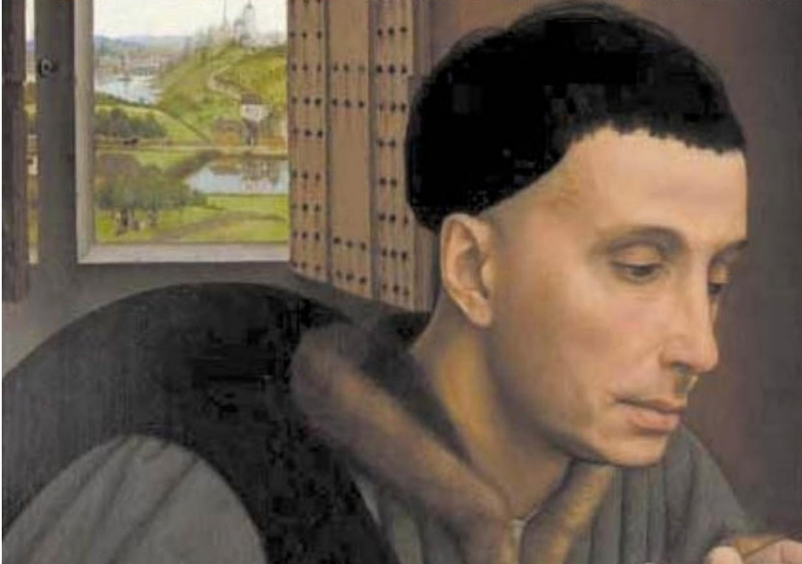
Handbooks of Developmental Psychology

THE SCIENCE OF READING
A Handbook



Edited by Margaret J. Snowling & Charles Hulme

Blackwell Publishing

Contents

20 *Learning to Read with a Hearing Impairment*

21 *Learning to Read with a Language Impairment*

PART VII *Teaching Reading*

26 *Teaching Children to Read: What Do We Know about How to Do It?*

27 *Recent Discoveries on Remedial Interventions for Children with Dyslexia*

Blackwell Handbooks of Developmental Psychology

This outstanding series of handbooks provides a cutting-edge overview of classic research, current research and future trends in developmental psychology.

- Each handbook draws together 25–30 newly commissioned chapters to provide a comprehensive overview of a subdiscipline of developmental psychology.
- The international team of contributors to each handbook has been specially chosen for its expertise and knowledge of each field.
- Each handbook is introduced and contextualized by leading figures in the field, lending coherence and authority to each volume.

The *Blackwell Handbooks of Developmental Psychology* will provide an invaluable overview for advanced students of developmental psychology and for researchers as an authoritative definition of their chosen field.

Published

Blackwell Handbook of Infant Development
Edited by Gavin Bremner and Alan Fogel

Blackwell Handbook of Childhood Social Development
Edited by Peter K. Smith and Craig H. Hart

Blackwell Handbook of Childhood Cognitive Development
Edited by Usha Goswami

Blackwell Handbook of Adolescence

Edited by Gerald R. Adams and Michael D. Berzonsky

The Science of Reading: A Handbook

Edited by Margaret J. Snowling and Charles Hulme

Forthcoming

Blackwell Handbook of Early Childhood Development
Edited by Kathleen McCartney and Deborah A. Phillips

The Science of Reading: A Handbook

Edited by

Margaret J. Snowling and Charles Hulme

Blackwell Publishing

BLACKWELL PUBLISHING
350 Main Street, Malden, MA 02148-5020, USA
9600 Garsington Road, Oxford OX4 2DQ, UK
550 Swanston Street, Carlton, Victoria 3053, Australia

First published 2005 by Blackwell Publishing Ltd

1 2005

Library of Congress Cataloging-in-Publication Data

The science of reading : a handbook / edited by Margaret J. Snowling and Charles Hulme. p. cm. — (Blackwell handbooks of developmental psychology)

Includes bibliographical references and indexes.

ISBN 13: 978-1-4051-1488-2 (alk. paper)

ISBN-10: 1-4051-1488-6 (alk. paper)

1. Reading. 2. Reading—Research. 3. Reading, Psychology of. I. Snowling, Margaret J. II. Hulme, Charles. III. Series.

LB1050.S365 2005

428.4—dc22

2005001421

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

For further information on

Blackwell Publishing, visit our website:

www.blackwellpublishing.com

Contributors

Judith A. Bowey
School of Psychology
University of Queensland
St Lucia
Queensland 4072
Australia
email: j.bowey@psy.uq.edu.au

Brian Byrne
School of Psychology
University of New England
Armidale NSW 2351
Australia
email: bbyrne@pobox.une.edu.au

Markéta Caravolas
Department of Psychology
University of Liverpool
Liverpool L69 7ZA
UK
email: M.C.Caravolas@liverpool.ac.uk

Max Coltheart
Macquarie Centre for Cognitive Science
Macquarie University
Sydney NSW 2109
Australia
email: max@maccs.mq.edu.au

Anna Maria Di Betta
Neurosciences Research Institute
School of Life and Health Sciences
Aston University
Birmingham B4 7ET
UK
email: a.m.dibetta@aston.ac.uk

Linnea C. Ehri

Graduate Center of the City University of New York
Program in Educational Psychology
CUNY Graduate Center
365 Fifth Ave.
New York, NY 10016
USA
email: LEhri@gc.cuny.edu

Jack M. Fletcher Center for Academic and Reading Skills
Department of Pediatrics
University of Texas Health Science Center at Houston
7000 Fannin UCT 2478
Houston TX 77030
USA
email: Jack.M.Fletcher@uth.tmc.edu

