-to-phoneme conversion route is that it is useful for children learning to read by sounding out words letter by letter.

KEY TERM

Sublexical: correspondences in spelling and sound beneath the level of the whole word.

Given that neither route can in itself adequately explain reading performance, it seems that we must use both. Modern dual-route theorists see reading as a “race” between these routes. When we see a word, both routes start processing it. For skilled readers, most of the time the direct route is much faster, so it will usually win the race and the word will be pronounced the way that it recommends. The indirect route will only be apparent in exceptional circumstances, such as when we see a very unfamiliar word; in that case, if the direct route is slower than normal, then the direct and GPC routes will produce different pronunciations at about the same time, and these words might be harder to pronounce.

Relation of the dual-route model to other models

In the previous chapter we examined a number of models of word recognition. These can all be seen as theories of how the direct, lexical access reading route operates. The dual route is the simplest version of a range of possible multi-route or *parallel coding* models, some of which posit more than two reading routes. Do we really need a non-lexical route at all for routine reading? Although we appear to need it for reading nonwords, it seems a costly procedure. We have a mechanism ready to use for something we rarely do—pronouncing new words or nonwords. Perhaps it is left over from the development of reading, or perhaps it is not as costly as it first appears. We will see later that the non-lexical route is also apparently needed to account for the neuropsychological data. Indeed, whether or not two routes are necessary for reading is a central issue of the topic of reading. Models that propose that we can get away with only one (such as connectionist models) must produce a satisfactory account of how we can pronounce nonwords.

Of course, except for reading aloud, the primary goal of reading is not getting the sound of a word, but getting the meaning. As we shall see in Chapter 8, in the early stages of learning to read children get to the meaning through the sound; that is, they spell out the sound of the words, and then access meaning as they recognize those sounds. Some researchers believe that even skilled

adults primarily get to meaning by going from print to phonology and then to meaning, an idea called *phonological mediation* (discussed in more detail below). Most researchers however believe that in skilled adults, most of the time, there is a direct route from print to semantics. Indeed, as we shall see below, most researchers believe that there is a direct route from print to sound, and a direct route via semantics; what is debated is the role of the indirect route in normal reading (see Taft & van Graan, 1998, for further discussion of these issues).

THE PROCESSES OF NORMAL READING

According to the dual-route model, there are two independent routes when naming a word and accessing the lexicon: a lexical or direct access route and a sublexical or grapheme–phoneme conversion route. This section looks at how we name nonwords and words.

Reading nonwords

According to the dual-route model, the pronunciation of all nonwords should be assembled using the GPC route. This means that all pronounceable nonwords should be alike and their similarity to words should not matter. However, pronounceable nonwords are not all alike.

The pseudohomophone effect

Pseudohomophones are pronounceable nonwords that sound like words when pronounced (such as “brane,” which sounds like the word “brain” when spoken). The behavior of the pseudohomophone “brane” can be compared with the very similar nonword “brame,” which does not sound like a word when it is spoken. Rubenstein et al.

KEY TERM

Pseudohomophones: a *nonword* that sounds like a word when pronounced (e.g., “nite”).

(1971) showed that pseudohomophones are more confusable with words than other types of nonwords are. Participants are faster to name them, but slower to reject them as nonwords than control nonwords.

Is the effect caused by the phonological or visual similarity between the nonword and word? Martin (1982) and Taft (1982) argued that it is visual similarity that is important. Pseudohomophones are more confusable with words than other nonwords are because they *look* more similar to words than non-pseudohomophones do, rather than because they *sound* the same. Pring (1981) alternated the case of letters within versus across graphemes, such as the “AI” in “grait,” to produce “GraIT” or “GRaiT.” These strings look different but still sound the same. Alternating letter cases within a grapheme or spelling unit (ai) eliminates the pseudohomophone effect; alternating letters elsewhere in the word (aiT) does not. Hence we are sensitive to the visual appearance of spelling units of words.

The pseudohomophone effect suggests that not all nonwords are processed in the same way. The importance of the visual appearance of the nonwords further suggests that something else apart from phonological recoding is involved here. It remains to be seen whether the phonological recoding route is still necessary, but if it is, then it must be more complex than we first thought.

Glushko’s (1979) experiment: Lexical effects on nonword reading

Glushko (1979) performed a very important experiment on the effect of the regularity of the word-neighbors of a nonword on its pronunciation. Consider the nonword “taze.” Its word-neighbors include “gaze,” “laze,” and “maze”; these are all themselves regularly pronounced words. Now consider the word-neighbors of the nonword “tave.” These also include plenty of regular words (e.g., “rave,” “save,” and “gave”) but there is an exception word-neighbor (“have”). As another example, compare the nonwords “feal” and “fead”: both have regular neighbours (e.g., “real,” “seal,” “deal,” and “bead”) but the pronunciation of “fead” is influenced by its irregular neighbour “dead.” Glushko (1979) showed that naming latencies to

nonwords such as “tave” were significantly slower than to ones such as “taze.” That is, reaction times to nonwords that have orthographically irregular spelling-to-sound correspondence word-neighbors are slower than to other nonword controls. Also, people make pronunciation “errors” with such nonwords: “pove” might be pronounced to rhyme with “love” rather than “cove”; and “heaf” might be pronounced to rhyme with “deaf” rather than “leaf.” In summary, Glushko found that the pronunciation of nonwords is affected by the pronunciation of similar words, and that nonwords are not the same as each other. Subsequent research has shown that the proportion of regular pronunciations of nonwords increases as the number of orthographic neighbours increases (McCann & Besner, 1987). In summary, there are lexical effects on nonword processing.

More on reading nonwords

The nonword “yead” can be pronounced to rhyme with “bead” or “head.” Kay and Marcel (1981) showed that its pronunciation can be affected by the pronunciation of a preceding prime word: “bead” biases a participant to pronounce “yead” to rhyme with it, whereas the prime “head” biases participants to the alternative pronunciation. Rosson (1983) primed the nonword by a semantic relative of a phonologically related word. The task was to pronounce “louch” when preceded either by “feel” (which is associated with “touch”) or by “sofa” (which is associated with “couch”). In both cases “louch” tended to be pronounced to rhyme with the appropriate relative.

Finally, nonword effects in complex experiments are sensitive to many factors, such as the pronunciation of the surrounding words in the list. This also suggests that nonword pronunciation involves more than just grapheme-to-phoneme conversion.

Evaluation of research on reading nonwords

These data do not fit the simple version of the dual-route model. The pronunciation of nonwords is affected by the pronunciation of visually similar words. That is, there are lexical effects in nonword

processing; the lexical route seems to be affecting the non-lexical route.

Word processing

According to the dual-route model, words are accessed directly by the direct route. This means that all words should be treated the same in respect of the regularity of their spelling-to-sound correspondences. An examination of the data reveals that this prediction does not stand up.

One problem for the simple dual-route model is that pronunciation regularity affects response times, although in a complex way. Baron and Strawson (1976) provided an early demonstration of this problem, finding that a list of regular words was named faster than a list of frequency-matched exception words (e.g., “have”). This task is a simplified version of the naming task, with response time averaged across many items rather than taken from each one individually. There have been many other demonstrations of the influence of regularity on naming time (e.g., Forster & Chambers, 1973; Frederiksen & Kroll, 1976; Stanovich & Bauer, 1978). A well-replicated finding is that of an interaction between regularity and frequency: regularity has little effect on the pronunciation of high-frequency words, but low-frequency regular words are named faster than low-frequency irregular words (e.g., Andrews, 1982; Seidenberg, Waters, Barnes, & Tanenhaus, 1984), even when we control for age-of-acquisition (Monaghan & Ellis, 2002). Jared (1997b) found that high-frequency words can be sensitive to regularity, but the effect of regularity is moderated by the number and frequencies of their “friends” and “enemies” (words with similar or conflicting pronunciations). That is, it is important to control for the neighborhood characteristics of the target words as well as their regularity in order to observe the interaction. On the other hand, it is not clear whether there are regularity effects on lexical decision. They have been obtained by, for example, Stanovich and Bauer (1978), but not by Coltheart et al. (1977), or Seidenberg et al. (1984). In particular, a word such as “yacht” looks unusual, as well as having an irregular pronunciation. The letter pairs “ya” and “ht” are not frequent in

English; we say they have a *low bigram frequency*. Obviously the visual appearance of words is going to affect the time it takes for direct access, so we need to control for this when searching for regularity effects. Once we control for the generally unusual appearance of irregular words, regularity and consistency only seem to affect naming times, not lexical decision times. Age-of-acquisition has a similar effect to frequency, and gives rise to a similar interaction: Consistency has a much bigger impact on naming time for late-acquired than early-acquired words (Monaghan & Ellis, 2002). Why do late-acquired and low-frequency inconsistent words stand out? One possibility is that late-acquired low-frequency consistent words can make use of the network structure of other consistent words; inconsistent items cannot, and need new associations to be learned between input and output (Monaghan & Ellis, 2002).

In general, regularity effects are more likely to be found when participants have to be more conservative, such as when accuracy rather than speed is emphasized. The finding that regularity affects naming might appear problematic for the dual-route model, but makes sense if there is a race between the direct and indirect routes. Remember that there is an interaction between regularity and frequency. The pronunciation of common words is directly retrieved before the indirect route can construct any conflicting pronunciation. Conflict arises when the lexical route is slow, as when retrieving low-frequency words, and when the pronunciation of a low-frequency word generated by the lexical route conflicts with that generated by the non-lexical route (Norris & Brown, 1985).

Glushko's (1979) experiment: Results from words

Glushko (1979) also found that words behave in a similar way to nonwords, in that the naming times of words are affected by the phonological consistency of neighbors. The naming of a regular word is slowed down relative to that of a control word of similar frequency if the test word has irregular neighbors. For example, the word “gang” is regular, and all its neighbors (such as “bang,” “sang,” “hang,” and “rang”) are also regular.

TABLE 7.1

Classification of word pronunciations depending on regularity and consistency (based on Patterson & Morton, 1985)

Word type	Example	Characteristics
Consistent	gaze	All words receive the same regular pronunciation of the body
Consensus	lint	All words with one exception receive the same regular pronunciation
Heretic	pint	The irregular exception to the consensus
Gang	look	All words with one exception receive the same irregular pronunciation
Hero	spook	The regular exception to the gang
Gang without a hero	cold	All words receive the same irregular pronunciation
Ambiguous: conformist	cove	Regular pronunciation with many irregular exemplars
Ambiguous: independent	love	Irregular pronunciation with many regular exemplars
Hermit	yacht	No other word has this body

Consider on the other hand “base”; this itself has a regular pronunciation (compare it with “case”), but it is inconsistent, in that it has one irregular neighbour, “vase” (in British English pronunciation). We could say that “vase” is an *enemy* of “base.” This leads to a slowing of naming times. In addition, Glushko found true naming errors of over-regularization: for example “pint” was sometimes given its regular pronunciation—to rhyme with “dint.”

Pronunciation neighborhoods

Continuing this line of research, Brown (1987) argued that the number of consistently pronounced neighbors (*friends*) determines naming times, rather than whether a word has *enemies* (that is, whether or not it is regular). It is now thought that the number of both friends and enemies affects naming times (Brown & Watson, 1994; Jared, McRae, & Seidenberg, 1990; Kay & Bishop, 1987).

Andrews (1989) found effects of neighborhood size in both the naming and the lexical decision tasks. Responses to words with large neighborhoods were faster than words with small neighborhoods (although this may be moderated by frequency, as suggested by Grainger, 1990). Not all readers produce the same results. Barron (1981) found that good and poor elementary school readers both read regular words more quickly than irregular words. However, once he controlled for neighborhood effects, he found that there was no longer any regularity effect in the good readers, although it persisted in the poor readers.

Parkin (1982) found more of a continuum of ease-of-pronunciation than a simple division between regular and irregular words. All this work suggests that a binary division into words with regular and irregular pronunciations is no longer adequate. Patterson and Morton (1985) provided a more satisfactory but complex categorization rather than a straightforward dichotomy between regular and irregular words (see Table 7.1). This classification reflects two factors: first, the regularity of the pronunciation with reference to spelling-to-sound correspondence rules; second, the agreement with other words that share the same **body**. (This is the end of a monosyllabic word, comprising the central vowel plus final consonant or consonant cluster; e.g., “aint” in “saint” or “us” in “plus.”) We need to consider not only whether a word is regular or irregular, but also whether its neighbors are regular or irregular. The same classification scheme can be applied to nonwords.

In summary, just as not all nonwords behave in the same way, neither do all words. The regularity of pronunciation of a word affects the ease with which we can name it. In addition, the pronunciation of a word’s neighbors can affect its naming. The number of friends and enemies affects how easy it is to name a word.

KEY TERM

Body: the same as a *rime*: the final vowel and terminal consonants.

*The role of sound in accessing meaning:
Phonological mediation*

There is some experimental evidence suggesting that a word's sound may have some influence on accessing the meaning (Frost, 1998; van Orden, 1987; van Orden, Johnstone, & Hale, 1988; van Orden, Pennington, & Stone, 1990). In a category decision task, participants have to decide if a visually presented target word is a member of a particular category. For example, given "A type of fruit" you would respond "yes" to "pear," and "no" to "pour." If the "no" word is a homophone of a "yes" word (e.g., "pair"), participants make a lot of false positive errors—that is, they respond "yes" instead of "no." Participants seem confused by the sound of the word, and category decision clearly involves accessing the meaning. The effect is most noticeable when participants have to respond quickly. Lesch and Pollatsek (1998) found evidence of interference between homophones in a semantic relatedness task (e.g., SAND–BEECH). We take longer to respond to homophones in a lexical decision task (e.g., MAID), presumably because the homophones are generating confusion in lexical access, perhaps through feedback from phonology to orthography (Pexman, Lupker, & Jared, 2001; Pexman, Lupker, & Reggin, 2002).

Hence there is considerable evidence that the recognition of a word can be influenced by its phonology. The dominant view is that this influence arises through the indirect route, although word recognition is primarily driven by the direct route (or routes)—a view that has been labeled the weak phonological perspective (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Rastle & Brysbaert, 2006). Most of the models described in this chapter subscribe to the weak phonological view. The alternative, strong phonological view, that we primarily get to the meaning through sound, is called *phonological mediation*. The most extreme form of this idea is that visual word recognition cannot occur in the absence of computing the sound of the word.

There is a great deal of controversy about the status of phonological mediation. Other experiments support the idea. Folk (1999) examined

eye movements as participants read sentences containing either "soul" or "sole." Folk found that the homophones were read with longer gaze duration—that is, they were processed as though they were lexically ambiguous, even though the orthography should have prevented this. This result is only explicable if the phonology is in some way interfering with the semantic access.

On the other hand, Jared and Seidenberg (1991) showed that prior phonological access only happens with low-frequency homophones. In an examination of proof-reading and eye movements, Jared, Levy, and Rayner (1999) also found that phonology only plays a role in accessing the meanings of low-frequency words. In addition, they found that poor readers are more likely to have to access phonology in order to access semantics, whereas good readers primarily activate semantics first. Daneman, Reingold, and Davidson (1995) reported eye fixation data on homophones that suggested the meaning of a word is accessed first whereas the phonological code is accessed later, probably post-access. They found that gaze duration times were longer on an incorrect homophone (e.g., "brake" was in the text when the context demanded "break"), and that the fixation times on the incorrect homophone were about the same as on a spelling control (e.g., "broke"). This means that the appropriate meaning must have been activated before the decision to move the eyes, and that the phonological code is not activated at this time. (If the phonological code had been accessed before meaning then the incorrect homophone would sound all right in the context, and gaze durations should have been about the same.) The phonological code is accessed later, however, and influences the number of *regressions* (when the eyes look back to earlier material) to the target word. (However, see Rayner, Pollatsek, & Binder, 1998, for different conclusions. It is clear that these experiments are very sensitive to the materials used.)

Taft and van Graan (1998) used a semantic categorization task to examine phonological mediation. Participants had to decide whether or not words belonged to a category of "words with definable meanings" (e.g., "plank," "pint") or the category of "given names" (e.g., "Pam," "Phil"). There was no difference in the decision times

between regular definable words (e.g., “plank”) and irregular definable words (e.g., “pint”), although a regularity effect was shown in a word naming task. This suggests that the sound of a word does not need to be accessed on the route to accessing its meaning.

A number of studies have tried to decide between the strong and weak phonological views using masked phonological priming. In this technique, targets (e.g., “clip”) are preceded by phonologically identical nonword primes (e.g., “klijp”). Responses to the targets are faster and more accurate than when the target is preceded by an unrelated word. Several studies have found priming effects occur even when the primes have been masked and presented so briefly that they cannot be consciously observed and reported, suggesting that the phonological stimulus must occur automatically and extremely quickly (e.g., Lukatela & Turvey, 1994a, 1994b; Perfetti, Bell, & Delaney, 1988). While some researchers interpret masked phonological priming as supporting phonological mediation—Why else should early phonological activation happen so early unless it is essential?—other researchers point out that these effects are very sensitive to environmental conditions, and are not always reliably found (see Rastle & Brysbaert, 2006, for a review). In a meta-analysis of the literature, Rastle and Brysbaert (2006) do find small but significant masked phonological priming effects.

These data suggest that the sound of a word is usually accessed at an early stage. However, there is much evidence suggesting that phonological recoding cannot be obligatory in order to access the word’s meaning (Ellis, 1993). For example, some dyslexics cannot pronounce nonwords, yet can still read many words. Hanley and McDonnell (1997) described the case of a patient, PS, who understood the meaning of words in reading without being able to pronounce them correctly. Critically, PS did not have a preserved *inner* phonological code that could be used to access the meaning. Some patients have preserved inner phonology and preserved reading comprehension, but make errors in speaking aloud (Caplan & Waters, 1995b). Hanley and McDonnell argued that PS did not have access to his phonological

code because he was unable to access both meanings of a homophone from seeing just one in print. Thus PS could not produce the phonological forms of words aloud correctly, and did not have access to an internal phonological representation of those words, yet he could still understand them when reading them. For example, he could give perfect definitions of printed words. In general, a review of the neuropsychological literature suggests that people can recognize words in the absence of phonology (Coltheart, 2004). Hence it is unlikely that phonological recoding is an obligatory component of visual word recognition (Rastle & Brysbaert, 2006).