Ram Frost
Department of Psychology
The Hebrew University
Jerusalem 91905
Israel
email: frost@mscc.huji.ac.il

J. Richard Hanley
Department of Psychology
University of Essex
Wivenhoe Park
Colchester CO4 3SQ
UK
email: rhanley@essex.ac.uk

Charles Hulme
Department of Psychology
York University
York Y010 5DD
UK
email: ch1@york.ac.uk

Connie Juel
School of Education
Stanford University
485 Lasuen Mall
Stanford, CA 94305-3096
USA

email: cjuel@stanford.edu

Barbara J. Juhasz
Department of Psychology
University of Massachusetts
Amherst, MA 01003
USA
email: bjjuhasz@psych.umass.edu

Brett Kessler
Psychology Department
Washington University in St. Louis
Campus Box 1125
One Brookings Drive
St. Louis, MO 63130-4899
USA
email: bkessler@wustl.edu

Walter Kintsch
Department of Psychology
University of Colorado
Boulder, CO 80309-0344
USA
email: wkintsch@clipr.colorado.edu

Heidi Kloos
Center for Cognitive Sciences
Ohio State University
211 G Ohio Stadium East
1961 Tuttle Park Place
Columbus, OH 43210
email: Kloos.6@osu.edu

Régine Kolinsky
UNESCOG (CP 191)
Université libre de Bruxelles
50 Av. F. D. Roosevelt
B-1050 Brussels
Belgium
email: rkolins@ulb.ac.be

Matthew A. Lambon Ralph
Department of Psychology
University of Manchester

Manchester M13 9PL
UK
email: matt.lambon-ralph@man.ac.uk

Nicole Landi
Learning Research and Development Center
University of Pittsburgh
Pittsburgh, PA 15260
USA
email: nil3@pitt.edu

Jacqueline Leybaert
LAPSE
Université libre de Bruxelles
50 Av. F. D. Roosevelt
B-1050 Brussels
Belgium
email: leybaert@ulb.ac.be

Christopher J. Lonigan
Department of Psychology
Florida State University
One University Way
Tallahassee, FL 32306-1270
USA
email: lonigan@psy.fsu.edu

Stephen J. Lupker
Department of Psychology
University of Western Ontario
London
Ontario N6A 5C2
Canada
email: lupker@uwo.ca

Eamon McCrory
Department of Psychology
Institute of Psychiatry
De Crespigny Park
London SE5 8AF
UK
email: eamonmccrory@hotmail.com

José Morais

UNESCOG (CP 191)
Université libre de Bruxelles
50 Av. F. D. Roosevelt
B-1050 Brussels
Belgium
email: jmorais@ulb.ac.be

Kate Nation
Department of Experimental Psychology
University of Oxford
South Parks Road
Oxford OX1 3UD
UK
email: kate.nation@psy.ox.ac.uk

Jane Oakhill
Department of Psychology
University of Sussex
Falmer House
Brighton BN1 9RH
UK
email: J.Oakhill@sussex.ac.uk

Andrew Olson
Department of Psychology
University of Birmingham
Edgbaston
Birmingham B15 2TT
UK
email: a.l.o.olson@bham.ac.uk

Richard K. Olson
Department of Psychology
University of Colorado, UCB 345
Boulder, CO 80309
USA
email: rolson@psych.colorado.edu

Karalyn Patterson
MRC Cognition and Brain Sciences Unit
15 Chaucer Road
Cambridge CB2 2EF
UK
email: karalyn.patterson@mrccbu.cam.ac.uk

Bruce F. Pennington
Department of Psychology
University of Denver
2155 S. Race St.
Denver, CO 80210-4638
USA
email: bpenning@nova.psy.du.edu

Charles A. Perfetti
Learning Research and Development Center
University of Pittsburgh
Pittsburgh, PA 15260
USA
email: Perfetti@pitt.edu

Beth M. Phillips
Florida Center for Reading Research
City Centre Building Suite 7250
227 North Bronough St.
Tallahassee, FL 32301
USA
email: bphillips@fcrr.org

David C. Plaut
Departments of Psychology and Computer Science and Center for the Neural Basis of Cognition
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213-3890
USA
email: plaut@cmu.edu

Alexander Pollatsek
Department of Psychology
University of Massachusetts
Amherst, MA 01003
USA
email: pollatsek@psych.umass.edu

Cathy J. Price
Wellcome Department of Imaging Neuroscience
University College London
Institute of Neurology

12 Queen Square
London WC1N 3BG
UK
email: c.price@fil.ion.ucl.ac.uk

Katherine Rawson
Kent State University
Department of Psychology
P.O. Box 5190
Kent, OH 44242-0001
USA
email: krawson1@kent.edu

Keith Rayner
Department of Psychology
University of Massachusetts
Amherst, MA 01003
USA
email: rayner@psych.umass.edu

Cristina Romani
Department of Psychology
Aston University
Aston Triangle
Birmingham B4 7ET
UK
email: c.romani@aston.ac.uk

Philip H. K. Seymour
Department of Psychology
University of Dundee
Dundee DD1 4HN
UK
email: phks@edenfield65.freeserve.co.uk

Catherine E. Snow
Harvard Graduate School of Education
Larsen 3
Cambridge, MA 02138
USA
email: snowcat@gse.harvard.edu

Margaret J. Snowling
Department of Psychology

York University
York YO10 5DD
UK
email: m.snowling@psych.york.ac.uk

Joseph K. Torgesen
Florida Center for Reading Research at Florida State University
227 N. Bronough St., Suite 7250
Tallahassee, FL 32301
USA
email: torgesen@fcrr.org

Rebecca Treiman
Psychology Department
Washington University in St. Louis
Campus Box 1125
One Brookings Drive
St. Louis, MO 63130-4899
USA
email: rtreiman@wustl.edu

Guy C. Van Orden
Department of Psychology
Arizona State University
Tempe, AZ 85287-1104
USA
email: guy.van.orden@asu.edu

Frank R. Vellutino
Department of Psychology
State University of New York
1535 Western Avenue
Albany, NY 12203
USA
email: frv@csc.albany.edu

Preface

> "To completely analyse what we do when we read would almost be the acme of the psychologist's achievements, for it would be to describe very many of the most intricate workings of the human mind"
>
> (Huey, 1968).

The science of reading is mature and healthy as the contributions to this volume make clear. Together they provide an assessment of how far we have come in meeting the challenge laid down by Huey more than a century ago. Different chapters illustrate how some old issues remain alive, how new questions have been raised and how some problems have been solved. Many of the issues discussed here would undoubtedly have been familiar to Huey. Discussions of how skilled readers recognize printed words rapidly, of how eye movements in reading are controlled, the factors limiting reading comprehension, and arguments about how best to teach reading, all featured prominently in early studies of reading. These are important topics and ones that remain current, as several chapters in this book attest. There is little doubt that the technical advances made in many of these areas would be a source of pleasure to Huey and his contemporaries in the field of reading research. On the other hand, a number of issues dealt with in this book would probably have seemed totally foreign to people in the field of reading a century ago. For example, studies imaging the brain while it reads, studies examining the molecular genetics of reading disorders, and computational models of different aspects of the reading process would have seemed like science fiction a hundred years ago.

This Handbook provides a state-of-the-art overview of scientific studies of reading. The book is divided into seven sections. Part I deals with word recognition processes and is concerned largely with theories developed in studies of fluent adult reading. Such theories have heavily influenced (and been influenced by) studies of reading development, which are dealt with in Part II. Efficient word recognition processes are necessary, but not sufficient, for reading comprehension (Gough & Tunmer, 1986) and the chapters in Part III go beyond single word processing to consider reading comprehension processes in both adults and children, with an emphasis on the problems that may be encountered in children learning to comprehend what they read. Studies of reading and reading development have until recently been concerned only with reading English. Gough and Hillinger (1980) suggested that learning to read was an "unnatural act"; if that is true there is growing evidence that learning to read in English is a *particularly* unnatural act! Part IV of the book brings together work exploring how reading and reading development may differ across languages. This section highlights a number of issues and confronts the question of whether we can hope for a universal cognitive theory of reading and reading development – such a hope seems closer than some may have believed.

One justification for much research in psychology is that it helps us to understand, and in turn to prevent and to treat, disorders in psychological processes. The chapters in Part V look at our understanding of developmental and acquired disorders of reading and spelling. An important question here is the extent to which common forms of explanation may be valid for both acquired and developmental disorders. Part VI of the book examines the biological substrates of reading. It brings together work on brain imaging, which has revealed with new clarity the brain regions involved in different aspects of reading, with work on the genetic basis of dyslexia. The final section of the book, Part VII, examines how scientific studies of reading can contribute to improving the teaching of reading both in normally developing children and children with dyslexia.

We hope that the overviews of research presented here will be of value to psychologists and educationalists studying reading, their students, and to practitioners and others who want to find out about the current status of The Science of Reading.

Acknowledgments

We would like to thank Mark Seidenberg who played an invaluable role in helping to shape the form of this book in the early stages of its development.

We have learned a great deal from editing this book and would like to thank all our contributors for their excellent chapters, which made our task so easy and pleasurable.

Maggie Snowling and Charles Hulme

PART I

Word Recognition Processes in Reading

Editorial Part I

Word recognition is the foundation of reading; all other processes are dependent on it. If word recognition processes do not operate fluently and efficiently, reading will be at best highly inefficient. The study of word recognition processes is one of the oldest areas of research in the whole of experimental psychology (Cattell, 1886). The chapters in this section of the Handbook present an overview of current theories, methods, and findings in the study of word recognition processes in reading.

What do we mean by recognition here? Recognition involves accessing information stored in memory. In the case of visual word recognition this typically involves retrieving information about a word's spoken form and meaning from its printed form. The first two chapters, by Coltheart and Plaut, outline the two most influential theoretical frameworks for studies of visual word recognition.

Coltheart outlines the history and evolution of dual-route models of reading *aloud* (i.e., how the pronunciation of a printed word is generated). These dual-route models posit that there are two routes from print to speech: a lexical and nonlexical route. Broadly the lexical route involves looking up the pronunciation of a word stored in a lexicon or mental dictionary. In contrast, the nonlexical route involves translating the graphemes (letters or letter groups) into phonemes and assembling the pronunciation of a word from this sequence of phonemes. Such a process should work just as well for nonwords as for words, just so long as the word follows the spelling pattern of the language (a nonlexical reading of YACHT, will not yield the pronunciation for a kind of boat with a sail on it). This idea is embodied in an explicit computational model (the DRC model) that Coltheart describes in detail. It may be worth emphasizing that this highly influential model is a model of how adults read aloud; it is not concerned with how the knowledge allowing this to happen is acquired. A major focus of the model is how different disorders of reading aloud, which arise after brain damage in adults, can be accounted for.