How then can we explain the data showing phonological mediation? There are a number of alternative explanations. First, although phonological recoding prior to accessing meaning may not be obligatory, it might occur in some circumstances. Given there is a race between the lexical and sublexical routes in the dual-route model, if for some reason the lexical route is slow in producing an output, the sublexical route might have time to assemble a conflicting phonological representation. Second, there might be feedback from the speech production system to the semantic system, or the direct-access route causes inner speech that interferes with processing. Third, it is possible that lexical decision is based on phonological information (Rastle & Brysbaert, 2006).

Silent reading and inner speech

Although it seems unlikely that we have to access sound before meaning, we do routinely seem to access some sort of phonological code after accessing meaning in silent reading. Subjective evidence for this is the experience of “inner speech” while reading. Tongue-twisters such as (1) take longer to read silently than sentences where there is variation in the initial consonants (Haber & Haber, 1982). This suggests that we are accessing some sort of phonological code as we read.

- (1) Boris burned the brown bread badly.

However, this inner speech cannot involve exactly the same processes as overt speech because



The experience of “inner speech” while reading demonstrates that we can access some sort of phonological code after accessing meaning in silent reading.

we can read silently much faster than we can read aloud (Rayner & Pollatsek, 1989), and because overt articulation does not prohibit inner speech while reading. Furthermore, although most people who are profoundly deaf read very poorly, some read quite well (Conrad, 1972). Although this might suggest that eventual phonological coding is optional, it is likely that these deaf able readers are converting printed words into some sign language code (Rayner & Pollatsek, 1989). Evidence for this is that deaf people are troubled by the silent reading of word strings that correspond to hand-twisters (Treiman & Hirsh-Pasek, 1983). (Interestingly, deaf people also have some difficulty with signing phonological tongue-twisters, suggesting that difficulty can arise from lip-reading sounds.)

Hence, when we read we seem to access a phonological code that we experience as inner speech. That is, when we gain access to a word’s representation in the lexicon, all its attributes become available. The activation of a phonological code is not confined to alphabetic languages. On-line experimental data using priming and semantic judgment tasks suggest that phonological information about ideographs is automatically activated in both Chinese (Perfetti & Zhang, 1991, 1995) and Japanese kanji (Wydell, Patterson, & Humphreys, 1993).

Inner speech seems to assist comprehension; if it is reduced, comprehension suffers for all but

the easiest material (Rayner & Pollatsek, 1989). McCutchen and Perfetti (1982) argued that whichever route is used for lexical access in reading, at least part of the phonological code of each word is automatically accessed—in particular we access the sounds of beginnings of words. Although there is some debate about the precise nature of the phonological code and how much of it is activated, it does seem that silent reading necessarily generates some sort of phonological code (Rayner & Pollatsek, 1989). This information is used to assist comprehension, primarily by maintaining items in sequence in working memory.

The role of meaning in accessing sound

Phonological mediation means that we might access meaning via sound. Sometimes need to access the meaning before we can access a word’s sound. Words such as “bow,” “row,” and “tear,” have two different pronunciations. This type of word is called a *homograph*. How do we select the appropriate pronunciation? Consider sentences (2) and (3):

- (2) When his shoelace came loose, Vlad had to tie a bow.
- (3) At the end of the play, Dirk went to the front of the stage to take a bow.

Clearly here we need to access the word’s meaning before we can select the appropriate pronunciation. Further evidence that semantics can affect reading is provided by a study by Strain, Patterson, and Seidenberg (1995). They showed that there is an effect of imageability on skilled reading such that there is a three-way interaction between frequency, imageability, and spelling consistency. People are particularly slow and make more errors when reading low-frequency exception words with abstract meanings (e.g., “scarce”). Although a subsequent study by Monaghan and Ellis (2002), suggests that this semantic effect might be at least in part the result of a confound with age-of-acquisition, as abstract low-frequency exception words tend to have late AOA, this interaction is still found when we control for AOA (Strain, Patterson, & Seidenberg, 2002). Hence, at least some of the time, we need to access a

word's semantic representation before we can access its phonology.

Does speed reading work?

Occasionally you might notice advertisements in the press for techniques for improving your reading speed. The most famous of these techniques is known as “speed reading.” Proponents of speed reading claim that you can increase your reading speed from the average of 200–350 words a minute to 2000 words a minute or even faster, yet retain the same level of comprehension. Is this possible? Unfortunately, the preponderance of psychological research suggests not. As you increase your reading speed above the normal rate, comprehension declines. Just and Carpenter (1987) compared the understanding of speed readers and normal readers on an easy piece of text (an article from *Reader's Digest*) and a difficult piece of text (an article from *Scientific American*). They found that normal readers scored 15% higher on comprehension measures than the speed readers across both passages. In fact, the speed readers performed only slightly better than a group of people who skimmed through the passages. The speed readers did as well as the normal readers on the general gist of the text, but were worse at details. In particular, speed readers could not answer questions when the answers were located in places where their eyes had not fixated.

Speed reading, then, is not as effective as normal reading. Eye movements are the key to why speed reading confers limited advantages (Rayner & Pollatsek, 1989). For a word to be processed properly, its image has to land close to the fovea and stay there for a sufficient length of time. Speed reading is nothing more than skimming through a piece of writing (Carver, 1972). This is not to say that readers obtain nothing from skimming: if you have sufficient prior information about the material, your level of comprehension can be quite good. If you speed read and then read normally, your overall level of comprehension and retention might be better than if you had just read the text normally. It is also a useful technique for preparing to read a book or article in a structured way (see Chapter 12). Finally,

associated techniques such as relaxing before you start to read might well have beneficial effects on comprehension and retention.

Evaluation of experiments on normal reading

There are two major problems with a simple dual-route model. First, we have seen that there are lexical effects on reading nonwords, which should be read by a non-lexical route that is insensitive to lexical information. Second, there are effects of regularity of pronunciation on reading words, which should be read by a direct, lexical route that is insensitive to phonological recoding.

A race model fares better. Regularity effects arise when the direct and indirect routes produce an output at about the same time, so that conflict arises between the irregular pronunciation proposed by the lexical route and the regular pronunciation proposed by the sublexical route. However, it is not clear how a race model where the indirect route uses grapheme–phoneme conversion can explain lexical effects on reading nonwords. Neither is it clear how semantics can guide the operation of the direct route.

Skilled readers have a measure of attentional or strategic control over the lexical and sublexical routes such that they can attend selectively to lexical or sublexical information (Baluch & Besner, 1991; Monsell, Patterson, Graham, Hughes, & Milroy, 1992; Zevin & Balota, 2000). For example, Monsell et al. found that the composition of word lists affected naming performance. High-frequency exception words were pronounced faster when they were in pure blocks than when they were mixed with nonwords. Monsell et al. argued that this was because participants allocated more attention to lexical information when reading the pure blocks. Participants also made fewer regularization errors when the words were presented in pure blocks (when they can rely solely on lexical processing) than in mixed blocks (when the sublexical route has to be involved).

At first sight, then, this experiment suggests that in difficult circumstances people seem able to change their emphasis in reading from using lexical information to sublexical information.

However, Jared (1997a) argued that people need not change the extent to which they rely on sublexical information, but instead might be responding at different points in the processing of the stimuli. She argued that the faster pronunciation latencies found in Monsell et al.'s experiment in the exception-only condition could just be the result of a general increase in response speed, rather than a reduction in reliance on the non-lexical route.

However, there is further evidence for strategic effects in the choice of route when reading. Using a primed naming task, Zevin and Balota (2000) found that nonword primes produce a greater dependence on sublexical processing, but low-frequency exception word primes produce a greater dependence on lexical processing. Coltheart and Rastle (1994) suggested that lexical access is performed so quickly for high-frequency words that there is little scope for sublexical involvement, but with low-frequency words or in difficult conditions people can devote more attention to one route or the other.

THE NEUROPSYCHOLOGY OF ADULT READING DISORDERS: ACQUIRED DYSLEXIA

What can studies of people with brain damage tell us about reading? This section is concerned with disorders of processing written language. We must distinguish between **acquired disorders** (which, as a result of head trauma such as stroke, operation, or head injury, lead to disruption of processes that were functioning normally beforehand) and **developmental disorders** (which do not result from obvious trauma, and which disrupt the development of a particular function). Disorders of reading are called the **dyslexias**; disorders of writing are called the **dysgraphias**. Damage to the left hemisphere will generally result in dyslexia, but as the same sites are involved in speaking, dyslexia is often accompanied by impairments to spoken language processing.

We can distinguish *central* dyslexias, which involve central, high-level reading processes, from

peripheral dyslexias, which involve lower level processes. Peripheral dyslexias include visual dyslexia, attentional dyslexia, letter-by-letter reading, and neglect dyslexia, all of which disrupt the extraction of visual information from the page. As our focus is on understanding the central reading process, we will limit discussion here to the central dyslexias. In addition, we will only look at acquired disorders in this section, and defer discussion of developmental dyslexia until our examination of learning to read.

If the dual-route model of reading is correct, then we should expect to find a double dissociation of the two reading routes. That is, we should find some patients have damage to the lexical route but can still read by the non-lexical route only, whereas we should be able to find other patients who have damage to the non-lexical route but can read by the lexical route only. The existence of a double dissociation is a strong prediction of the dual-route model, and a real challenge to any single-route model.

Surface dyslexia

People with **surface dyslexia** have a selective impairment in the ability to read irregular (exception) words. Hence they would have difficulty with “steak” compared with a similar regular relative word such as “speak.” Marshall and Newcombe (1973) and Shallice and Warrington (1980) described some early case histories. Surface dyslexics often make *over-regularization errors* when trying to read irregular words aloud. For example, they pronounce “broad” as “brode,” “steak” as

KEY TERMS

Acquired disorders: a disorder caused by brain damage is acquired if it affects an ability that was previously intact (contrasted with *developmental disorder*).

Developmental disorders: a disorder where the normal development or acquisition of a process (e.g., reading) is affected.

Dyslexias: disorder of reading.

Surface dyslexia: a type of dyslexia where the person has difficulty with exception words.

“steek,” and “island” as “eyesland.” On the other hand, their ability to read regular words and nonwords is intact. In terms of the dual-route model, the most obvious explanation of surface dyslexia is that these patients can only read via the indirect, non-lexical route: that is, it is an impairment of the lexical (direct access) processing route. The comprehension of word meaning is intact in these patients. They still know what an “island” is, even if they cannot read the word, and they can still understand it if you say the word to them.

The effects of brain damage are rarely localized to highly specific systems, and, in practice, patients do not show such clear-cut behavior as the ideal of totally preserved regular word and nonword reading, and the total loss of irregular words. The clearest case yet reported is that of a patient referred to as MP (Bub, Cancelliere, & Kertesz, 1985). She showed completely normal accuracy in reading nonwords, and hence her non-lexical route was totally preserved. She was not the best possible case of surface dyslexia, however, because she could read some irregular words (with an accuracy of 85% on high-frequency items, and 40% on low-frequency exception words). This means that her lexical route must have been partially intact. The pure cases are rarely found. Other patients show considerably less clear-cut reading than this, with even better performance on irregular words, and some deficit in reading regular words.

If patients were reading through a non-lexical route, we would not expect lexical variables to affect the likelihood of reading success. Kremin (1985) found no effect of word frequency, part of speech (noun versus adjective versus verb), or whether or not it is easy to form a mental image of what is referred to (called **imageability**), on the likelihood of reading success. Although patients such as MP, from Bub et al. (1985), show a clear frequency effect in that they make few regularizations of high-frequency words, other patients, such as HTR, from Shallice, Warrington, and McCarthy (1983) do not. Patients also make homophone confusions (such as reading “pane” as “to cause distress”).

Surface dyslexia may not be a unitary category. Shallice and McCarthy (1985) distinguished

between Type I and Type II surface dyslexia. Patients of both types are poor at reading exception words. The more pure cases, known as Type I patients, are highly accurate at naming regular words and pseudowords. Other patients, known as Type II, also show some impairment at reading regular words and pseudowords. The reading performance of Type II patients may be affected by lexical variables such that they are better at reading high-frequency, high-imageability words, better at reading nouns than adjectives and at reading adjectives than verbs, and better at reading short words than long. Type II patients must have an additional, moderate impairment to the non-lexical route, but the dual-route model can nevertheless still explain this pattern.

Phonological dyslexia

People with **phonological dyslexia** have a selective impairment in the ability to read pronounceable nonwords, called *pseudowords* (such as “sleeb”), while their ability to read matched words (e.g., “sleep”) is preserved. Phonological dyslexia was first described by Shallice and Warrington (1975, 1980), Patterson (1980), and Beauvois and Derouesné (1979). Phonological dyslexics find irregular words no harder to read than regular ones. These symptoms suggest that these patients can only read using the lexical route, and therefore that phonological dyslexia is an impairment of the non-lexical (GPC) processing route. As with surface dyslexia, the “perfect patient,” who in this case would be able to read all words but no nonwords, has yet to be discovered. The clearest case yet reported is that of patient WB (Funnell, 1983), who could not read nonwords at all; hence the non-lexical GPC route must have been

KEY TERMS

Imageability: a semantic variable concerning how easy it is to form a mental image of a word: “rose” is more imageable than “truth.”

Phonological dyslexia: a type of dyslexia where people can read words quite well but are poor at reading nonwords.

completely abolished. He was not the most extreme case possible of phonological dyslexia, however, because there was also an impairment to his lexical route; his performance was about 85% correct on words.

For those patients who can pronounce some nonwords, nonword reading is improved if the nonwords are pseudohomophones (such as “nite” for “night,” or “brane” for “brain”). Those patients who also have difficulty in reading words have particular difficulty in reading the function words that do the grammatical work of the language. Low-frequency, low-imageability words are also poorly read, although neither frequency nor imageability seems to have any overwhelming role in itself. These patients also have difficulty in reading morphologically complex words—those that have syntactic modifications called inflections. They sometimes make what are called *derivational errors* on these words, where they read a word as a grammatical relative of the target, such as reading “performing” as “performance.” Finally, they also make *visual errors*, in which a word is read as another with a similar visual appearance, such as reading “perform” as “perfume.”

There are different types of phonological dyslexia. Derouesné and Beauvois (1979) suggested that phonological dyslexia can result from disruption of either orthographic or phonological processing. Some patients are worse at reading graphemically complex nonwords (e.g., CAU, where a phoneme is represented by two letters; hence this nonword requires more graphemic parsing) than graphemically simple nonwords (e.g., IKO, where there is a one-to-one mapping between letters and graphemes), but show no advantage for pseudohomophones. These patients suffer from a disruption of graphemic parsing. Another group of patients are better at reading pseudohomophones than non-pseudohomophones, but show no effect of orthographic complexity. These patients suffer from a disruption of phonological processing. Friedman (1995) distinguished between phonological dyslexia arising from an impairment of orthographic-to-phonological processing (characterized by relatively poor function word reading but good nonword repetition) from that arising from an impairment of general

phonological processing (characterized by the reverse pattern).

Following this, a three-stage model of sublexical processing has emerged (Beauvois & Derouesné, 1979; Coltheart, 1985; Friedman, 1995). First, a *graphemic analysis* stage parses the letter string into graphemes. Second, a *print-to-sound conversion* stage assigns phonemes to graphemes. Third, in the *phonemic blending stage* the sounds are assembled into a phonological representation. There are patients whose behavior can best be explained in terms of disruption of each of these stages (Lesch & Martin, 1998). MS (Newcombe & Marshall, 1985) suffered from disruption to graphemic analysis. Patients with disrupted graphemic analysis find nonwords in which each grapheme is represented by a single letter easier to read than nonwords with multiple correspondences. WB (Funnell, 1983) suffered from disruption in the print-to-sound conversion stage; here nonword repetition is intact. ML (Lesch & Martin, 1998) was a phonological dyslexic who could carry out tasks of phonological assembly on syllables, but not on sub-syllabic units (onsets, bodies, and phonemes). MV (Bub, Black, Howell, & Kertesz, 1987) suffered from disruption to the phonemic stage.

Why do some people with phonological dyslexia have difficulty reading function words? One possibility is that function words are difficult because they are so abstract (Friedman, 1995). However, patient MC (Druks & Froud, 2002) had great difficulty in reading nonwords, morphologically complex words, and function words in isolation. Crucially he could read highly abstract content words, so it cannot be the abstractness of the function words that caused his problems. Nevertheless, he could understand the meaning of function words that he could not read, and his deficit was confined to reading single words. His reading of function words in continuous text was much better. It is likely that MC at least has a problem with syntactic processing such that when producing words in isolation he is unable to access syntactic information.

People with phonological dyslexia show complex phonological problems that have nothing to do with orthography. Indeed, it has been

proposed that phonological dyslexia is a consequence of a general problem with phonological processing (Farah, Stowe, & Levinson, 1996; Harm & Seidenberg, 2001; Patterson, Suzuki, & Wydel, 1996). If phonological dyslexia arises solely because of problems with ability to translate orthography into phonology, then there must be brain tissue dedicated to this task. This implies that this brain tissue becomes dedicated by school-age learning, which is an unappealing prospect. The alternative view is that phonological dyslexia is just one aspect of a general impairment of phonological processing. This impairment will normally be manifested in performance on non-reading tasks such as rhyming, nonword writing, phonological short-term memory, nonword repetition, and tasks of phonological synthesis (“what does ‘c – a – t spell out?’”) and phonological awareness (“what word is left if you take the ‘p’ sound out of ‘spoon?’”). This proposal also explains why pseudohomophones are read better than non-pseudohomophones. An important piece of evidence in favor of this hypothesis is that phonological dyslexia is never observed in the absence of a more general phonological deficit (but see Coltheart, 1996, for a dissenting view). A general phonological deficit makes it difficult to assemble pronunciations for nonwords. Words are spared much of this difficulty because of support from other words and top-down support from their semantic representations. Repeating words and nonwords is facilitated by support from auditory representations, so some phonological dyslexics can still repeat some nonwords. However, if the repetition task is made more difficult so that patients can no longer gain support from the auditory representations, repetition performance declines markedly (Farah et al., 1996). This idea that phonological dyslexia is caused by a general phonological deficit is central to the connectionist account of dyslexia, discussed later.