Plaut gives an overview of a different class of models of reading aloud that employ connectionist architectures (models that learn to pronounce words by training associations between distributed representations of orthography and phonology). One particularly influential model of this type is the so-called triangle model (Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989). This model abandons the distinction between a lexical and nonlexical procedure for translating visual words into pronunciations; instead the same mechanism is used to convert words and nonwords into pronunciations, based on patterns of connections between orthographic inputs and phonological outputs. One other critical difference between the triangle model and the DRC model is that the triangle model explicitly embodies a learning procedure and thus can be considered a model of both adult reading and reading development. It is clear that these are very different conceptions of how the mind reads single words. Both approaches deal with a wide range of evidence. Arguably, the DRC model is more successful in dealing with the detailed form of reading impairments observed after brain damage in adults, while the ability to think about development and adult performance together in the triangle model is a considerable attraction. There is no doubt that differences between these models will be a source of intense interest in the coming years.

Lupker's chapter moves on to review a huge body of experimental evidence concerned with how adults recognize printed words. Many of these experiments investigate what is a remarkably rapid and accurate process in most adults, by measuring reaction time, or by impairing performance by using masking (preventing participants from seeing a word clearly by superimposing another stimulus immediately after the word has been presented). Any complete model of word recognition ultimately will have many phenomena from such experiments to explain. These include the fact that people perceive letters more efficiently when they are embedded in words, that high-frequency (i.e., more familiar) words are recognized easier than less familiar words, and that recognition of words is influenced by previously presented words (seeing a prior word that is related in form or meaning helps us to recognize a word that follows it). One conclusion that emerges powerfully from Lupker's review is the need for interactive models in which activation of orthographic and phonological information reciprocally influence each other. This is an issue that Van Orden and Kloos take up in detail, presenting a wealth of evidence that converges on the idea that there is intimate and perpetual interaction between representations of orthography and phonology (spelling and sound) during the process of recognizing a printed word.

Moving on from the recognition of isolated words, Rayner, Juhasz, and Pollatsek discuss eye movements in reading. Eye movements provide a fascinating window on how word recognition processes operate in the more natural context of reading continuous text. It appears that the pattern of eye movements in reading is heavily influenced by the cognitive processes subserving both word recognition and text comprehension. The majority of words in text are directly fixated (usually somewhere in the first half of the word). For readers of English the area of text processed during a fixation (the perceptual span) is about 3 or 4 letters to the left of fixation and some 14 or 15 letters to the right of fixation. This limit seems to be a basic one determined by acuity limitations, and useful information about letter identity is extracted only from a smaller area, perhaps 7 or 8 letters to the right of the fixation point. It appears that only short, frequent, or highly predictable words are identified prior to being fixated (so that they can be skipped). However, partial information (about a word's orthography and phonology but typically not its meaning) about the word following the fixation point often is extracted and combined with information subsequently extracted when the word is directly fixated. These studies are consistent with the view that the speed and efficiency of word recognition processes (as well as higher-level text-based processes) place fundamental constraints on how quickly even skilled readers read text.

Arguably the central question in the study of word recognition in reading is the role of phonology. All of the chapters in Part I address this issue explicitly. It appears that a consensus has been reached: phonological coding is central to word recognition, though opinions are divided on many details of how phonology is accessed and its possible importance in providing access to semantic information.

1

Modeling Reading: The Dual-Route Approach

Max Coltheart

Reading is information-processing: transforming print to speech, or print to meaning. Anyone who has successfully learned to read has acquired a mental information-processing system that can accomplish such transformations. If we are to understand reading, we will have to understand the nature of that system. What are its individual information-processing components? What are the pathways of communication between these components?

Most research on reading since 1970 has investigated reading aloud and so sought to learn about the parts of the reading system that are particularly involved in transforming print to speech. A broad theoretical consensus has been reached: whether theories are connectionist (e.g., Seidenberg & McClelland, 1989; Plaut, this volume) or nonconnectionist (e.g., Coltheart, Curtis, Atkins & Haller, 1993), it is agreed that within the reading system there are two different procedures accomplishing this transformation – there are dual routes from print to speech. (The distinction between connectionist and nonconnectionist theories of cognition is discussed later in this chapter.)

In the Beginning…

The dual-route conception of reading seems first to have been enunciated by de Saussure (1922; translated 1983, p. 34):

> there is also the question of reading. We read in two ways; the new or unknown word is scanned letter after letter, but a common or familiar word is taken in at a glance, without bothering about the individual letters: its visual shape functions like an ideogram.

However, it was not until the 1970s that this conception achieved wide currency. A clear and explicit expression of the dual-route idea was offered by

Forster and Chambers (1973):

> The pronunciation of a visually presented word involves assigning to a sequence of letters some kind of acoustic or articulatory coding. There are presumably two alternative ways in which this coding can be assigned. First, the pronunciation could be computed by application of a set of grapheme–phoneme rules, or letter-sound correspondence rules. This coding can be carried out independently of any consideration of the meaning or familiarity of the letter sequence, as in the pronunciation of previously unencountered sequences, such as flitch, mantiness and streep. Alternatively, the pronunciation may be determined by searching long-term memory for stored information about how to pronounce familiar letter sequences, obtaining the necessary information by a direct dictionary look-up, instead of rule application. Obviously, this procedure would work only for familiar words. (Forster & Chambers, 1973, p. 627)

> Subjects always begin computing pronunciations from scratch at the same time as they begin lexical search. Whichever process is completed first controls the output generated. (Forster & Chambers, 1973, p. 632)

In the same year, Marshall and Newcombe (1973) advanced a similar idea within a box-and arrow diagram. The text of their paper indicates that one of the routes in that model consists of reading “via putative grapheme–phoneme correspondence rules” (Marshall & Newcombe, 1973, p. 191). Since the other route in the model they proposed involves reading via semantics, and is thus available only for familiar words, their conception would seem to have been exactly the same as that of Forster and Chambers (1973).

This idea spread rapidly:

> We can… distinguish between an orthographic mechanism, which makes use of such general and productive relationships between letter patterns and sounds as exist, and a lexical mechanism, which relies instead upon specific knowledge of pronunciations of particular words or morphemes, that is, a lexicon of pronunciations (if not meanings as well). (Baron & Strawson, 1976, p. 386)

> It seems that both of the mechanisms we have suggested, the orthographic and lexical mechanisms, are used for pronouncing printed words. (Baron & Strawson, 1976, p. 391)

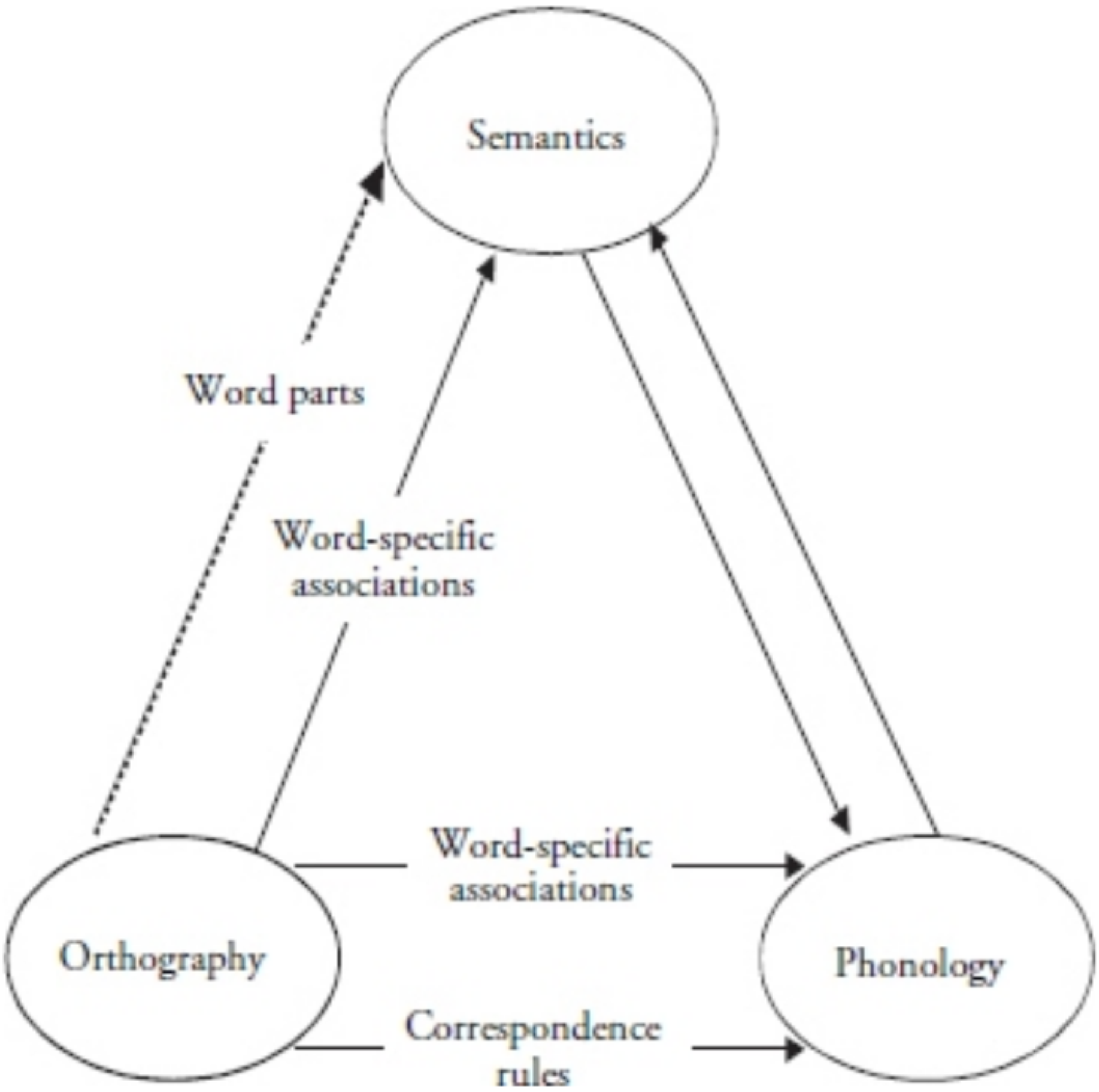
> Naming can be accomplished either by orthographic-phonemic translation, or by reference to the internal lexicon. (Frederiksen & Kroll, 1976, p. 378)

In these first explications of the dual route idea, a contrast was typically drawn between words (which can be read by the lexical route) and nonwords (which cannot, and so require the nonlexical route). Baron and Strawson (1976) were the first to see that, within the context of dual-route models, this is not quite the right contrast to be making (at least for English):

> The main idea behind Experiment 1 was to compare the times taken to read three different kinds of stimuli: (a) regular words, which follow the “rules” of

English orthography, (b) exception words, which break these rules, and (c) nonsense words, which can only be pronounced by the rules, since they are not words. (Baron & Strawson, 1976, p. 387)

Figure 1.1 An architecture of the reading system (redrawn from Baron, 1977).