Deep dyslexia

At first sight, surface and phonological dyslexia appear to exhaust the possibilities of the consequences of damage to the dual-route model. There is, however, another even more surprising type

of dyslexia called **deep dyslexia**. Marshall and Newcombe (1966, 1973) first described deep dyslexia in two patients, GR and KU, although it is now recognized that the syndrome had been observed in patients before this (Marshall & Newcombe, 1980). In many respects deep dyslexia resembles phonological dyslexia. Patients have great difficulty in reading nonwords, and considerable difficulty in reading the grammatical, function words. Like phonological dyslexics, they make visual and derivational errors. However, the defining characteristic of deep dyslexia is the presence of semantic reading errors or **semantic paralexias**, when people produce a word related in meaning to the target instead of the target, as in examples (4) to (7):

- (4) DAUGHTER “sister”
- (5) PRAY “chapel”
- (6) ROSE “flower”
- (7) KILL “hate”

The imageability of a word is an important determinant of the probability of reading success in deep dyslexia. The easier it is to form a mental image of a word, the easier it is to read. Note that just an imageability effect in reading does not mean that patients with deep dyslexia are better at all tasks involving more concrete words. Indeed, Newton and Barry (1997) described a patient (LW) who was much better at reading high-frequency concrete words than abstract words, but who showed no impairment in comprehending those same abstract words.

Coltheart (1980) listed 12 symptoms commonly shown by deep dyslexics: They make semantic errors, they make visual errors, they substitute incorrect function words for the target, they make derivational errors, they can’t pronounce nonwords, they show an imageability effect, they

KEY TERMS

Deep dyslexia: disorder of reading characterized by semantic reading errors.

Semantic paralexias: a reading error based on a word’s meaning.

find nouns easier to read than adjectives, they find adjectives easier to read than verbs, they find function words more difficult to read than content words, their writing is impaired, their auditory short-term memory is impaired, and their reading ability depends on the context of a word (e.g., FLY is easier to read when it is a noun in a sentence than a verb).

There has been some debate about the extent to which deep dyslexia is a syndrome (a syndrome is a group of symptoms that cluster together). Coltheart (1980) argued that the clustering of symptoms is meaningful, in that they suggest a single underlying cause. However, although these symptoms tend to occur in many patients, they do not apparently necessarily do so. For example, AR (Warrington & Shallice, 1979) did not show concreteness and content word effects and had intact writing and auditory short-term memory. A few patients make semantic errors but very few visual errors (Caramazza & Hillis, 1990). Such patients suggest that it is unlikely that there is a single underlying deficit. Like phonological dyslexics, deep dyslexics obviously have some difficulty in obtaining non-lexical access to phonology via grapheme–phoneme recoding, but they also have some disorder of the semantic system. We nevertheless have to explain why these symptoms are so often associated. One possibility is that the different symptoms of deep dyslexia arise because of an arbitrary feature of brain anatomy: Different but nearby parts of the brain control processes such as writing and auditory short-term memory, so that damage to one is often associated with damage to another. As we will see, a more satisfying account is provided by connectionist modeling.

Shallice (1988) argued that there are three subtypes of deep dyslexia that vary in the precise impairments involved. *Input* deep dyslexics have difficulties in reaching the exact semantic representations of words in reading. In these patients, auditory comprehension is superior to reading. *Central* deep dyslexics have a severe auditory comprehension deficit in addition to their reading difficulties. *Output* deep dyslexics can process words up to their semantic representations, but then have difficulty producing the appropriate

phonological output. In practice it can be difficult to assign particular patients to these subtypes, and it is not clear what precise impairment of the reading systems is necessary to produce each subtype (Newton & Barry, 1997).

The right-hemisphere hypothesis

Does deep dyslexia reflect attempts by a greatly damaged system to read normally, as has been argued by Morton and Patterson (1980), among others? Or does it instead reflect the operation of an otherwise normally suppressed system coming through? Perhaps deep dyslexics do not always use the left hemisphere for reading. Instead, people with deep dyslexia might use a reading system based in the right hemisphere that is normally suppressed (Coltheart, 1980; Saffran, Bogyo, Schwartz, & Marin, 1980; Zaidel & Peters, 1981). This right-hemisphere hypothesis is supported by the observation that the more of the left hemisphere that is damaged, the more severe the deep dyslexia observed (Jones & Martin, 1985; but see Marshall & Patterson, 1985). Furthermore, the reading performance of deep dyslexics resembles that of split-brain patients when words are presented to the left visual field, and therefore to the right hemisphere. Under such conditions they also make semantic paralexias, and have an advantage for concrete words. Finally, Patterson, Vargha-Khadem, and Polkey (1989) described the case of a patient called NI, a 17-year-old girl who had had her left hemisphere removed for the treatment of severe epilepsy. After recovery she retained some reading ability, but her performance resembled that of deep dyslexics.

In spite of these points in its favor, the right-hemisphere reading hypothesis has never won wide acceptance. In part this is because the hypothesis is considered a negative one, in that if it were correct, deep dyslexia would tell us nothing about normal reading. In addition, people with deep dyslexia read much better than split-brain patients who are forced to rely on the right hemisphere for reading. The right-hemisphere advantage for concrete words is rarely found, and the imageability of the target words used in these experiments might have been confounded with

length (Ellis & Young, 1988; Patterson & Besner, 1984). Finally, Roeltgen (1987) described a patient who suffered from deep dyslexia as a result of a stroke in the left hemisphere. He later suffered from a second left hemisphere stroke, which had the effect of destroying his residual reading ability. If the deep dyslexia had been a consequence of right hemisphere reading, it should not have been affected by the second stroke in the left hemisphere.

Summary of research on deep dyslexia

There has been debate as to whether the term “deep dyslexia” is a meaningful label. The crucial issue is whether or not its symptoms must necessarily co-occur because they have the same underlying cause. Are semantic paralexias always found associated with impaired nonword reading? So far they seem to be; in all reported cases semantic paralexias have been associated with all the other symptoms. How then can deep dyslexia be explained by one underlying disorder? In terms of the dual-route model, there would need to be damage to both the semantic system (to explain the semantic paralexias and the imageability effects) and the non-lexical route (to explain the difficulties with nonwords). We would also then have to specify that for some reason damage to the first is always associated with damage to the second (e.g., because of an anatomical accident that the neural tissue supporting both processes is in adjoining parts of the brain). This is inelegant. As we shall see, connectionist models have cast valuable light on this question. A second issue is whether we can make inferences from deep dyslexia about the processes of normal reading, as we can for the other types of acquired dyslexia. We have seen that the dual-route model readily explains surface and phonological dyslexia, and that their occurrence is as expected if we were to lesion that model by removing one of the routes. Hence it is reasonable to make inferences about normal reading on the basis of data from such patients. There is some doubt, however, as to whether we are entitled to do this in the case of deep dyslexia; if the right-hemisphere hypothesis were correct, deep

dyslexia would tell us little about normal reading. The balance of evidence is at present that deep dyslexia does not reflect right-hemisphere reading, but does reflect reading by a greatly damaged left hemisphere. Deep dyslexia suggests that normally we can in some way read through meaning; that is, we use the semantic representation of a word to obtain its phonology. This supports our earlier observation that with homographs (e.g., “bow”) we use the meaning to select the appropriate pronunciation.

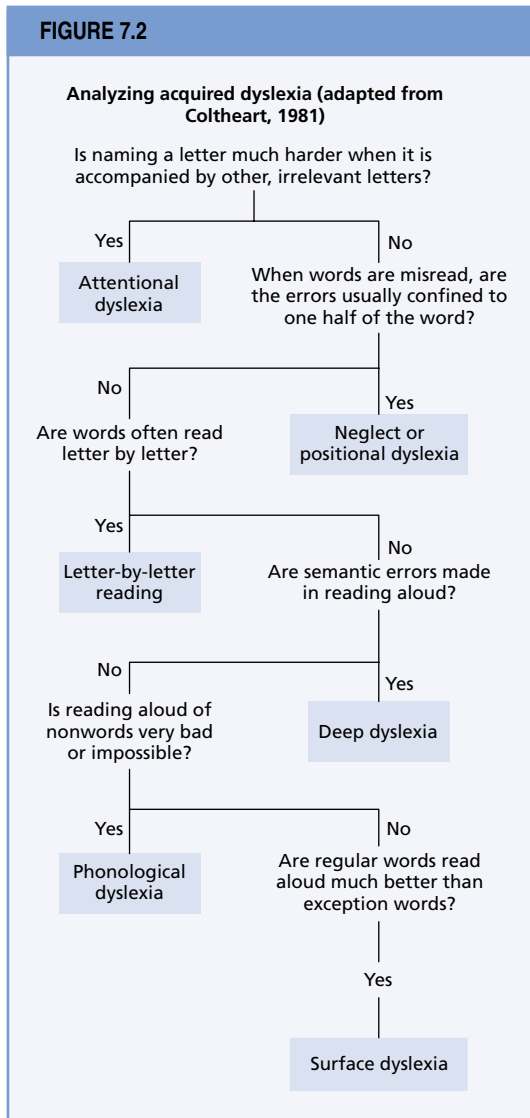
Non-semantic reading

Schwartz, Marin, and Saffran (1979), and Schwartz, Saffran, and Marin (1980a) described WLP, an elderly patient suffering from progressive dementia. WLP had a greatly impaired ability to retrieve the meaning of written words; for example, she was unable to match written animal names to pictures. She could read those words out aloud almost perfectly, getting 18 out of 20 correct and making only minor errors, even on low-frequency words. She could also read irregular words and nonwords. In summary, WLP could read words without any comprehension of their meaning. Coslett (1991) described a patient, WT, who was virtually unable to read nonwords, suggesting an impairment of the indirect route of the dual-route model, but who was able to read irregular words quite proficiently, even though she could not understand those words. These case studies suggest that we must have a direct access route from orthography to phonology that does not go through semantics.

Summary of the interpretation of the acquired dyslexias

We have looked at four main types of adult central dyslexia: surface, phonological, deep, and non-semantic reading. We have seen how a dual-route model explains surface dyslexia as an impairment of the lexical, direct access route, and explains phonological dyslexia as an impairment of the non-lexical, phonological recoding route. The existence of non-semantic reading suggest that the simple dual-route model needs refinement. In

particular, the direct route must be split into two. There must be a *non-semantic direct access route* that retrieves phonology given orthography, but which does not pass through semantics first, and a *semantic direct access route* that passes through semantics and allows us to select the appropriate sounds of non-homophonic homographs (e.g., “wind”). In non-semantic reading, the semantic direct route has been abolished but the non-semantic direct route is intact. An analysis of acquired dyslexia by Coltheart (1981) is shown in Figure 7.2.

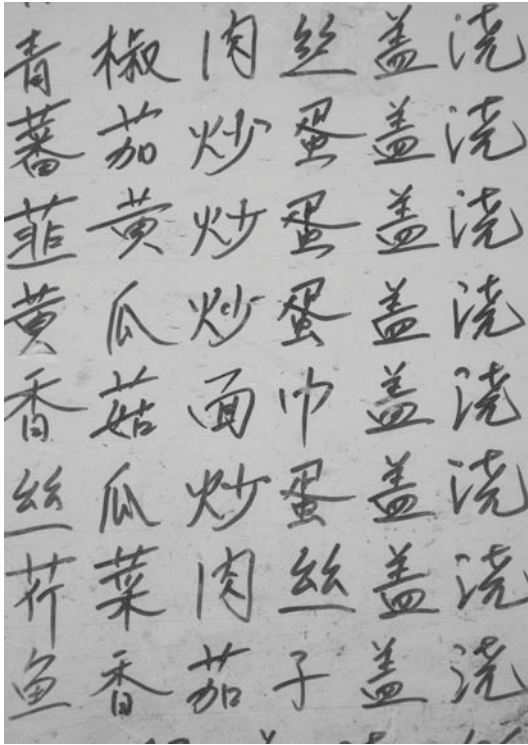
FIGURE 7.2

Acquired dyslexia in other languages

Languages such as Italian, Spanish, or Serbo-Croat, which have totally transparent or *shallow* alphabetic orthographies—that is, where every grapheme is in a one-to-one relation with a phoneme—can show phonological and deep dyslexia, but not surface dyslexia, defined as an inability to read exception words (Patterson, Marshall, & Coltheart, 1985a, 1985b). However, we can find the symptoms that can co-occur with an impairment of exception word reading, such as homophone confusions, in the languages that permit them (Masterson, Coltheart, & Meara, 1985).

Whereas languages such as English have a single, alphabetic script, Japanese has two different scripts, kana and kanji (see Coltheart, 1980; Sasanuma, 1980). Kana is a syllabic script, and kanji is a logographic or ideographic script. Therefore words in kanji convey no information on how a word should be pronounced. While kana allows sublexical processing, kanji must be accessed through a direct, lexical route. The right hemisphere is better at dealing with kanji, and the left hemisphere is better at reading kana (Coltheart, 1980). Reading of briefly presented kana words is more accurate when they are presented to the right visual field (left hemisphere), but reading of kanji words is better when they are presented to the left visual field (right hemisphere). The analog of surface dyslexia is found in patients where there is a selective impairment of reading kanji, but the reading of kana is preserved. The analog of phonological dyslexia is an ability to read both kana and kanji, but a difficulty in reading Japanese nonwords. The analog of deep dyslexia is a selective impairment of reading kana, while the reading of kanji is preserved. For example, patient TY could read words in both kanji and kana almost perfectly, but she had great difficulty with nonwords constructed from kana words (Sasanuma, Ito, Patterson, & Ito, 1996).

Chinese is an ideographic language. Butterworth and Wengang (1991) reported evidence of two routes in reading in Chinese. Ideographs can be read aloud either through a route that associates the symbol with its complete pronunciation, or through one that uses parts of the symbol.



Kanji (shown here) is a logographic or ideographic script, providing no information on word pronunciation.

(Although Chinese is non-alphabetic, most symbols contain some sublexical information on pronunciation.) Each route can be selectively impaired by brain damage, leading to distinct types of reading disorder.

The study of other languages that have different means of mapping orthography onto phonology is still at a relatively early stage, but it is likely to greatly enhance our understanding of reading mechanisms. The findings suggest that the neuropsychological mechanisms involved in reading are universal, although there are obviously some differences related to the unique features of different orthographies.

MODELS OF WORD NAMING

Both the classic dual-route and the single-route, lexical-instance models face a number of problems.

First, there are lexical effects for nonwords and regularity effects for words, and therefore reading cannot be a simple case of automatic grapheme-to-phoneme conversion for nonwords, and automatic direct access for all words. Single-route models, on the other hand, appear to provide no account of nonword pronunciation, and it remains to be demonstrated how neighborhood effects affect a word's pronunciation. Second, any model must also be able to account for the pattern of dissociations found in dyslexia. While surface and phonological dyslexia indicate that two reading mechanisms are necessary, other disorders suggest that these alone will not suffice. At first sight it is not obvious how a single-route model could explain these dissociations at all.

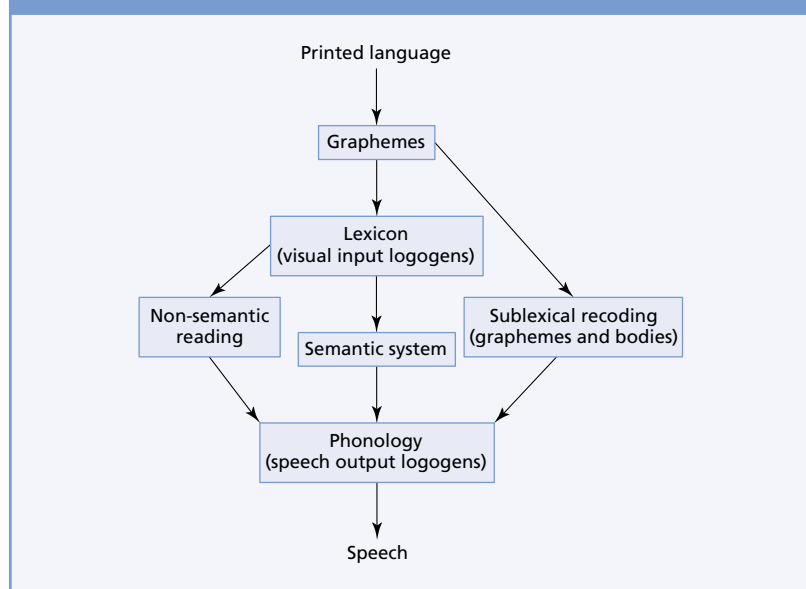
Theorists have taken two different approaches depending on their starting point. One possibility is to refine the dual-route model. Another is to show how word-neighborhoods can affect pronunciation, and how pseudowords can be pronounced in a single-route model. This led to the development of analogy models. More recently, a connectionist model of reading has been developed that takes the single-route, analogy-based approach to the limit.

The revised dual-route model

We can save the dual-route model by making it more complex. Morton and Patterson (1980) and Patterson and Morton (1985) described a three-route model (see Figure 7.3). First, there is a non-lexical route for assembling pronunciations from sublexical grapheme-phoneme conversion. The non-lexical route now consists of two subsystems. A standard *grapheme-phoneme conversion mechanism* is supplemented with a *body subsystem* that makes use of information about correspondences between orthographic and phonological rimes. This is needed to explain lexical effects on nonword pronunciation. Second, the direct route is split into a semantic and a non-semantic direct route.

The three-route model accounts for the data as follows. The lexical effects on nonwords and regularity effects on words are explained by cross-talk between the lexical and non-lexical routes. Two

FIGURE 7.3



The revised dual-route model of reading.

types of interaction are possible: interference during retrieval, and conflict in resolving multiple phonological forms after retrieval. The two sub-systems of the non-lexical route also give the model greater power. Surface dyslexia is the loss of the ability to make direct contact with the orthographic lexicon, and phonological dyslexia is the loss of the indirect route. Non-semantic reading is a loss of the lexical-semantic route. Deep dyslexia remains rather mysterious. First, we have to argue that these patients can only read through the lexical-semantic route. While accounting for the symptoms that resemble phonological dyslexia, it still does not explain the semantic paralexias. One possibility is that this route is used normally, but not always successfully, and that it needs additional information (such as from the non-lexical and non-semantic direct route) to succeed. So when this information is no longer available it functions imperfectly. It gets us to the right semantic area, but not necessarily to the exact item, hence giving paralexias. This additional assumption seems somewhat arbitrary. An alternative idea is that paralexias are the result of additional damage to the semantic system itself. Hence a complex pattern of impairments is still necessary

to explain deep dyslexia, and there is no reason to suggest that these are not dissociable.