Baron (1977) was the first to express these ideas in a completely explicit box-and-arrow model of reading, which is shown in figure 1.1. This model has some remarkably modern features: for example, it has a lexical-nonsemantic route for reading aloud (a route that is available only for words yet does not proceed via the semantic system) and it envisages the possibility of a route from orthography to semantics that uses word parts (Baron had in mind prefixes and suffixes here) as well as one that uses whole words.

Even more importantly, the diagram in figure 1.1 involves two different uses of the dual-route conception. The work previously cited in this chapter all concerned a dualroute account of reading aloud; but Baron's model also offered a dual-route account of reading comprehension:

> we may get from print to meaning either directly – as when we use pictures or maps, and possibly when we read a sentence like *I saw the son* – or indirectly, through sound, as when we first read a word we have only heard before. (Baron, 1977, p. 176)

> Two different strategies are available to readers of English for identifying a printed word. The phonemic strategy involves first translating the word into a full

> phonemic (auditory and/or articulatory) representation, and then using this representation to retrieve the meaning of the word. This second step relies on the same knowledge used in identifying words in spoken language. This strategy must be used when we encounter for the first time a word we have heard but not seen. The visual strategy involves using the visual information itself (or possibly some derivative of it which is not formally equivalent to overt pronunciation) to retrieve the meaning. It must be used to distinguish homophones when the context is insufficient, for example, in the sentence, "Give me a pair (pear)." (Baron & McKillop, 1975, p. 91)

The dual-route theory of reading aloud and the dual-route theory of reading comprehension are logically independent: the correctness of one says nothing about the correctness of the other. Further discussion of these two dual-route theories may be found in Coltheart (2000). The present chapter considers just the dual-route approach to reading aloud.

A final point worth making re Baron's chapter has to do with the analogy he used to illustrate why two routes might be better than one (even when one is imperfect – the nonlexical route with irregular words, for example):

> A third – and to me most satisfying – explanation of the use of the indirect path… is that it is used in parallel with the direct path. If this is the case, we can expect it to be useful even if it is usually slower than the direct path in providing information about meaning. If we imagine the two paths as hoses that can be used to fill up a bucket with information about meaning, we can see that addition of a second hose can speed up filling the bucket even if it provides less water than the first. (Baron, 1977, p. 203)

An analogy commonly used to describe the relationship between the two routes in dual-route models has been the horse race: the lexical and nonlexical routes race, and whichever finishes first is responsible for output. But this analogy is wrong. In the reading aloud of irregular words, on those occasions where the nonlexical route wins, according to the horse race analogy the response will be wrong: it will be a regularization error. But what is typically seen in experiments on the regularity effect in reading aloud is that responses to irregular words are correct but slow. The horse race analogy cannot capture that typical result, whereas Baron's hose-and-bucket analogy can. The latter analogy is equally apt in the case of the dual-route model of reading comprehension.

"Lexical" and "Nonlexical" Reading Routes

This use of the terms "lexical" and "nonlexical" for referring to the two reading routes seems to have originated with Coltheart (1980). Reading via the lexical route involves looking up a word in a mental lexicon containing knowledge about the spellings and pronunciations of letter strings that are real words (and so are present in the lexicon); reading via the nonlexical route makes no reference to this lexicon, but instead involves making use of rules relating segments of orthography to segments of phonology. The quotation from de Saussure with which this chapter began suggested that the orthographic segments used by the nonlexical route are single letters, but, as discussed by Coltheart (1978), that cannot be right, since in most alphabetically written languages single phonemes are frequently represented by sequences of letters rather than single letters. Coltheart (1978) used the term "grapheme" to refer to any letter or letter sequence that represents a single phoneme, so that TH and IGH are the two graphemes of the two-phoneme word THIGH. He suggested that the rules used by the nonlexical reading route are, specifically, grapheme–phoneme correspondence rules such as TH → */θ/* and IGH → /ai/.

Phenomena Explained via the Dual-Route Model

This model was meant to explain data not only from normal reading, but also facts about disorders of reading, both acquired and developmental.

Reaction times in reading-aloud experiments are longer for irregular words than regular words, and the dual-route model attributed this to that fact that the two routes generate conflicting information at the phoneme level when a word is irregular, but not when a word is regular: resolution of that conflict takes time, and that is responsible for the regularity effect in speeded reading aloud. Frequency effects on reading aloud were explained by proposing that access to entries for high-frequency words in the mental lexicon was faster than access for low-frequency words. From that it follows, according to the dualroute model, that low-frequency words will show a larger regularity effect, since lexical processing will be relatively slow for such words and there will be more time for the conflicting information from the nonlexical route to affect reading; and this interaction of frequency with regularity was observed.

Suppose brain damage in a previously literate person selectively impaired the operation of the lexical route for reading aloud while leaving the nonlexical route intact. What would such a person's reading be like? Well, nonwords and regular words would still be read with normal accuracy because the nonlexical route can do this job; but irregular words will suffer, because for correct reading they require the lexical route. If it fails with an irregular word, then the response will just come from the nonlexical route, and so will be wrong: *island* will be read as "iz-land," *yacht* to rhyme with "matched," and *have* to rhyme with "cave." Exactly this pattern is seen in some people whose reading has been impaired by brain damage; it is called surface dyslexia, and two particularly clear cases are those reported by McCarthy and Warrington (1986) and Behrmann and Bub (1992). The occurrence of surface dyslexia is good evidence that the reading system contains lexical and nonlexical routes for reading aloud, since this reading disorder is exactly what would be expected if the lexical route is damaged and the nonlexical route is spared.

Suppose instead that brain damage in a previously literate person selectively impaired the operation of the nonlexical route for reading aloud while leaving the lexical route intact. What would such a person's reading be like? Well, irregular words and regular words would still be read with normal accuracy because the lexical route can do this job; but nonwords will suffer, because for correct reading they require the nonlexical route. Exactly this pattern – good reading of words with poor reading of nonwords – is seen in some people whose reading has been impaired by brain damage; it is called phonological dyslexia (see Coltheart, 1996, for a review of such studies). This too is good evidence for a dual-route conception of the reading system.

The reading disorders just discussed are called acquired dyslexias because they are acquired as a result of brain damage in people who were previously literate. The term "developmental dyslexia," in contrast, refers to people who have had difficulty in learning to read in the first place, and have never attained a normal level of reading skill. Just as brain damage can selectively affect the lexical or the nonlexical reading route, perhaps also learning these two routes is subject to such selective influence. This is so. There are children who are very poor for their age at reading irregular words but normal for their age at reading regular words (e.g., Castles & Coltheart, 1996); this is developmental surface dyslexia. And there are children who are very poor for their age at reading nonwords but normal for their age at reading regular words and irregular words (e.g., Stothard, Snowling, & Hulme, 1996); this is developmental phonological dyslexia. Since it appears that difficulties in learning just the lexical and or just the nonlexical route can be observed, these different patterns of developmental dyslexia are also good evidence for the dual-route model of reading.

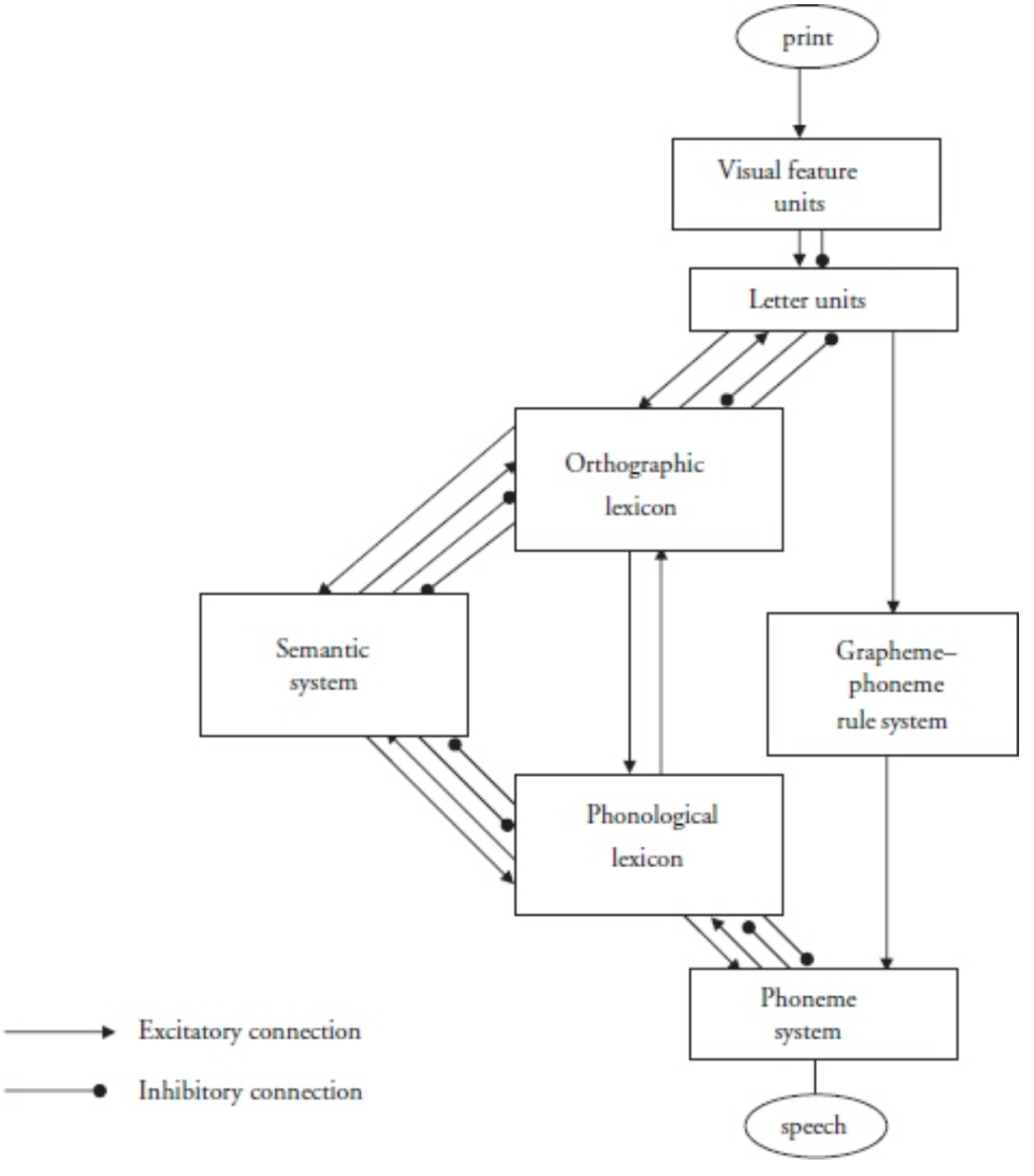
Computational Modeling of Reading

We have seen that the dual-route conception, applied both to reading aloud and to reading comprehension, was well established by the mid-1970s. A major next step in the study of reading was computational modeling.