Multi-route models are becoming increasingly complicated as we find out more about the reading process (for example, see Carr & Pollatsek, 1985). Another idea is that multiple levels of spelling-to-sound correspondences combine in determining the pronunciation of a word. In Norris's (1994a) *multiple-levels model*, different levels of spelling-to-sound information, including phoneme, rime (the final part of the word giving rise to the words with which it rhymes, e.g., "eak" in "speak"), and word-level correspondences, combine in an interactive activation network to determine the final pronunciation of a word. Such an approach develops earlier models that make use of knowledge at multiple levels, such as Brown (1987), Patterson and Morton (1985), and Shallice, Warrington, and McCarthy (1983).

The most recent version of the dual-route model is the *dual-route cascaded*, or *DRC*, model (Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart & Rastle, 1994; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). This is a computational model based on the architecture of the dual-route model—although it is in fact misleadingly

so called, as it is really based on the three-route model, with a non-lexical grapheme–phoneme rule system and a lexical system, which in turn is divided into one route that passes through the semantic system and a non-semantic route that does not. The model makes use of cascaded processing, in that as soon as there is any activation at the letter level, activation is passed on to the word level. The computational model can simulate performance on both lexical decision and naming tasks, showing appropriate effects of frequency, regularity, pseudohomophones, neighborhood, and priming. Regularity is now a central motivation of the model; words are either regular, or they are not. Irregular words take longer to pronounce than regular ones because the lexical and non-lexical routes produce conflicting pronunciations. The model accounts for surface dyslexia by making entries in the orthographic lexicon less available, and for phonological dyslexia by damaging the grapheme–phoneme conversion route.

There is not uniform agreement that it is necessary to divide the direct route into two. In the *summation model* (Hillis & Caramazza, 1991b; Howard & Franklin, 1988), the only direct route is reading through semantics. How does this model account for non-semantic reading? The idea is that access to the semantic system is not completely obliterated. Activation from the sublexical route combines (or is “summed”) with activation trickling down from the damaged direct semantic route to ensure the correct pronunciation.

It is difficult to distinguish between these variants of the original dual-route model, although the three-route version provides the more explicit account of the dissociations observed in dyslexia. There is also some evidence against the summation hypothesis. EP (Funnell, 1996) could read irregular words that she could not name, and priming the name with the initial letter did not help her naming, contrary to the prediction of the summation hypothesis. Many aspects of the dual-route model have been subsumed by the triangle model that serves as the basis of connectionist models of reading. The situation is complicated even more by the apparent co-occurrence of the loss of particular word meanings in dementia and surface dyslexia (see later).

The analogy model

The analogy model arose in the late 1970s when the extent of lexical effects on nonword reading and differences between words became apparent (Glushko, 1979; Henderson, 1982; Kay & Marcel, 1981; Marcel, 1980). It is a form of single-route model that provides an explicit mechanism for how we pronounce nonwords. It proposes that we pronounce nonwords and new words by analogy with other words. When a word (or nonword) is presented, it activates its neighbors, and these all influence its pronunciation. For example, “gang” activates “hang,” “rang,” “sang,” and “bang”; these are all consistent with the regular pronunciation of “gang,” and hence assembling a pronunciation is straightforward. When presented with “base,” however, “case” and “vase” are activated; these conflict and hence the assembly of a pronunciation is slowed down until the conflict is resolved. A nonword such as “taze” is pronounced by analogy with the consistent set of similar words (“maze,” “gaze,” “daze”). A nonword such as “mave” activates “gave,” “rave,” and “save,” but it also activates the conflicting enemy “have,” which hence slows down pronunciation of “mave.” In order to name by analogy, you have to find candidate words containing appropriate orthographic segments (like “-ave”); obtain the phonological representation of the segments; and assemble the complete phonology (“m + ave”).

Although attractive in the way they deal with regularity and neighborhood effects, early versions of analogy models suffered from a number of problems. First, the models did not make clear how the input is segmented in an appropriate way. Second, the models make incorrect predictions about how some nonwords should be pronounced. Particularly troublesome are nonwords based on gangs; “pook” should be pronounced by analogy with the great preponderance of the gang comprising “book,” “hook,” “look,” and “rook,” yet it is given the “hero” pronunciation (see Table 7.1)—which is in accordance with grapheme–phoneme correspondence rules—nearly 75% of the time (Kay, 1985). Analogy theory also appears to make incorrect predictions about how long it takes us

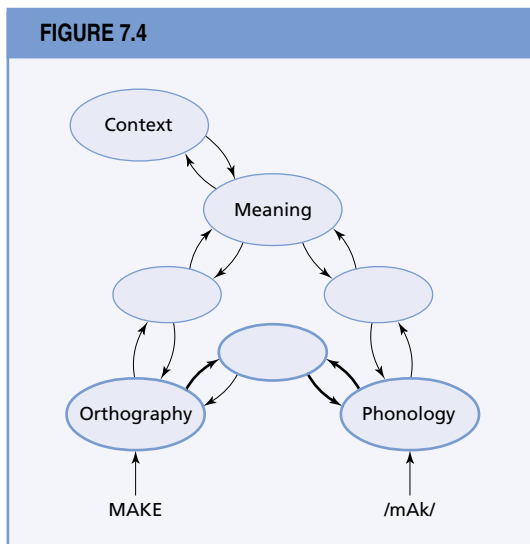
to make regularization errors (Patterson & Morton, 1985). Finally, it is not clear how analogy models account for the dissociations found in acquired dyslexia. Nevertheless, in some ways the analogy model was a precursor of connectionist models of reading.

Connectionist models

The original Seidenberg and McClelland (1989) model evolved in response to criticisms that I will examine after describing the original model.

Seidenberg and McClelland (1989)'s model of reading

The Seidenberg and McClelland (1989) model (often abbreviated to SM) shares many features with the interactive activation model of letter recognition discussed in Chapter 6. The SM model provides an account of how readers recognize letter strings as words and pronounce them. This first model simulated one route of a more general model of lexical processing (see Figure 7.4).



Seidenberg and McClelland's (1989) "triangle model" of word recognition. Implemented pathways are shown in bold. Reproduced with permission from Harm and Seidenberg (2001).

Reading and speech involve three types of code: orthographic, meaning, and phonological. These are connected with feedback connections. The shape of the model has given it the name of the *Triangle Model*. As in the revised dual-route model, there is a route from orthography to phonology by way of semantics. The key feature of the model is that there is only one other route from orthography to phonology; there is no route involving grapheme–phoneme correspondence rules.

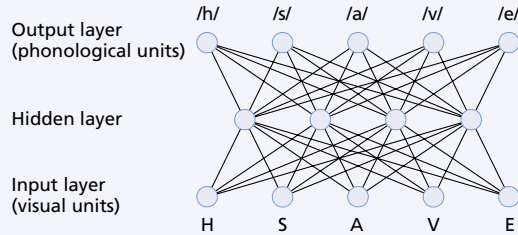
Seidenberg and McClelland (1989) just simulated the orthographic-to-phonology part of the overall triangle model. The model has three levels, each containing many simple units. These are the input, hidden, and output layers (see Figure 7.5). Each of the units in these layers has an activation level, and each unit is connected to all the units in the next level by a weighted connection, which can be either excitatory or inhibitory. An important characteristic of this type of model is that the weights on these connections are not set by the modelers, but are learned. This network learns to associate a phonological output with an orthographic input by being given repeated exposure to word-pronunciation pairs. It learns using an algorithm called **back-propagation**. This involves slowly reducing the discrepancy between the desired and actual outputs of the network by changing the weights on the connections. (See the Appendix for more information.)

Seidenberg and McClelland used 400 units to code orthographic information for input and 460 units to code phonological information for output, mediated by 200 hidden units. Phonemes and graphemes were encoded as a set of triples, so that each grapheme or phoneme was specified with its flanking grapheme or phoneme. This is

KEY TERM

Back-propagation: an algorithm for learning input–output pairs in connectionist networks. It works by alternately reducing the error between the actual output and the desired output of the network.

FIGURE 7.5



The layers of Seidenberg and McClelland's (1989) model of word recognition (simplified—see text for details). Based on Seidenberg and McClelland (1989).

a common trick to represent position-specificity (Wickelgren, 1969). For example, the word “have” was represented by the triples “#ha,” “hav,” “ave,” “ve#,” with “#” representing a blank space. A non-local representation was used: The graphemic representations were encoded as a pattern of activation across the orthographic units rather than corresponding directly to particular graphemes. Each phoneme triple was encoded as a pattern of activation distributed over a set of units representing phonetic features—a representation known as a *Wickelfeature*. The underlying architecture was not a simple feed-forward one, in that the hidden units fed back to the orthographic units, mimicking top-down word-to-letter connections in the IAC model of word recognition. However, there was no feedback from the phonological to the hidden units, so phonological representations could not directly influence the processing of orthographic-level representations.

The training corpus comprised all 2897 uninflected monosyllabic words of at least three or more letters in the English language present in the Kucera and Francis (1967) word corpus. Each trial consisted of the presentation of a letter string that was converted into the appropriate pattern of activation over the orthographic units. This in turn fed forward to the phonological units by way of the hidden units. In the training phase, words were presented a number of times with a probability proportional to the logarithm of their frequency. This means that the ease with which a word is learned by the network, and the effect it has on similar words, depends to some extent on its frequency. About 150,000 learning trials were

needed to minimize the differences between the desired and actual outputs.

After training, the network was tested by presenting letter strings and computing the orthographic and phonological *error scores*. The error score is a measure of the average difference between the actual and desired output of each of the output units, across all patterns. *Phonological error scores* were generated by applying input to the orthographic units, and measured by the output of the phonological units; they were interpreted as reflecting performance on a naming task. *Orthographic error scores* were generated by comparing the pattern of activation input to the orthographic units with the pattern produced through feedback from the hidden units, and were interpreted as a measure reflecting the performance of the model in a lexical decision task. Orthographic error scores are therefore a measure of orthographic familiarity. Seidenberg and McClelland showed that the model fitted human data on a wide range of inputs. For example, regular words (such as “gave”) were pronounced faster than exception words (such as “have”).

Note that the Seidenberg and McClelland model uses a single mechanism to read nonwords and exception words. There is only one set of hidden units, and only one process is used to name regular, exception, and novel items. As the model uses a distributed representation, there is no one-to-one correspondence between hidden units and lexical items; each word is represented by a pattern of activation over the hidden units. According to this model, lexical memory does not consist of entries for individual words. Orthographic

neighbors do not influence the pronunciation of a word directly at the time of processing; instead, regularity effects in pronunciation derive from statistical regularities in the words of the training corpus—all the words we have learned—as implemented in the weights of connections in the simulation. Lexical processing therefore involves the activation of information, and is not an all-or-none event.

Evaluation of the original SM model

Coltheart et al. (1993) criticized important aspects of the Seidenberg and McClelland (SM) model. They formulated six questions about reading that any account of reading must answer:

- How do skilled readers read exception words aloud?
- How do skilled readers read nonwords aloud?
- How do participants make visual lexical decision judgments?
- How does surface dyslexia arise?
- How does phonological dyslexia arise?
- How does developmental dyslexia arise?

Coltheart et al. then argued that Seidenberg and McClelland's model only answered the first of these questions.

Besner, Twilley, McCann, and Seergobin (1990) provided a detailed critique of the Seidenberg and McClelland model, although a reply by Seidenberg and McClelland (1990) answered some of these points. First, Besner et al. argued that in a sense the model still possesses a lexicon, where instead of a word corresponding to a unit, it corresponds to a pattern of activation. Second, they pointed out that the model “reads” nonwords rather poorly—certainly much less well than a skilled reader. In particular, it only produced the “correct,” regular pronunciation of a nonword under 70% of the time. This contrasts with the model's excellent performance on its original training set. Hence the model's performance on nonwords is impaired from the beginning. In reply, Seidenberg and McClelland (1990) pointed out that their model was trained on only 2987 words, as opposed to the 30,000 words that people know, and that this

may be responsible for the difference. Hence the model simulates the direct lexical route rather better than it simulates the indirect grapheme–phoneme route. Therefore any disruption of the model will give a better account of disruption to the direct route—that is, of surface dyslexia. The model's account of lexical decision is inadequate in that it makes far too many errors—in particular it accepts too many nonwords as words (Besner et al., 1990; Fera & Besner, 1992). The model did not perform as well as people do on nonwords, in particular on nonwords that contain unusual spelling patterns (e.g., JINJE, FAIJE). In addition, the model's account of surface dyslexia was problematic and its account of phonological dyslexia non-existent.

Forster (1994) evaluated the assumptions behind connectionist modeling of visual word recognition. He made the point that showing that a network model can successfully learn to perform a complex task such as reading does not mean that that is the way humans actually do it. Finally, Norris (1994b) argued that a major stumbling block for the Seidenberg and McClelland model was that it could not account for the ability of readers to shift strategically between reliance on lexical and sublexical information.

The revised connectionist model: PMSP

A revised connectionist model performs much better at pronouncing nonwords and at lexical decision than the original (Plaut, 1997; Plaut & McClelland, 1993; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg, Petersen, MacDonald, & Plaut, 1996; Seidenberg, Plaut, Petersen, McClelland, & McRae, 1994). The model, called PMSP for short, used more realistic input and output representations. Phonological representations were based on phonemes with phonotactic constraints (that constrain which sounds occur together in the language), and orthographic representations were based on graphemes with graphotactic constraints (that constrain which letters occur together in the language). The original SM model performed badly on nonwords because Wickelfeatures *disperse* spelling–sound regularities. For example, in GAVE, the A is

represented in the context of G and V, and has nothing in common with the A in SAVE (represented in the context of S and V). In the revised PMSP model, letters and phonemes activate the same units irrespective of context. A mathematical analysis showed that a response to a letter string input is a function that depends positively on the frequency of exposure to the pattern, positively to the sum of the frequencies of its friends, and negatively to the sum of the frequencies of its enemies. The response to a letter string is non-linear, in that there are diminishing returns: For example, regular words are so good they gain little extra benefit from frequency. This explains the interaction we observe between word consistency and frequency. As we shall see, the revised model also gives a much better account of dyslexia.

Accessing semantics

Of course the goal of reading is to access the meaning of words. The PSMP model simulates the orthography–phonology side of the triangle. Clearly, according to the model, we can access semantics either directly (OS: orthography–semantics) or indirectly (OPS: orthography–phonology–semantics—what we have also called phonological mediation). Hence there is a *division of labor* between the two routes. Harm and Seidenberg (2004) model the access of semantics. In the full model, all parts of the system operate simultaneously and contribute to the activation of meaning. The Harm and Seidenberg model is a complete implementation of the triangle model. It is trained to produce the correct pattern of activation across a set of semantic features given an orthographic input. In the first phase, the model is trained for a while on the phonology–semantics side of the triangle, to simulate the knowledge of young children who cannot yet read, but who know what words mean. These weights were then frozen. In the second phase, the orthography–phonology and orthography–semantics sides of the triangle were then trained.

How does the trained model perform? Perhaps not surprisingly, in simulations resembling the skilled reader in normal conditions, the OS route

is normally faster, with the OPS route lagging somewhat behind. Nevertheless, analysis of how activation of the input determines activation of the output shows that activation of the semantic system is driven by both pathways. Even if the OPS path is slower, it still always contributes to the final output. In addition, because of interactivity in the system, activation of the semantic system activates corresponding phonological representations, which in turn affect the semantic system. Simulations show that the relative contributions of the two pathways (OS and OPS) are modulated by a number of factors, including skill (phonological information is more important early on in training, corresponding to less skilled readers) and word frequency (for high-frequency words the OS pathway is more efficient). The model also simulates the response times of Van Orden (1987), where people are slow to say “no” to “Is it a flower? ROWS.”

Connectionist models of dyslexia

Over the last few years connectionist modeling has contributed to our understanding of deep and surface dyslexia.

Modeling surface dyslexia

Patterson, Seidenberg, and McClelland (1989) artificially damaged or “lesioned” the Seidenberg and McClelland (1989) network after the learning phase by destroying hidden units or connection weights, and then observed the behavior of the model. Its performance resembled the reading of a surface dyslexic. Patterson et al. (1989) explored three main types of lesion: damage to the connections between the orthographic input and hidden units (called early weights); damage to the connections between the hidden and output (phonological) units (called late weights), and damage to the hidden units themselves. Damage was inflicted by probabilistically resetting a proportion of the weights or units to zero. The greater the amount of damage being simulated, the higher the proportion of weights that was changed. The consequences were measured in two ways. First, the damage was measured by the

phonological error score, which as we have seen reflects the difference between the actual and target activation values of the phonological output units. Obviously, high error scores reflect impaired performance. Second, the damage was measured by the reversal rate. This corresponds to a switch in pronunciation by the model, so that a regular pronunciation is given to an exception item (for example, “have” is pronounced to rhyme with “gave”).

Increasing damage at each location produces near-linear increases in the phonological error scores of all types of word. On the whole, though, the lesioned model performed better with regular than with exception words. The reversal rate increased as the degree of damage increased, but nevertheless there were still more reversals occurring on exception words than on regular words. Damage to the hidden units in particular produced a large number of instances where exception words were produced with a regular pronunciation; this is similar to the result whereby surface dyslexics over-regularize their pronunciations. However, the number of regularized pronunciations that were produced by the lesioned model was significantly lower than that produced by surface dyslexic patients. No lesion made the model perform selectively worse on nonwords. Hence the behavior of the lesioned model resembles that of a surface dyslexic.

Patterson et al. also found that word frequency was not a major determinant of whether a pronunciation reversed or not. (It did have some effect, so that high-frequency words were generally more robust to damage.) As we have seen, some surface dyslexics show frequency effects on reading, while others do not. Patterson et al. found that the main determinant of reversals was the number of vowel features by which the regular pronunciation differs from the correct pronunciation, a finding verified from the neuropsychological data.

An additional point of interest is that the lesioned model produced errors that have traditionally been interpreted as “visual” errors. These are mispronunciations that are not over-regularizations and that were traditionally thought to result from an impairment of early graphemic

analysis. If this analysis is correct, then Patterson et al. should only have found such errors when there was damage to the orthographic units involved. In contrast, they found them even when the orthographic units were not damaged. This is an example of a particular strength of connectionist modeling; the same mechanism explains what were previously considered to be disparate findings. Here visual errors result from the same lesion that causes other characteristics of surface dyslexia, and it is unnecessary to resort to more complex explanations involving additional damage to the graphemic analysis system.

There are three main problems with this particular account. First, we have already seen that the original Seidenberg and McClelland model was relatively bad at producing nonwords before it was lesioned. We might say that the original model is already operating as a phonological dyslexic. Yet surface dyslexics are good at reading nonwords. Second, the model does not really over-regularize, it just changes the vowel sound of words. Third, Behrmann and Bub (1992) reported data that are inconsistent with this model. In particular, they showed that the performance of the surface dyslexic MP on irregular words does vary as a function of word frequency. They interpreted this frequency effect as problematic for connectionist models. Patterson et al. (1989) were quite explicit in simulating only surface dyslexia; their model does not address phonological dyslexia.

Exploring semantic involvement in reading

The revised model, abbreviated to PSMP, provides a better account of dyslexia. The improvements come about because the simulations implement both pathways of the triangle model in order to explain semantic effects on reading.

Surface dyslexia arises in the progressive neurological disease dementia (see Chapter 11 on semantics for details of dementia). Importantly, people with dementia find exception words difficult to pronounce and repeat if they have lost the meaning of those words (Hodges, Patterson, Oxbury, & Funnell, 1992; Patterson & Hodges,

1992; but see Funnell, 1996). Patterson and Hodges proposed that the integrity of lexical representations depends on their interaction with the semantic system: Semantic representations bind phonological representations together with a semantic glue; hence this is called the *semantic glue hypothesis*. As the semantic system gradually dissolves in dementia, so the semantic glue gradually comes unstuck, and the lexical representations lose their integrity. Patients are therefore forced to rely on a sublexical or grapheme–phoneme correspondence reading route, leading to surface dyslexic errors. Furthermore, they have difficulty in repeating irregular words for which they have lost the meaning, if the system is sufficiently stressed (by repeating lists of words), but they can repeat lists of words for which the meaning is intact (Patterson, Graham, & Hodges, 1994; but see Funnell, 1996, for a patient who does not show this difference).

PMSP showed that a realistic model of surface dyslexia depends on involving semantics in reading. Support from semantics normally relieves the phonological pathway from having to master low-frequency exception words by itself. In surface dyslexia the semantic pathway is damaged, and the isolated phonological pathway reveals itself as surface dyslexia.

Plaut (1997) further examined the involvement of semantics in reading. He noted that some patients have substantial semantic impairments but can read exception words accurately (e.g., DC of Lambon Ralph, Ellis, & Franklin, 1995; DRN of Cipolotti & Warrington, 1995; WLP of Schwartz, Marin, & Saffran, 1979). To explain why some patients with semantic impairments cannot read exception words but some can, Plaut suggested that there are individual differences in the division of labor between semantic and phonological pathways. Although the majority of patients with semantic damage show surface dyslexia (Graham, Hodges, & Patterson, 1994), some exceptions are predicted. He also argued that people use a number of strategies in performing lexical decision, one of which is to use semantic familiarity as a basis for making judgments. The revised model therefore takes into account individual differences between speakers, and shows how small differences

in reading strategies can lead to different consequences after brain damage.

Modeling phonological dyslexia

The triangle model provides the best connectionist account of phonological dyslexia. It envisages reading as taking place through the three routes conceptualized in the original SM model. The routes are orthography to phonology, orthography to semantics, and semantics to phonology (Figure 7.3). This approach sees phonological dyslexia as nothing other than a general problem with phonological processing (Farah et al., 1996, Sasanuma et al., 1996). Phonological dyslexia arises through impairments to representations at the phonological level, rather than to grapheme–phoneme conversion. This is called the *phonological impairment hypothesis*. People with phonological dyslexia can still read words because their weakened phonological representations can be accessed through the semantic level. (Hence this approach is also a development of the semantic glue hypothesis.) We have already noted that the original Seidenberg and McClelland (1989) model performed rather like a phonological dyslexic patient, in that it performed relatively poorly on nonwords. Consistent with the phonological deficit hypothesis, the explanation for this poor performance was that the source of these errors was the impoverished phonological representations used by the model.

An apparent problem with the phonological deficit hypothesis is that it is not clear that it would correctly handle the way in which people with phonological dyslexia read pseudohomophones better than other types of nonwords (Coltheart, 1996). Furthermore, patient LB of Derouesné and Beauvois (1985) showed an advantage for pseudohomophones, but no obvious general phonological impairment. There have also been effects of orthographic complexity and visual similarity, suggesting that there is also an orthographic impairment present in phonological dyslexia (Derouesné & Beauvois, 1985; Howard & Best, 1996). For example, Howard and Best showed that their patient Melanie-Jane read pseudohomophones that were visually

similar to the related word (e.g., GERL) better than pseudohomophones that were visually more distant (e.g., PHOCKS). There was no effect of visual similarity for control nonwords. However, Harm and Seidenberg (2001) show how phonological impairment in a connectionist model can give rise to such effects. A phonological impairment magnifies the ease with which different types of stimuli are read.

Modeling deep dyslexia

Hinton and Shallice (1991) lesioned another connectionist model to simulate deep dyslexia. Their model was trained by back-propagation to associate word pronunciations with a representation of the meaning of words. This model is particularly important, because it shows that one type of lesion can give rise to all the symptoms of deep dyslexia, particularly both paralexias and visual errors.

The underlying semantic representation of a word is specified as a pattern of activation across semantic feature units (which Hinton & Shallice called *sememes*). These correspond to semantic features or primitives such as “main-shape-2D,” “has-legs,” “brown,” and “mammal.” These can be thought of as atomic units of meaning (see Chapter 11). The architecture of the Hinton and Shallice (1991) model comprised 28 graphemic input units and 68 semantic output units with an intervening hidden layer containing 40 intermediate units. The model was trained to produce an appropriate output representation given a particular orthographic input using back-propagation. The model was trained on 40 uninflected monosyllabic words.

The structure of the output layer is quite complex. First, there were interconnections between some of the semantic units. The 68 semantic feature units were divided into 19 groups depending on their interpretation, with inhibitory connections between appropriate members of the group. For example, in the group of semantic features that define the size of the object denoted by the word, there are three semantic features: “max-size-less-foot,” “max-size-foot-to-two-yards,” and “max-size-greater-two-yards.” Each of these

features inhibits the others in the group, because obviously an object can only have one size. Second, an additional set of hidden units called *cleanup units* was connected to the semantic units. These permit more complex interdependencies between the semantic units to be learned, and have the effect of producing structure in the output layer. This results in a richer semantic space where there are strong semantic *attractors*. An attractor can be seen as a point in semantic space to which neighboring states of the network are attracted; it resembles the bottom of a valley or basin, so that objects positioned on the sides of the basin tend to migrate towards the lowest point. This corresponds to the semantic representation ultimately assigned to a word.

As in Patterson et al.’s (1989) simulation of surface dyslexia, different types of lesion were possible. There are two dimensions to remember: one is what is lesioned, the other is how it is lesioned. The connections involved were the grapheme–intermediate, intermediate–sememe, and sememe–cleanup. Three methods of lesioning the network were used. First, each set of connections was taken in turn, and a proportion of their weights was set to zero (effectively disconnecting units). Second, random noise was added to each connection. Third, the hidden units (the intermediate and cleanup units) were ablated by destroying a proportion of them.

The results showed that the closer the lesion was to the semantic system, the more effect it had. The lesion type and site interacted in their effects; for example, the cleanup circuit was more sensitive to added noise than to disconnections. Lesions resulted in four types of error: semantic (where an input gave an output word that was semantically but not visually close to the target; these resemble the classic semantic paralexias of deep dyslexics); visual (words visually but not semantically similar), mixed (where the output is both semantically and visually close to the

KEY TERM

Attractors: a point in the connectionist attractor network to which related states are attracted.

target), and others. All lesion sites and types (except for that of disconnecting the semantic and cleanup units) produced the same broad pattern of errors. Finally, on some occasions the lesions were so severe that the network could not generate an explicit response. In these cases, Hinton and Shallice tested the below-threshold information left in the system by simulating a forced-choice procedure. They achieved this by comparing the residual semantic output to a set of possible outputs corresponding to a set of words, one of which was the target semantic output. The model behaved above chance on this forced-choice test, in that its output semantic representation tended to be closer to that of the target than to the alternatives.

Hence the lesioned network behaves like a deep dyslexic patient, in particular in making semantic paralexias. The paralexias occur because semantic attractors cause the accessing of feature clusters close to the meanings of words that are related to the target. A “landscape” metaphor may be useful. Lesions can be thought of as resulting in the destruction of the ridges that separate the different basins of attraction. The occurrence of such errors does not seem to be crucially dependent on the particular lesion type or site under consideration. Furthermore, this account provides an explanation of why different error types, particularly semantic and visual errors, nearly always co-occur in such patients. Two visually similar words can point in the first instance to nearby parts of semantic space, even though their ultimate meanings in the basins may be far apart; if you start off on top of a hill, going downhill in different directions will take you to very different ultimate locations. Lesions modify semantic space so that visually similar words are then attracted to different semantic attractors.

Hinton and Shallice’s account is important for cognitive neuropsychologists for a number of reasons. First, it provides an explicit mechanism whereby the characteristics of deep dyslexia can be derived from a model of normal reading. Second, it shows that the actual site of the lesion is not of primary importance. This is mainly because of the “cascade” characteristics of these networks. Each stage of processing is continu-

ally activating the next, and is not dependent on the completion of processing by its prior stage (McClelland, 1979). Therefore, effects of lesions at one network site are very quickly passed on to surrounding sites. Third, it shows why symptoms that were previously considered to be conceptually distinct necessarily co-occur. Semantic and visual errors can result from the same lesion. Fourth, it thus revives the importance of syndromes as a neuropsychological concept. If symptoms co-occur as a result of any lesion to a particular system, then it makes sense to look for and study such co-occurrences.

Plaut and Shallice (1993a) extended this work to examine the effect of word abstractness on lesioned reading performance. As we have seen, the reading performance of deep dyslexic patients is significantly better on more imageable than on less imageable words. Plaut and Shallice showed that the richness of the underlying semantic representation of a word is an analog of imageability. They hypothesized that the semantic representations of abstract words contain fewer semantic features than those of concrete words; that is, the more concrete a word is, the richer its semantic representation. Jones (1985) showed that it was possible to account for imageability effects in deep dyslexia by recasting them as *ease-of-predication* effects. Ease-of-predication is a measure of how easy it is to generate things to say about a word, or predicates, and is obviously closely related to the richness of the underlying semantic representation. It is easier to find more things to say about more imageable words than about less imageable words. Plaut and Shallice (1993a) showed that when an attractor network similar to that of Hinton and Shallice (1991) is lesioned, concrete words are read better than abstract words. One exception was that severe lesions of the cleanup system resulted in better performance on abstract words. Plaut and Shallice argue that this is consistent with patient CAV (Warrington, 1981), who showed such an advantage. Hence this network can account for both the usual better performance of deep dyslexic patients on concrete words, and also the rare exception where the reverse is the case. They also showed that lesions closer to the grapheme units tended

to produce more visual errors, whereas lesions closer to the semantic units tended to produce more semantic errors. The model also provides an account of the behavior of normal participants reading degraded words (McLeod, Shallice, & Plaut, 2000). If words are presented very rapidly to people, they make both visual and semantic errors. The data fit the connectionist model well.

Connectionist modeling has advanced our understanding of deep dyslexia in particular, and neuropsychological deficits in general. The finding that apparently unrelated symptoms can necessarily co-occur as a result of a single lesion is of particular importance. It suggests that deep dyslexia may after all be a unitary condition. However, there is one fly in the ointment. The finding that at least some patients show imageability effects in reading but not in comprehension is troublesome for all models that posit a disturbance of semantic representations as the cause of deep dyslexia (Newton & Barry, 1997). Instead, in at least some patients, the primary disturbance may be to the speech production component of reading.

Comparison of models

A simple dual-route model provides an inadequate account of reading, and needs at least an additional lexical route through imageable semantics. The more complex a model becomes, the greater the worry that routes are being introduced on an arbitrary basis to account for particular findings. Analogy models have some attractive features, but their detailed workings are vague and they do not seem able to account for all the data. Connectionist modeling has provided an explicit, single-route model that covers most of the main findings, but has its problems. At the very least it has clarified the issues involved in reading. Its contribution goes beyond this, however. It has set the challenge that only one route is necessary in reading words and nonwords, and that regularity effects in pronunciation arise out of statistical regularities in the words of the language. It may not be a complete or correct account, however it is certainly a challenging one.

Currently we are faced with two serious alternatives: a connectionist model such as the triangle model, and a variant of the dual-route

model such as the dual-route cascaded model. The literature is full of claim and counter-claim, and it would be presumptuous for a text like this to say that one is clearly right and the other wrong. There are many studies providing support for and against one or the other of the models. Many of them focus on how we read nonwords (Besner et al., 1990; Seidenberg et al., 1994), because the division of labor in the DRC model between a lexical route with knowledge of individual words and a non-lexical route with spelling rules is absent in connectionist models, and this difference is the key one between the two sorts of models. The DRC emphasizes *regularity* (does the word obey the rule?), which is a categorical concept—either the word obeys the spelling-sound rules or it does not, with nonwords having to be pronounced by the rule. The triangle model emphasizes *consistency* of rimes and other units (how often is -AVE pronounced in a certain way?), which is a statistical concept. According to Zevin and Seidenberg (2006), consistency effects such as those shown in Glushko's (1979) and Jared's (1997b) study are the critical test between models. Words like PAVE are regular but inconsistent; according to the DRC model they should be as easy to pronounce as regular and consistent words such as PANE; according to the triangle model they should not. Now of course we know from Glushko's study that regular inconsistent words are slower to pronounce than regular consistent ones, but Coltheart et al. (2001) argue that these differences are an artifact arising from several confounding factors (e.g., the presence of exception words in the materials, and an increase in the number of times it is necessary to reanalyze inconsistent words as we read them from left to right). Zevin and Seidenberg (2006) argued that graded sensitivity to consistency effects in nonwords provides the critical test between the models, with only connectionist models correctly predicting the presence of such effects, and being able to account for individual differences in nonword pronunciation. However, doubtless this debate will run and run.

Perhaps the choice between the triangle and the dual-route cascaded model comes down to which one values most: explaining a wide range of data, or parsimony in design.

Balota (1990) asked if there is a magic moment when we recognize a word but do not yet have access to its meaning. He argued that the tasks most commonly used to study word processing (lexical decision and word naming) are both sensitive to post-access processes. This makes interpretation of data obtained using these tasks difficult (although not, as we have seen, impossible). Furthermore, deep dyslexia (discussed later) suggests that it is possible to access meaning without correctly identifying the word, while non-semantic reading suggests that we can recognize words without necessarily accessing their meaning. Whereas unique lexical access is a prerequisite of activating meaning in models such as the logogen and the serial search model, cascading

connectionist models permit the gradual activation of semantic information while evidence is still accumulating from perceptual processing. A model such as the triangle model (Patterson, Suzuki, & Wydell, 1996; Plaut, McClelland, Seidenberg, & Patterson, 1996) seems best able to accommodate all these constraints.

Finally, all of these models—particularly the connectionist ones—are limited in that they have focused on the recognition of morphologically simple, often monosyllabic words. Rastle and Coltheart (2000) have developed a rule-based model of reading bisyllabic words, emphasizing how we produce the correct stress, and Ans, Carbonnel, and Valdois (1998) have developed a connectionist model of reading polysyllabic words.

SUMMARY

- Different languages use different principles to translate words into sounds; languages such as English use the alphabetic principle.
- Regular words have a regular grapheme-to-phoneme correspondence, but exception words do not.
- According to the dual-route model, words can be read through a direct lexical route or a sublexical route; in adult skilled readers the lexical route is usually faster.
- The sublexical route was originally thought to use grapheme–phoneme conversion, but now it is considered to use correspondences across a range of sublexical levels.
- There are effects of lexical similarity in reading certain nonwords (pseudohomophones), while not all words are read with equal facility (the consistency of the regularity of a word's neighbors affects its ease of pronunciation).
- It might be necessary to access the phonological code of a word before we can access its meaning; this process is called phonological mediation.
- Phonological mediation is most likely to be observed with low-frequency words and with poor readers.
- Readers have some attentional control over which route they emphasize in reading.
- Access to some phonological code is mandatory, even in silent reading, but normally does not precede semantic access.
- Increasing reading speed above about 350 words a minute (by speed reading, for example) leads to reduced comprehension.
- Surface dyslexia is difficulty in reading exception words; it corresponds to an impairment of the lexical route in the dual-route model.
- Phonological dyslexia is difficulty in reading nonwords; it corresponds to an impairment of the sublexical route in the dual-route model.
- Deep dyslexic readers display a number of symptoms including making visual errors, but the most important characteristic is the presence of semantic reading errors or paralexias.
- There has been some debate as to whether deep dyslexia is a coherent syndrome.
- Non-semantic readers can pronounce irregular words even though they do not know their meaning.
- The revised dual-route model uses multiple sublexical correspondences and permits direct access through a semantic lexical route and a non-semantic lexical route.

- The dual-route cascaded model allows activation to trickle through levels before processing is necessarily completed at any level.
- Seidenberg and McClelland (SM) produced an important connectionist model of reading; however, it performed poorly on nonwords and pseudohomophones.
- Lesioning the SM network gives rise to behavior resembling surface dyslexia, but its over-regularizations differ from those made by humans.
- The revised version of this model, PSMP, gives a much better account of normal reading and surface dyslexia; it uses a much more realistic representation for input and output than the original model.
- There are clear semantic influences on normal and impaired reading, and recent connectionist models are trying to take these into account.
- The triangle model accounts for phonological dyslexia as an impairment to the phonological representations: this is the phonological impairment hypothesis.
- Deep dyslexia has been modeled by lesioning semantic attractors; the lesioned model shows how the apparently disparate symptoms of deep dyslexia can arise from one type of lesion.
- More imageable words are relatively spared because they have richer semantic representations.
- There has been considerable debate as to whether developmental dyslexia is qualitatively different from very poor normal reading, and whether there are subtypes that correspond to acquired dyslexias; the preponderance of evidence suggests that developmental dyslexia is on a continuum with normal reading.
- Connectionist modeling shows how two distinct types of damage can lead to a continuum of impairment between development surface and phonological dyslexia extremes.

SOME QUESTIONS TO THINK ABOUT

1. Is there a “magic moment” when we recognize a word but cannot access its meaning?
2. Why might reading errors occur? Keep a record of any errors you make and try to relate them to what you have learned in this and the previous chapter.
3. What practical tips could help adult dyslexic readers to read more effectively?

FURTHER READING

Many of the references at the end of Chapter 6 will also be relevant here. There are a number of works that describe the orthography of English, and discuss the rules whereby certain spelling-to-sound correspondences are described as regular and others as irregular. One of the best known of these is Venezky (1970). For an example of work on reading in a different orthographic system, see Kess and Miyamoto (1999).

For a general introduction to reading, writing, spelling, and their disorders, see Ellis (1993). For more discussion of dyslexia, including peripheral dyslexias, see Ellis and Young (1988). Two volumes (entitled *Deep dyslexia*, 2nd edition, by Coltheart, Patterson, & Marshall, 1987, and *Surface dyslexia* by Patterson, Marshall, & Coltheart, 1985b) cover much of the relevant material. A special issue of the journal *Cognitive Neuropsychology* (1996, volume 13, part 6) was devoted to phonological dyslexia.

For recent overviews of reading, see Andrews (2006) and Snowling and Hulme (2007).

Subject Index

Entries given in **bold** indicate
glossary definitions

- AB, 351
- Abstraction-based theories, 338
- AC, 343
- Access files, 193
- Accessibility, 377–378
- Accommodation, 80, 401, 403
- Acoustic invariance, 258–259
- Acoustic phonetics, 29
- Acoustics, 27, **491**
- Acquired disorders, 220, **491**
- Acquired dyslexia, 220–227
- Acquisition and learning
 - distinction hypothesis, 159
- ACT/ACT* models, 379
- Activation, 13, 190, 265, 327–328, 425, 485, 486, **491**
- Active-filler strategy, 314
- Active filter hypothesis, 160
- Adjacency pairs, 458
- Adjectives, 36, **491**
- Adverbs, 36, **491**
- Affirmative declarative sentences, 40
- Affixes, 403, **491**
 - addition/loss, 411
 - stripping, 191
- Affordances, 345
- Affricatives, 33
- Age-of-acquisition (AOA), 174–175, 218
- Agent, 83, 136, 287, 288, **491**
- Agnosia, 185, 344, **491**
- Agrammatism, 316–318, 437–439, **491**
- Akeakamai (dolphin), 58
- Alex (parrot), 57
- Alexia, **491**
- Algorithms, 485
- Alignment, 459
- Allan, 250
- Allophones, 28, **491**
- Alphabetic principle, 247–248
- Alphabetic reading, 241
- Alphabetic scripts, 210
- Alveolars, 32
- Alzheimer's disease (AD), 146, 350–353, 351, 448, **491**
- Ambiguity, 38, 199–208, 289–291, 298–313, 321, 329, 374–377, 460
- American English, 29
- American Sign Language (ASL/AMESLAN), 59–60, 61, 62, 78, **492**
- Analogy models, 229–230, 238
- Analysis-by-synthesis, 267–268
- Analytic phonics, 248
- Anaphor resolution, 374
- Anaphora, 374–277, **491**
- Aneurysms, 443, **491**
- Angular gyrus, 69
- Animal language, 54–67
- Animate-inanimate dissociation, 346–350
- Anomia, 71, 344, 417, 441–442, 468–469, **491**
- Antecedents, 374, **491**
- Antonyms, 323
- Ants, 54
- Apes, 53–54, 58–67
- Aphasia, 68, 70–71, 281, 411, 435–437, 439–440, 445–446, **491**
 - bilinguals, 158
 - connectionist models, 442–445
- Apraxia, 435, **491**
- AR, 224, 343
- Arabic, 111, 149, 210
- Arcuate fasciculus, 14, 15, 69
- Argument structure, 42, 141–144, 287, **492**
- Articulatory apparatus, 52, **492**
- Articulatory phonetics, 29
- Artificial intelligence (AI), 11–13, 328–329
- Artificial languages, 58, 118
- AS, 281
- ASL, *see* American Sign Language
- Aspect, 405, **492**
- Aspirated/Aspiration, 28, **492**
- Assimilation, 80, 259, **492**
- Associative priming, 185, 186–187
- Attachment, 309, **492**
- Attachment preference, 307–308, **492**
- Attentional priming, 180, 181
- Attentional processing, 178–181, 190, 318, **492**
- Attractors, 236, 354, **492**
 - network, 354
- Audience design, 291, 459–460
- Audiolingual method, 159
- Auditory phonetics, 29
- Auditory short-term memory (ASTM), 472–474, **492**

- Augmented transition network (ATN), **492**
 Austin (chimp), 61, 66
 Autism, 82
 Automata theory, 43–45
 Automatic priming, 180, 181, 186–187
 Automatic processing, 178–181, 318, **492**
 Autonomous access model, 203
 Autonomous models, 21, 193–195, 198–199, 266, 288, 289, 299–301
 Auxiliary verbs, 40, **492**
 AW, 411

 Babbling, 123, **492**
 Baby talk, *see* Child-directed speech (CDS)
 Back-channel communication, 458
 Back-propagation, 146–147, 230, 485, 487–489, **492**
 Backtrack, **492**
 Backward inferences, 369
 Backward-looking centers, 378
 Backward translation, 157
 Backwards masking, 172
 Balanced ambiguity, 203
 Basic color terms, 94, 95
 Basic level, 130, 336–337, **492**
 Basque, 7
 Bassa, 94
 Batonic gestures, 434
 BC, 346
 Bees, 54–55
 Behaviorism, 10
 Berinmo, 96
 Berlitz method, 159
 Bias, 423, 488
 Bilabial sounds, 32
 Bilingual, 153, **492**
 Bilingualism, 153–158, 161–162, 260, 417, 423, **492**
 Binding, 314
 Bins, 193
 Blindness, 84–86, 87
 Blocking hypothesis, 417
 Blocksworld, 12
 BO, 391, 475
 Bobby, 249
 Body, 215, **492**
 subsystem, 228
 Boltzman machines, 489
 Bonding, 372
 Bonobos, 64, 66

 Bootstrapping, 121, 122–123, 130, 136, 137, **492**
 Bottom-up priority, 269
 Bottom-up processing, 21, 479–480, **492**
 Bound morphemes, 403, **492**
 Boundary effect, 386
 Bow-wow theory, 52
 Box-and-arrow diagrams, 11, 15, 481
 Box-and-candle problem, 92, 93
 Boxology, 11, 481
 Brain
 evolution, 52
 hemisphere dominance, 53–54, 68, 72–75, 158
 imaging, 16–19, 68, 482
 lesion studies, 14–16, 68, 73
 localization of functions, 67–71
 structure, 15
 Bridging inferences, 369
 Brightness masking, 172
 British English, 29
 Broca's aphasia, 68, 435–437, 445–446, **492**
 Broca's area, 14, 15, 52, 53, 68, 70, 72, 438
 Buffers, **492**

 Canonical sentence strategy, 296
 Cantonese, 33
 Capacity theory, 474
 Cascade models, 21, 420, 421, 423, 425, 427, **492**
 Case grammar, 379
 CAT (computerized axial tomography), 17
 Categorical perception, 260–261, 262–263, **493**
 Categorization, 127, 322–323
 Category-specific disorders, 346–350, 355–356
 Causal coherence, 362
 Causative verbs, 144
 CAV, 237
 CB, 440, 444
 CELEX database, 173
 Center-embedding, 39
 Centering theory, 378
 Central dyslexias, 220
 Central executive, 439, 471
 Cerebral vascular accident (CVA), **493**
 Characteristic features, 330, 332
 Child-directed speech (CDS), 83, 108–110, 121, 148, **493**
 Chimpanzees, 58–67

 Chinchillas, 122
 Chinese, 33, 92, 93, 97, 210, 218, 226–227, 244
 “Chinese readers”, 251
 Classification, 335–338
 Clauses, 37, 294, **493**
 Cleanup units, 236
 Cleft sentences, **493**
 “Clever Hans” effect, 57
 Click displacement, 294–295
 Clitics, **493**
 Closed-class items, 37, **493**
 Closed-class words, *see* Function words
 Closure, 297
 Clue words, 246
 Co-articulation, 259, **493**
 Coda, 33
 Code switching, 154–155
 Cognates, 156, **493**
 Cognition hypothesis, 79–82, 87
 Cognitive constraint, 323
 Cognitive cycles, 434
 Cognitive economy, 322
 Cognitive intrusions, 404
 Cognitive linguistics, 43
 Cognitive neuropsychology, 14–16, 480, 481, 482
 Cognitive neuroscience, 14
 Cognitive psychology, 10
 Cognitive science, 11–13
 Coherence, 279, 362
 Coherence graph, 386
 Cohesion, 362
 Cohort model, 268–273, 280
 Co-indexation, **493**
 Collective monologues, 80
 Color, 94–96
 Common ground, 377
 Common-store models, 155–156
 Communication, 5, 9, 54, 161
 Communication accuracy, 94
 Comparative linguistics, 10
 Compensation for co-articulation, 276
 Competence, 34–35, **493**
 Competition, 305–306, 309
 Competition-integration model, 306, 309
 Competitive queuing, 429
 Complement, **493**
 Complementizers, 290–291, **493**
 Complex word processing, 191–192
 Compound-cue theory, 190
 Comprehensible input hypothesis, 160

- Comprehension, 12–13, 218, 361–393
- Computational metaphor, 10–11
- Concepts, 322, 336, 338–339
- Conceptual dependency theory, 379–380
- Conceptual mediation, 157
- Conceptual pacts, 457
- Conceptual Selection Model (CSM), 414
- Conceptualization, 397–398
- Concrete operational stage, 79–80
- Conditioned head turn technique, 104–105
- Conditioning, 10, 44, 106–108
- Conduction aphasia, 446
- Conduit d'approche, 446
- Confidence, 161
- Conjoint frequency, 327
- Conjunctions, 36, **493**
- Connectionism, 13, 23–24, 481, 485–489, **493**
- age-of-acquisition, 175
 - aphasia, 442–445
 - dementia, 352–353, 355–356
 - distributional analysis, 139
 - distributional information, 117–118
 - dyslexia, 233–238
 - lexical ambiguity, 206–207
 - past tense acquisition, 146–148
 - phonological encoding, 429
 - reading, 230–233, 238–239
 - semantic memory, 352–353
 - semantics, 353–358
 - speech production, 425, 426, 429
 - speech recognition, 273–280
- Connectionist neuropsychology, 481, 482
- Connotations, 324
- Conservation, 79–80, 82
- Consolidated alphabetic phase, 242
- Consonantal scripts, 210
- Consonants, 31–33, **493**
- processing, 281
- Constituents, 37, **493**
- Constraints, 43, 298–299, 309
- Construction-integration model, 386–388
- Constructivist-semantic view, 137
- Content words, 36–37, 318, 401–402, **493**
- Context, 161, 187–190, 203–206, 261, 262–264, 266–267
- Context-free grammars, 40, 45
- Context-guided single-reading lexical access model, 200
- Context-sensitive grammars, 40, 45
- Context-sensitive model, 204–205
- Contingent negative variation (CNV), 17
- Continuity hypothesis, 111, 112, 123
- Contrastive hypothesis, 134, 158–159
- Controlled processing, 178, **493**
- Conversation, 362, 454–460
- analysis, 458
 - back channels, 458
 - collaboration, 459–460
 - cooperation, 456–457
 - implicatures, 457, **495**
 - inferences, 454–458
 - maxims, 456–458, **493**
 - structure, 458–459
 - turn-taking, 83–84, 458–459
- Conversational hypothesis, 109
- Cooperative principle, 456–457
- Coordinates, 323
- Copulative verbs, **493**
- Core description, 330, 331
- Co-reference, 374, **493**
- Counterfactual, **493**
- Creoles, 114, **493**
- Critical period, 72–79
- Cross-fostering, 59
- Cross-linguistic differences, 120, 148–149, 307–308, 480, 483, **493**
- Cross-modal priming, 186, 271
- Cross-nurturing, 59
- Cross-sectional studies, 105
- Crossed aphasia, 73, 158, 446
- Cue validity, 180
- Cultural transmission, 60
- Culture, 161
- Culture-specific information, 368
- CW, 348
- D/P (declarative/procedural) model, 71
- Dani, 94–95, 96
- Dante, 417
- Data, 4
- Data-driven, 21
- Dative verbs, **493**
- David, 112
- Deafness, 76, 78, 86–87, 109, 112, 114, 123, 218, 281, 468
- Dean, John, 363
- Decay, 487
- Decompositional theories, 328–335
- Deep dysgraphia, 448
- Deep dyslexia, 223–225, 228, 236–238, 239, 354, **493**
- subtypes, 224
- Deep dysphasia, 443, 468, **493**
- Deep structure (d-structure), 41
- Defining features, 330, 332
- Degenerate input, 108
- Deictic function, 312–313
- Dementia, 234–235, 300, 344, 350–353, 355–356, 391–392, 470, 475
- Denotation, 324
- Dentals, 32
- Dependencies, 314–315, 407–409
- Derivation, **493**
- Derivational errors, 222
- Derivational morphology, 5, 6, **494**
- Derivational theory of complexity (DTC), 292, 293
- Design features, 55–57
- Determiners, 36, **494**
- Detransformation, 292, 293
- Developmental disorders, 220, **494**
- Developmental dysgraphia, 249
- Developmental dyslexia, 249–255
- DF, 255
- Dialects, 29
- Dialog, 459–460
- Diary studies, 154
- Dichotic-listening task, 201
- Digging-in, 306
- Ding-dong theory, 52
- Diphthongs, 31, **494**
- Direct methods, 159
- Direct object, 37
- Direct speech acts, 456
- Disambiguation region, 290
- Disconnection syndrome, 69
- Discontinuity hypothesis, 124
- Discourse, 361–362, **494**
- analysis, 458
 - neuropsychology of processing, 390–392
- Discrete stage models, 21, **494**
- Dissociation, 22, 403, **494**
- Distinctive features, 31
- Distinguishing features, 355, 356
- Distributional analysis, 24, 138–139
- Distributional information, 117–118, 121, **494**
- Ditransitive verbs, 37

- DM, 300, 351, 445
Dogs, 54, 57
Dolphins, 55, 58
Domain-specific knowledge
 hypothesis (DSKH), 350
Double dissociation, 16, 220,
 347–348, 411, 445–446, **494**
Down's syndrome, 81
DP, 445
DSMSG model, 443–445
DT, 470
Dual-code hypothesis, 343
Dual-pathway hypothesis, 191
Dual-route cascaded (DRC)
 model, 228–229, 238
Dual-route models
 of reading, 211–212, 214, 219,
 225–226, 227–229, 238
 of spelling, 248
Dutch, 7, 96, 120, 307
Dyirbal, 337
Dysarthria, **494**
Dysgraphias, 220, 448, **494**
 developmental, 249
Dyslexias, 185, 217, 220, **494**
 acquired, 220–227
 connectionist models, 233–238
 developmental, 249–255
Dysphasia, 443, 468
Dysprosody, 435, **494**
- E-language (externalised
 language), 35
E-Z Reader model, 170
Early-syntax theories, 143
Ease-of-predication effects, 237
Echolalia, 80
EDE, 282
EE, 468
EEGs (Electroencephalography),
 17, **494**
Egocentric speech, 80
Elaborative inferences, 369
ELIZA, 11–12, 13
Elizabeth (chimp), 61
EM, 343, 351
Emergentist account, 73
Empiricism, 105, 106
Endogenous oscillators, 458–459
Enemies, 215
Energy masking, 172
English, 7, 34, 42
Entrainment, 457, 459
Entrenchment hypothesis, 143
Environmental contamination, 404
EP, 229
Epilinguistic knowledge, 243
Episodic memory, 322, **494**
Equipotentiality hypothesis, 73
ERPs (Event-related potentials),
 17, **494**
 lateralization, 74, 158
 sentence processing, 295, 301
 speech recognition, 272
 word recognition, 156, 189
Error-reduction learning, 487
Error scores, 231
Eskimo, 90, 91
EST, 442
Evolution, 52–54
Excitation, 486
Execution, 397, 398
Exemplar theory, 337–338
Exercise hypothesis, 75
Expressive, **494**
Extension, 324
Eye fixation, 169, 216, 297
Eye movements
 comprehension, 372, 375
 reading, 168–170, 188–189,
 216, 219
 sentence processing, 295, 299,
 300, 303–304, 306, 309–
 310, 312, 461–462
 word recognition, 203, 267
Eye-tracking, 149, 157
Eye-witness testimony, 363–364,
 372
- Face management, 457
Face perception, 185
Facilitation, 14, 177, 486, **494**
Familiarity, 173–174
Familiarity bias, 423
Family resemblance, 335–338
Fan effect, 379
Fast-mapping, 127
Featural and Unitary Semantic
 Space hypothesis, 335
Feature comparison theory, 330–
 331, 332
Feature lists, 328, 330–331
Feature masking, 172
Feedback, 21, 83, 106–108, 423–
 424, 427–428
Felicia, 349
Felicity conditions, 454
Feral children, 76–77
Figurative speech, 340–342, **494**
Filipino, 90
Filled hesitations, 432
Filled pauses, 432, 458
Fillers, 314, **494**
Finite-state devices, 44–45
- Finnish, 7, 149
Finno-Ugric languages, 7
First mention, 377–378
Fixations, 169, 216
Fixed-choice two-stage models,
 309
Fixed structure, 297
Flow diagrams, 11
Fluent aphasia, 411, 436
Fluent restorations, 270
fMRI (functional magnetic
 resonance imaging), 18, 146,
 156, 307, 350
Focal colors, 94–96
Focus, 376
Foregrounding, 376
Form-based priming, 177
Formal operations stage, 80
Formal paraphasias, 443, **494**
Formal universals, 112–113
Formants, 27–28, **494**
Formulation, 397, 398
Forward-looking centers, 378
Forward translation, 157
Four Cs, 161
Fovea, 169
FOXP2 gene, 53, 67, 115
Frame-based models, 428–429
Frame problem, 369
French, 7, 34, 111, 210, 307
Frequency attenuation, 176
Frequency effect, 173–175, 182–
 184
Freudian slips, 399
Fricatives, 33
Friends, 215
Full alphabetic phase, 242
Full-listing hypothesis, 191
Function words, 36–37, 318,
 401–402, **494**
Functional architecture, 15
Functional core hypothesis, 133
Functional fixedness, 92
Fuzzy prototypes, 276
- Gaps, 313–315, **494**
Garden path model, 298, 299–
 300, 309
Garden path sentences, 290–291,
 494
Gating tasks, 271, **494**
Gaze, 83, 84, 405, 458
Gender, 92, 138, 375, 417–418,
 494
Generalized delta rule, 488
Generative grammar, **494**
Generative semantics, 379

- Genetic linguistics, 114–115
 Genie, 77–78
 German, 7, 92, 146, 245, 246
 Gerunds, **494**
 Gestures, 53–54, 58, 86, 434, 458
 Given–new contract, 378
 GL, 442
 Glides, 33
 Global aphasia, 446
 Global modularity, 421
 Globality assumption, 445
 Glottal stops, 32–33, **494**
 Glottis, 32
 Goal, 288, **494**
 Gold's theorem, 115–116
 Gorillas, 61
 Government and Binding (GB) theory, 41
 GR, 223
 Grammar, 34–45, **494**
 learning, 83, 108
 Grammatical constraints, 323
 Grammatical differences, 91–92
 Grammatical elements, **494**
 Grammaticality judgments, 66, 316
 Grapheme-to-phoneme conversion (GPC) route, 211, 212
 Grapheme–phoneme conversion mechanism, 227
 Grapheme–phoneme correspondence, 29, 210, 245–246
 Graphemes, 209–210, **494**
 analysis, 222
 Greek, 97, 246
 Grounding, 357–358
 Group monologues, 80
 Gua (chimp), 59
 Guessing model, 183

 HAL, 11
 HAM (Human Associative Memory), 379
 Hard-wired modules, 22
 Harm and Seidenberg model, 233
 Harmonics, 29, 31
 HDM (high-dimensional memory) models, 357
 Head-first/head-last languages, 42
 Head of phrase, 42
 Hearing impairment, *see* Deafness
 Heave-ho theory, 52
 Hebbian learning, 489
 Hebrew, 138, 169, 210
 Hemidecortication, 73–74, **495**

 Hemisphere dominance, 53–54, 68, 72–75, 158
 Here-and-now model, 385
 Hesitations, 432–435
 Heterographic homophones, 199, **495**
 HIC (High Interactional Content), 365
 Hidden units, 488, **495**
 Holophrastic speech, 136
 Homographs, 199, 218, **495**
 Homonyms, **495**
 Homophones, 191, 199, 216, 418, **495**
 Honey bees, 54–55
 Horizontal information flow, 424
 HTR, 221
 Hungarian, 7
 Hybridization, 339
 Hypothesis, 4

 I-language (internalised language), 35
 Iconic gestures, 434
 Ideas, 434
 Identification procedures, 330–331
 Identification semantics hypothesis, 346
 Ideographic languages, 210, 218
 Idioms, 341–342, **495**
 IFA, 281
 IG, 442
 IL, 442
 Illocutionary force, 454, 455
 Imageability, 177, 218, 221, 223, **495**
 Imitation, 106
 Immersion method, 159, 161
 Implicatures, 457, **495**
 Importance, 365–366
 Indirect object, 37
 Indirect speech acts, 456
 Indo-European languages, 7
 Induction, 115
 Infarct, **495**
 Inference, 361, 368–373, 454–458, **495**
 Infinitive, **495**
 Inflection, 6, **495**
 Inflectional morphology, 5–6, **495**
 Information flow, 420, 423, 424
 Information theory, 10
 Informational load, 376–377
 Informationally encapsulated, 463
 Informative signals, 54
 Inhibition, 14, 177, 486, **495**

 Initial contact, 264–265
 Innateness, 22–23, 110–118, 136–137, 350, 480
 Inner speech, 22, 217–218, 465, **495**
 Input account, 344
 Instance theories, 337–338, 339
 Instantiation principle, 338
 Integration, 203–204, 265
 Intension, 324
 Interaction, 262, 266, 288, 290, 301–308, 420
 Interactive activation and competition (IAC) model, 197–198, 485–487
 Interactive activation models, 197–198, 424–425, 426, 480, 485–487
 Interactive direct access models, 192–193, 195–199
 Interactive processing, 21, 262, 480
 Intercorrelated features, 332, 355–356
 Interference hypothesis, 416
 Interlopers, 416
 International Phonetic Alphabet (IPA), 29, 30
 Interpretative context, 267
 Interrogative sentences, 40
 Intersubjectivity, 129
 Intertranslatability, 90
 Intonation, 120
 Intransitive verbs, 37, 38, **495**
 Inuit, 90, 91
 Invariance hypothesis, 73
 Irregular forms, 6
 Irregular words, 211, 220–221
 Irreversible determinism, 73
 Irreversible passives, 293
 ISA link, 325
 Isabelle, 78
 Isolated children, 76–78
 Isolation point, 265
 Italian, 7, 92, 98, 111, 210, 226
 Iteration, 39

 James IV, 76
 Japanese, 28, 33, 42, 120, 210, 226
 Jargon aphasia, 439–440
 JB, 473
 JBN, 445
 JBR, 346, 347, 348
 JCU, 441–442
 Jim, 76
 JJ, 348

- JM, 252
 Joint attention, 83, 129
 JV, 344
- Kana, 210, 226
 Kanji, 210, 218, 226, 227
 Kannada, 144
 Kanzi (chimp), 64, 65, 66
 KC, 440
 KE (family), 114–115
 KE (patient), 427
 Kernel sentences, 40, 292
 Keywords, 161
 Kikuyu, 120
 Kinship terms, 328
 Koko (gorilla), 61
 Korean, 113
 KR, 349
 KU, 223
- L1 and L2, 154, **495**
 Labile stage, 170
 Labiodentals, 32
 Lana (chimp), 61, 63
 Language
 biological basis, 67–79
 before birth, 119–120
 change, 7–8
 cognitive basis, 79–82
 defined, 5, 55
 described, 27–47
 design features, 55–57
 development, 72–79, 84–87,
 103–151, 480, 482
 differences, 8, 19–20
 functions, 3–4, 9
 mixing, 154
 origin and evolution, 7, 51–54
 production, 397–450
 social aspects, 76, 82–84, 129
 switching, 155
 system structure, 463–476,
 482–483
 and thought, 87–98
 of thought, 288
 units, 6
 Language acquisition device
 (LAD), 67, 110–118, 483,
 495
 Language acquisition socialization
 system (LASS), 83
 Language bioprogram hypothesis,
 114
 Language-specific processing, 24,
 480
 Larynx, 29
 Last resort strategy, 339
- Late bilingualism, 154
 Late closure, 298, 307–308
 Late-syntax theories, 143
 Latent semantic analysis (LSA),
 356–357, **495**
 Laura, 81–82
 Layering, 453, 456
 LB, 235
 Learnability theory, 116
 Learning difficulties, 81–82
 Learning rule, 485
 Learning theory, 106–108, 116
 Lemmas, 412, **495**
 selection, 412, 418
 Lesion studies, 14–16, 68, 73
 Less-is-more-theory, 79, 118
 Levels of processing, 11,
 19–20
 Lexemes, 412, **495**
 Lexical, **495**
 Lexical access, 167–168, 265,
 495
 Lexical ambiguity, 199–208
 Lexical bias, 423
 Lexical boost, 406
 Lexical-category ambiguity, 311–
 313
 Lexical causatives, 333–334
 Lexical decision, 170, 173, 176,
 178, 181–182, 183–184, 186,
 188, 480
 Lexical development, 125–136,
 144–145
 Lexical entrainment, 457
 Lexical guidance, 300
 Lexical hermits, 211
 Lexical identification shift, 262–
 263
 Lexical instance models, 192
 Lexical neighborhoods, 272
 Lexical principles, 127–129
 Lexical route, 211
 Lexical selection, 264–265, 412
 Lexical-semantic anomia, 441–
 442
 Lexicalization, 398, 412–428,
 495
 feedback, 423–424, 427–428
 interaction, 420–424
 interactive activation model,
 424–425
 stages, 412–420
 time course, 420–423
 Lexicons, 7, 71, 322, **495**
 bilingual, 155–157
 number of, 465–471
 Lexigrams, 61, 62, 64, 65
- LIC (Low Interactional Content),
 365
 Limbus tracking, 168
 Linear-bounded automata, 44
 Linguistic determinism,
 89–90
 Linguistic feedback hypothesis,
 109
 Linguistic relativism, 90
 Linguistic universals, 112–114
 Linguistics, 4, 10
 Linking rules, 116, 136
 Lip-reading, 281, 462
 Liquids, 33
 Listening for mispronunciations
 task, 270–271
 Literacy, 168, 243–244
 Literal language, 340
 Living-nonliving dissociation,
 346–350
 Local interaction, 421
 Localization of function, 14, 15
 Location, 288, **495**
 Locational coherence, 362
 Locutionary force, 454, 455
 Logical inferences, 369
 Logistic function, 488
 Logogen model, 195–196, 198
 Logographic languages, 210
 Longitudinal studies, 105
 Look-and-say method, 247, 248
 Loulis (chimp), 60
 Low bigram frequency, 214
 LT, 473
 LW, 223
- Macroplanning, 398, 432, 435
 Macrorules, 386
 Macrostructure, 386
 “Magic moment”, 167
 Main clause, **495**
 Malapropisms, 413, **495**
 Mandarin, 97
 Manner of articulation, 33, **495**
 Mapping hypothesis, 316–317
 Mapping problem, 127, 129
 Marking, 409
 Masking, 172–173
 Mass nouns, 130
 Master file, 193
 Maturation, 72, 111, 112, **495**
 Maturation state hypothesis,
 75–76, 79
 MC, 222
 McGurk effect, 462
 MD, 346

- Mean length of utterance (MLU), 145
- Meaning
frequency, 203–205
postulates, 333
and sound, 216–217, 218–219
- Meaning-first view, 137
- Meaning through syntax (MTS), 311
- Mediated priming, 182, 190, 420–421, **495**
- MEG (magnetoencephalography), 17
- Melanie-Jane, 235–236, 252
- Memory
for color, 94–96
inferences, 368–373
text, 363–368
see also Episodic memory; Semantic memory; Short-term memory; Working memory
- Memory organization packets (MOPs), 383–384
- Mental dictionary, *see* Lexicon
- Mental encyclopedia, 322
- Mental grammar, 71
- Mental models, 384–386
- MERGE, 280
- Message-level processes, 397
- Meta-analysis, **495**
- Metalinguistic knowledge, 243
- Metaphor, 340–342, **496**
- Metrical segmentation strategy, 259
- MH, 185, 469
- Michelangelo, 347
- Microplanning, 398, 432, 435
- Microstructure, 386
- Milliseconds, **496**
- Mini-theories, 338, 339
- Minimal attachment, 298
- Minimal pairs, 29, **496**
- Minimalist hypothesis, 370
- Minimalist program, 35, 42–43
- Mirror neurons, 54
- Misreading repetition blindness, 190
- Mixed substitutions, 423
- MK, 442, 468
- ML, 222
- Modality-specific anomia, 344
- Modality-specific content hypothesis, 345
- Modality-specific context hypothesis, 345
- Modality-specific format hypothesis, 345
- Model-theoretic semantics, 325
- Models, 4, 23
- Modifiers, 42, **496**
- Modularity, 19, 20–22, 297, 301, 463–465, 480, **496**
- Modules, 21
- Mohawk, 45
- Monitor hypothesis, 160
- Monitor model, 159–160
- Monkeys, 55, 66–67, 122
- Monologues, 80
- Monosyllabic, 33, **496**
- Morphemes, 5, 6, **496**
stranding, 403
- Morphing, 409
- Morphology, 5, 191–192, 410–412, **496**
- MOSAIC, 139
- Motherese, *see* Child-directed speech (CDS)
- Motor theory, 267–268
- Movement, 97
- Moving masks, 169
- MP, 221, 234, 344
- MRI (magnetic resonance imaging), 17
- MS, 473
- Multiple-access model, 200
- Multiple components, 476
- Multiple-levels model, 228
- Multiple meanings, 199
- Multiple senses, 199
- Multiple-stores hypothesis, 344
- Mutual-exclusivity assumption, 128
- MV, 222
- N-statistic, 175–176
- N3C principle, 129
- N400, 17, 189, 272, 301
- Naming errors, 401
- Naming tasks, 170, 173, 176, 178, 181–182, 183–184, 186, 188
- Nasal sounds, 33
- Nativism, 79, 117, **496**
- Natural kind, 325, **496**
- Natural order in acquisition hypothesis, 159–160
- Natural selection, 52
- Nature–nurture debate, 105
- Navaho, 91
- NC, 443, 444
- Negative evidence, 83, 107–108, 116
- Negative sentences, 40
- Neighborhood density, 272
- Neighborhood effects, 175–176, 184, 213–215, 272
- Neighborhood size, 176
- Neologisms, 8, 439–440, **496**
- Network architecture, 485
- Neural networks, 13, 485
- Neuroimaging, 16–19, 68, 146
- Neuromodulatory systems, 343
- Neuropsychology, 4, 14, 16
cognitive neuropsychology, 14–16, 480, 481, 482
connectionist neuropsychology, 481, 482
discourse processing, 390–392
dissociation, 22
lexicons, 467–471
reading, 220–227
semantics, 342–353
speech production, 435–446
speech recognition, 281–282
text processing, 390–392
writing, 448
- Neuroscience, 482
cognitive neuroscience, 14
parsing, 315–319
- New nodes, 296
- NI, 224
- Nim Chimpsky, 61, 66
- Nodes, 38
- Non-associative semantic priming, 185, 186
- Noncognate translation equivalents, 162
- Non-fluent aphasia, 436
- Non-linguistic ambiguity, 460
- Non-nutritive sucking, 119
- Non-plan-internal errors, 404
- Non-semantic reading, 225, 228, 239
- Non-structural context, 266
- Non-terminal elements, 36
- Nonwords, 176, 211, **496**
processing, 212–214
- Noun-bar units, 42
- Noun phrases, 37, **496**
- Nouns, 36, 130, **496**
acquisition, 126, 127–128, 135
combinatorial history, 339–340
- Novel name–nameless category principle, 129
- Nucleus, 33

- Number, 405, **496**
 Number-agreement errors, 407–409
 Number systems, 93–94
- O, Dr., 469
 Object, 37, 38, 83, **496**
 Object permanence, 79, 80–81, 85
 Object recognition, 184–185, 352–353
 Obligatory decomposition hypothesis, 191
 Obligatory transformations, 292
 Off-line, **496**
 On-line, **496**
 Onomatopoeia, 51–52
 Onset, 33, **496**
 Ontogenetic roots, 88
 Open-class words, *see* Content words
 Open discourse roles, 371–372
 Open words, 140
 Optic aphasia, 344, 345
 Optimality Theory, 43
 Optional transformations, 292
 Ordered-access model, 200
 Orthographic error scores, 231
 Orthographic priming, 177
 Orthographic relatedness, 177
 Orthography/orthographic, 7, **496**
 Orthotactic, **496**
 Orton-Gillingham-Stillman multisensory method, 254
 Ostensive model, 127, **496**
 OUCH, 345
 Output simplification, 125
 Over-extensions, 131–134, **496**
 Over-regularization errors, 220–221
 Overlap of processing, 21, 420, 480
- P300, 17
 P600, 301
 Palatal sounds, 32
 Parafovea, 169, 297
 Paralexia, **496**
 Parallel autonomous models, 289
 Parallel coding, 212
 Parallel distributed processing (PDP), 13, 485
 Parallel function, 374
 Parallel transmission, 259
 Parameters/parameter setting, 41, 42, 111–112, **496**
 Paraphasias, 439–440, 443, **496**
 Parapraxes, 399
 Parkinson's disease, 69–70, 146
 Parrots, 56–57
 PARRY, 12
 Parsing, 287–320, **496**
 models, 298–299, 308–311
 neuroscience, 315–319
 psycholinguistics, 294–298
 units, 294–295
 working memory, 474–476
 Partial activation hypothesis, 416–417
 Partial alphabetic reading phase, 242
 Participles, 290, **496**
 Passives, 38, 149, 293
 Passivization, 40
 Past tense acquisition, 145–148
 Patients, 136, **496**
 Pattern masking, 172, 176
 Pauses, 432–435
 PC, 346
 Perception without awareness, 172–173
 Perceptual code, 343
 Performance, 34–35, **496**
 Performative verbs, 454
 Peripheral dyslexias, 220
 Periphery, 169
 Perlocutionary force, 454, 455
 PET (positron emission tomography), 18, 146, 350, 415
 Pheromones, 54
 “Phoenicians”, 251
 Phoenix (dolphin), 58
 Phonemes, 28–29, **496**
 blending, 222
 constancy, 243
 monitoring, 200–201, 261–262
 restoration, 263–264
 Phones, 28–29
 Phonetic encoding, 430
 Phonetics, 5, 28, 29, **496**
 Phonic mediation theory, 448
 Phonic method, 247–248
 Phonological anomia, 442
 Phonological attractors, 427
 Phonological awareness, 243–245, **497**
 Phonological blocking, 417
 Phonological buffers, 472–474
 Phonological deficit model, 254
 Phonological development, 75, 76, 120–125, 244
 Phonological dysgraphia, 448
 Phonological dyslexia, 221–223, 228, 235–236, **497**
 Phonological encoding, 398, 412, 415, 428–432
 Phonological error scores, 231
 Phonological facilitation, 404
 Phonological form selection, 412
 Phonological impairment hypothesis, 235–236
 Phonological input store, 468
 Phonological loop, 471, 472, 474, 475
 Phonological mediation, 212, 216–217
 Phonological output store, 468
 Phonological priming, 217
 Phonological recoding, 211, 217
 Phonological transparency, 192
 Phonological word, 431
 Phonology, 5, 28, 29, **497**
 Phonotactic cues, 121–122, **497**
 Phrase marker/structure, 39
 Phrase-structure grammar, 35–40
 Phrase-structure rules, 35–36, 39
 Phrases, 37, **497**
 Physical modularity, 22
 Picture naming, 184
 in dementia, 351–353
 Picture-word interference, 157, 414–415
 Pidgins, 114, **497**
 Pigeons, 65
 Piraha, 67, 91
 Pitch, 33, 34
 Pivot grammar, 140
 Pivot words, 140
 Place of articulation, 31, 32, **497**
 Planning units, 433
 Planum temporale, 250, 251
 PMSP model, 232–233
 Polysemous words, 199, **497**
 Possible-word constraint, 259
 Post-access effects, 181
 Postalveolar sounds, 32
 Postlexical code, 261–262
 Poverty of the stimulus, 108, 115
 PQ4R method, 390
 PR, 448
 Pragmatic inferences, 369
 Pragmatics, 5, 453–462, **497**
 Pre-alphabetic phase, 242
 Pre-intellectual stage, 88
 Pre-speech, 84
 Predicate, 37, **497**
 Predicate calculus, 323
 Predictive validity, 180

- Preferential looking technique, 104
 Prefixes, 403, **497**
 Prelexical code, 261–262
 Preliminary phrase packager (PPP), 297–298
 Premackese, 60–61
 Preoperational stage, 79
 Prepositional phrases, 309, **497**
 Prepositions, 36, **497**
 Presupposition, 372
 Primes, 171, 177
 Priming, 13–14, 171, 177, 459, **497**
 PRIMIR model, 122–123
 Principle of Economy, 42
 Principles and parameters theory, 41–42
 Print-to-sound conversion, 222
 Prior knowledge, 366–368
 Privileged information, 457
 Pro-drop parameter, 111
 Probabilistic feature model, 330–331
 Probabilistic models, 24
 Processing in cascade, 420
 Processing modularity, 22, 297, 301
 Production systems, 379
 Progressive aphasia, 444
 Pronounceable nonwords, 211
 Pronouns, 36, 376, **497**
 Pronunciation, 29
 Pronunciation neighborhoods, 215
 Property, 339
 inheritance, 326
 Proportion effect, 180
 Propositional network models, 379–380
 Propositions, 379, **497**
 Prosody, 120, 122, **497**
 Prosopagnosia, 185
 Proto-Indo-European, 7
 Protolanguage, 52, 67
 Prototype hypothesis, 133–134
 Prototypes, 134, 335–337, **497**
 Prototypicality effects, 134, 327
 PS, 217, 348, 349
 Pseudohomophone effect, 212–213, **497**
 Pseudowords, 176, 211, **497**
 PSMP model, 234–235
 Psycholinguistic grain size theory, 246
 Psycholinguistics, 4, 9–14, 19–24, 294–298, 480–481, 483–484, **497**
 Pure definitional negatives (PDNs), 333
 Pure word deafness, 281, 468
 Purkinje system, 168
 Push-down automata, 44
 PW, 427, 445, 469
 Pyramids and palm trees test, 441
 Queen's English, 29
 Ramu, 77
 Random noise masking, 172
 Rank hypothesis, 194
 Rationalism, 105, 106
 RB, 444
 RC, 350
 RE, 252, 475
 Reaction times, 11, 13, 170–171, 180
 Readability, 387–388
 Reading, 209–240
 eye movements, 168–170, 188–189, 216, 219
 improving, 254–255, 389–390
 learning, 241–248
 models, 211–212, 227–239
 neuropsychology, 220–227
 normal processes, 212–220
 silently, 217–218
 span, 389, 471
 speed, 219
 teaching methods, 247–248
 Reanalysis, 309
 Received Pronunciation (RP), 29, 33
 Recency, 378
 Recent-filler strategy, 314
 Receptive, **497**
 Receptive aphasia, 391
 Recipient, 287, 288
 Recognition point, 265, 269
 Recurrent networks, 278–279, 429, 489, **497**
 Recursion, 39, 67
 Reduced relative clauses, 290–291, **497**
 Reduplicated babble, 123
 Reference, 373–378, **497**
 Referent, **497**
 Referential coherence, 362
 Referential processing, 362
 Referential theory, 303–304, 324–325
 Refractoriness, 343
 Refractory period, 343, **497**
 Regressions, 170, 216
 Regular words, 211, 214, 215
 Reinforcement, 10
 Relatedness effect, 327
 Relative clauses, 39, 290, **497**
 Relative pronouns, 39
 Relativism, 22
 Reordered access model, 203–204, 205
 Repair, 401
 Repetition, 80
 Repetition blindness, 162, 190
 Repetition disorder, 472
 Repetition priming, 155–156, 176, 414, **497**
 Representational modularity, 301
 Reproduction conduction aphasia, 446
 Requests, 341–342
 Resolution, 372, 374
 Response, 10
 Response bias, 183
 Restricted search hypothesis, 377
 Restructured knowledge, 159
 Resyllabification, 431–432
 Reverse flow of information, 420, 423
 Reversible passives, 293
 Rewrite rules, 36, 40
 RG, 448
 Rhyme, 242
 Rhythm, 34
 RIA model, 428
 Rich interpretation, 140–141
 Rico (dog), 57
 Right association, 296
 Right-hemisphere hypothesis, 224–225
 Rimes, 33, 242–243, **497**
 RM, 346
 Romance languages, 7
 Rote repetition, 161
 RSVP (rapid serial visual presentation) technique, 189–190
 Rules, 23, 24, 34, 140, 480
 Russian, 149
 Saccades, 169, 170, **498**
 Sapir-Whorf hypothesis, 9, 89–98
 Sarah (chimp), 58–59, 60–61, 63
 Sausage machine model, 297–298
 Scan-copier mechanism, 428–429, 440
 Scenes, 383
 Schema-based theories, 382–384
 Schemas, 336, 382, **498**
 Scripts, 382–384, **498**

- Second language acquisition, 75–76, 158–162
- Segmentation, 259–260, **498**
- Segmentation markers, 386
- Selection phase, 265
- Selection restrictions, 329
- Selective access model, 203
- Selective adaptation, 264
- Semantic, **498**
- Semantic assimilation theory, 137, 138
- Semantic bias, 301–302, 313
- Semantic bootstrapping, 136, 137
- Semantic categorization tasks, 171, 216–217
- Semantic complexity, 135
- Semantic context, 266–267
- Semantic dementia, 344, 350–353
- Semantic development, 126–136
- Semantic feature hypothesis, 133
- Semantic features, 328–335, **498**
- Semantic glue hypothesis, 235
- Semantic memory, 322, 343–346, 350–353, **498**
- Semantic microfeatures, 353–355
- Semantic networks, 325–328
- Semantic opacity, 192
- Semantic paraplexias, 223, 237
- Semantic-pragmatic disorder, 84, 391
- Semantic priming, 14, 177, 178–181, 185–187, 190, **498**
- Semantic primitives, 328–329
- Semantic processing, 362
- Semantic reactivation, 314
- Semantic roles, 38
- Semantic sets, 198
- Semantic transparency, 192
- Semantic verb class hypothesis, 142
- Semantics, 5, 321–360, **498**
 connectionism, 353–358
 neuropsychology, 342–353
- Semantics comes first, 310
- Sememes, 236
- Semi-vowels, 33
- Sense, 324
- Sensitive period hypothesis, 78–79
- Sensorimotor period, 79
- Sensory sets, 198
- Sensory–functional theory, 348–350
- Sentence, 27, 40, **498**
 analysis, **500**
 hierarchy, 36
 parsing, *see* Parsing
- structure supervisor, 297–298
- surface structure, 295–296
- verification, 325–326, 330–331, 332
- Separate-store models, 155–156
- Sequential bilingualism, 154, **498**
- Serbo-Croat, 98, 210, 226
- Serial autonomous models, 289
- Serial search model, 192, 193–195, 198
- Sex differences, 71–72
- Sexist language, 98
- Shadowing, 270
- Sherman (chimpanzee), 61, 66
- Shona, 94
- Short-term memory (STM), 471–476, **498**
 aphasia, 446
 impairments, 391–392
- SHORTLIST, 279–280
- SHRDLU, 12, 13
- Sign language, 59–60, 61, 68, 76, 86–87, 109, 112, 114, 124
- Signal detection, 263
- Signals, 54
- Silent reading, 217–218
- Similarity effects, 423–424
- Simile, 340
- Simultaneous bilingualism, 154, **498**
- Situation models, 384, 386–387
- SJD, 418
- SL, 36, 347
- SLIP, 401
- SM model, 230–232
- Social aspects, 76, 82–84, 129
- Social interactionism, 83, 84
- Songbirds, 72
- Sound
 contrasts, 124
 and meaning, 216–217, 218–219
 spectrograms, 27
 speech sounds, 27–34
- Source, **498**
- SP, 255
- Span, **498**
- Spanish, 7, 226, 246, 307
- Spatial information, 384–385
- Spatial terms, 96
- Specific language impairment (SLI), 114–115, 146, 391, **498**
- Speech acts, 454–456, **498**
- Speech apraxia, 435
- Speech dysfluencies, 432–435
- Speech errors, 399–404, 411, 413–414, 427
- Speech perception, 71, 257, 267
 in infancy, 120–123
- Speech production, 29, 69, 397–450
 hesitations, 432–435
 lexicalization, 412–428
 models, 401–404, 425, 426, 429
 neuropsychology, 435–446
 phonological encoding, 428–432
 syntactic planning, 404–412
- Speech recognition, 257–283
 context, 258–259, 261, 262–264, 266–267, 271–272
 models, 267–281
 neuropsychology, 281–282
 speed, 258, 261
 time course, 264–265
- Speech sounds, 27–34
- Speed reading, 219
- Speed–accuracy trade-offs, 170
- Spelling, 243, 248–249
- Spelling-to-sound correspondence, 211
- Spoonerisms, 399, **498**
- Spreading activation, 13, 190, 327–328, 379, 425, 485
- ST, 473
- Stage 1 speech, 149
- Standard theory, 41
- Stem, **498**
- Stimulus, 10
 degradation, 172
- Stimulus-onset asynchrony (SOA), 171, **498**
- STM conduction aphasia, 446
- Stochastic output, 275, **498**
- Stops, 33
- Stories, 361, 362–363
- Story grammars, 380–381
- Stranding, 400, 403
- Stress, 33–34, 120
- Stress-based segmentation, 259–260
- Stressed-timed languages, 34
- Stroke, **498**
- Stroop task, 178
- Structural ambiguity, 289–291, 298–313
- Structural context, 266
- Structural priming, 143, 405–406
- Structuralism, 10
- Subcategorization frame, 287
- Subject, 37, 38, 287, **498**

- Sublexical route, 211, **498**
 Subliminal perception, 171, 172
 Submersion method, 159
 Subordinate clauses, **498**
 Subordinates, 323
 Substantive universals, 112
 Subtraction method, 19
 Sucking habituation, 74, 104, **498**
 Suffixes, 403, **498**
 Summation model, 229
 Super-additive impairment, 345–346
 Superordinates, 323
 Suppression, 389, **498**
 Suprasegmental features, 33
 Surface dysgraphia, 448
 Surface dyslexia, 220–221, 228, 233–234, **498**
 Surface structure (s-structure), 41
 Syllabary, 430
 Syllabic scripts, 210
 Syllabification, 429–430
 Syllable-based segmentation, 260
 Syllable-timed languages, 34
 Syllables, 6, 33, **498**
 monitoring tasks, 260
 onset and rime, 33
 role in phonological encoding, 431
 Symbols, 5
 Syndrome, 68, **499**
 Synecdoche, 340
 Synonymy, 321
 Syntactic bootstrapping, 130
 Syntactic category
 ambiguity, 311–313
 learning, 136–139
 Syntactic comprehension, 149
 Syntactic development, 75–76, 78, 136–149
 Syntactic persistence, 405–406
 Syntactic planning, 398, 401–402, 404–412
 chunking/incremental, 409–410
 Syntactic priming, 364, 405–407
 Syntactic processing, 288, 299–308
 Syntactic roles, 136, **499**, **500**
 Syntactically ambiguous sentences, 38
 Syntax, 5, 27, 34–43, 56, 288, **499**
 autonomy of, 292, 293
 Synthetic phonics, 248
 Tabula rasa, 105
 Tachistoscopes, 171, **499**
 Targets, 171, 177
 Taxonomic constraint, 128
 TB, 475
 Telegraphic speech, 103, 136, 140, **499**
 Templates, 12, 258, 267
 matching, 267
 Temporal coherence, 362
 Temporal discreteness, 420
 Temporal gyrus, 69, 71
 Tense, 405, **499**
 Terminal elements, 36
 Text, 361
 memory for, 363–368
 Text processing
 models, 378–388
 neuropsychology, 390–392
 Thai, 28
 Thematic organization points (TOPs), 383
 Thematic relations, 339
 Thematic roles, 136, 287, 288, **499**, **500**
 Theme, 38, 287, 288, **499**
 Theory, 4
 Theory of mind, 82
 Theory theories, 338, 339
 Theta role, *see* Thematic roles
 Thought, 87–98, 288
 Time, 97
 Tip-of-the-tongue (TOT), 416–418, 433, **499**
 TMS (transcranial magnetic stimulation), 18
 TOB, 343
 Tone languages, 34
 Tongue-twisters, 217, 218
 Top-down processing, 21, 296, 479–480, **499**
 TRACE, 273–278, 280, 485
 Trace-deletion hypothesis, 316
 Traces, 42, 313–315
 Transcortical aphasia, 446, 473, **499**
 Transformation, 40, 45, 292, **499**
 Transformational grammar, 10, 40, 292–294, **499**
 Transition relevance place, 458
 Transitive verbs, 37, **499**
 Translating between languages, 156–157
 Tree diagrams, 38–39
 Trevor, 249
 Triangle model, 230–232, 235, 238
 Truth-theoretic semantics, 325
 Truth value, 379
 TU, 349
 Tuning hypothesis, 308
 Turing machines, 44, 45
 Turn-taking, 83–84, 458–459
 Twin studies, 117, 252
 Two-word grammars, 139–141
 TY, 226
 Type 1/Type 2 grammars, 44
 Type 3 language, 44, 116
 Type-B spelling disorder, 251
 Tzeltal, 96
 U-shaped development, 108, 141, 146, 147, 159
 Ultra-cognitive neuropsychology, 15
 Unaspirated, 28, **499**
 Unbalanced ambiguity, 203
 Unbounded dependencies, 314–315
 Under-extensions, 132–134
 Unfilled pauses, 432–433
 Unilingual, **499**
 Unimodal store hypothesis, 343
 Uniqueness point, 265, 269
 Unit, **499**
 Universal grammar, 111, 112–114, **499**
 Unrestricted-race model, 309, 310
 Unrestricted search hypothesis, 377
 Unvoiced consonants, 32, **499**
 Uralic languages, 7
 Variable-choice one-stage models, 309
 Variegated babble, 123
 Vegetative sounds, 103
 Velars, 32
 Velum, 32
 Verb-bar units, 42
 Verbal code, 343
 Verbal short-term memory, 474
 Verbs, 36, **499**
 acquisition, 142
 argument structure, 141–144, 287, **499**, **500**
 auxiliary, 40, **492**
 bias, 305
 causative, 144
 copulative, **493**
 dative, **493**
 direct object, 305
 ditransitive, 37
 inflections, 145–148
 intransitive, 37, 38, **495**

- islands, 142
- NP1/NP2, 375
- performative, 454
- sentence complement, 305
- transitive, 37, **499**
- Verification models, 198
- Vertical information, 424
- Viki (chimp), 58, 59
- Visual deficits, 250, 255
- Visual errors, 222, 234
- Visual impairment, 84–86, 87
- Visual word form area, 185
- Visual word recognition, 167–208
- Visual world, 295, 306–307, 405, 414, 460–462, 483
- Visuo-spatial sketchpad, 471
- Vocabulary
 - differentiation, 90–91
 - learning, 474
- Vocal cords, 32
- Vocal play, 103
- Vocal tract, 29, 59
- Voice onset time (VOT), 32, 260–261, 262, **499**
- Voiceless consonants, 32
- Voiceless glottal fricative, 32
- Voicing, 32, **499**
- Vowels, 29, 31, **499**
 - processing, 281
- Washoe (chimp), 59–60, 62–63
- WB, 221–222
- WEAVER++ model, 428, 429–430, 431
- Weights, 486
- Welsh, 93–94
- Wernicke-Geschwind model, 14, 69–71
- Wernicke's aphasia, 69, 391, 436–437, 445–446, **499**
- Wernicke's area, 14, 15, 68–69
- WH-words, 60
- Whale song, 55
- Whole-object assumption, 127–128
- Whole word method, 247, 248
- Wickelfeatures, 231, 232–233
- Wild Boy of Aveyron, 77
- Williams syndrome, 82, 146
- WLP, 225
- WMA, 469
- Wolf children, 77
- Word, 6, **499**
 - association, 157
 - association norms, 185–186
 - class, 36, **499**
 - defined, 6
 - exchange errors, 403
 - first words, 125–131
 - frequency, 173–175, 265
 - length, 175
 - order, 113, 404
 - processing, 214–219
 - recognition, 167–208, 264–265
 - substitutions, 404, 413–414
 - superiority effect, 197
- Word meaning deafness, 281, 468
- Working memory, 22, 82, 387–388, 389, 391, 471–472, 474–476, **499**
- Writing, 209–211, 447–448
- WT, 225
- X-bar syntax, 42
- X-rays, 16, 17
- Yerkish, 61, 63, 64
- YOT, 347–348