A computational model of some form of cognitive processing is a computer program which not only executes that particular form of processing, but does so in a way that the modeler believes to be also the way in which human beings perform the cognitive task in question. Various virtues of computational modeling are generally acknowledged – for example, it allows the theorist to discover parts of a theory that are not explicit enough; inexplicit parts of a theory cannot be translated into computer instructions. Once that problem is solved and a program that can actually be executed has been written, the modeler can then determine how closely the behavior of the model corresponds to the behavior of humans. Do all the variables that influence the behavior of humans as they perform the relevant cognitive task also affect the behavior of the program, and in the same way? And do all the variables that influence the behavior of the program as it performs the relevant cognitive task also affect the behavior of humans, and in the same way? Provided that the answer to both questions is yes, studying the behavior of the computational model has demonstrated that the theory from which the model was generated is sufficient to explain what is so far known about how humans perform in the relevant cognitive domain. That does not mean that there could not be a different theory from which a different computational model could be generated which performed just as well. If that happens, the time has come for working out experiments about which the theories make different predictions – that is, whose outcomes in simulations by the two computational models are in conflict.

Of all cognitive domains, reading is the one in which computational modeling has been most intensively employed. This began with the interactive activation and competition (IAC) model of McClelland and Rumelhart (1981) and Rumelhart and McClelland (1982). This was a model just of visual word recognition, not concerned with semantics or phonology. The latter domains were introduced in the much more extensive computational model developed in a seminal paper by Seidenberg and McClelland (1989). One influence their paper had was to prompt the development of a computational version of the dual-route model: the DRC (“dual-route cascaded”) model (Coltheart et al., 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001).

Figure 1.2 The DRC model.



The Dual-Route Cascaded (DRC) Model

The DRC is a computational model that computes pronunciation from print via two procedures, a lexical procedure and a nonlexical procedure (see figure 1.2).

The lexical procedure involves accessing a representation in the model's orthographic lexicon of real words and from there activating the word's node in the model's phonological lexicon of real words, which in turn activates the word's phonemes at the phoneme level of the model. Nonwords cannot be correctly read by this procedure since they are not present in these lexicons, but that does not mean that the lexical route will simply not produce any phonological output when the input is a nonword. A nonword such as SARE can produce some activation of entries in the orthographic lexicon for words visually similar to it, such as CARE, SORE, or SANE; this in turn can activate the phonological lexicon and hence the phoneme level. Such lexically generated activation cannot produce the correct pronunciation for a nonword, but there is evidence that it does influence the reading aloud of nonwords. For example, a nonword like SARE which is similar to many entries in the orthographic lexicon will be read aloud with a shorter reaction time (RT) than a nonword like ZUCE which is similar to few (McCann & Besner, 1987).

The nonlexical procedure of the DRC model applies grapheme–phoneme correspondence rules to the input string to convert letters to phonemes. It does so in serial left-to-right fashion, initially considering just the first letter in the string, then the first two letters, then the first three letters, and so on, until it gets past the last letter in the input. It correctly converts nonwords from print to sound, and also regular words (those that obey its grapheme–phoneme correspondence rules). Irregular (exception) words are "regularized" by the nonlexical procedure – that is, their rule-based pronunciations, which will be incorrect.

Processing along the lexical route occurs as follows:

Cycle 0: set all the units for visual features that are actually present in the input string to 1; set all others to zero.

Cycle 1: every visual feature set to 1 contributes activation to all the letters in the letter units to which it is connected. The connections are inhibitory when the letter does not contain that feature, and so the activation contributed is negative; the connections are excitatory when the letter does contain that feature, and so the activation contributed is positive.

Cycle 2: what happens on Cycle 1 again happens here. In addition, every letter unit contributes activation to all the word units in the orthographic lexicon to which it is connected. The connections are inhibitory when the word does not contain that letter, and so the activation contributed from letter unit to word unit is negative; the connections are excitatory when the word does contain that letter, and so the activation contributed from letter unit to word unit is positive.

Cycle 3: everything that happens on Cycle 1 and Cycle 2 happens again here. In addition:

(a) Feedforward: each unit in the orthographic lexicon contributes activation to its corresponding unit in the phonological lexicon.

(b) Feedback: every word unit in the orthographic lexicon unit contributes activation back to all the letter units to which it is connected. The connections are inhibitory when the word does not contain that letter, and so the activation contributed from word unit to letter unit is negative; the connections are excitatory when the word does contain that letter, and so the activation contributed from word unit to letter unit is positive.

Cycle 4: everything that happens on Cycles 1, 2, and 3 happens again here. In addition:

(a) Feedforward: every unit in the phonological lexicon contributes activation to all the phoneme units to which it is connected. The connections are inhibitory when the word's pronunciation does not contain that phoneme, and so the activation contributed from word unit to phoneme unit is negative; the connections are excitatory when the word's pronunciation does contain that phoneme, and so the activation contributed from word unit to phoneme unit is positive.

(b) Feedback: every unit in the phonological lexicon contributes feedback activation to its corresponding unit in the orthographic lexicon.

Cycle 5: everything that happens on Cycles 1, 2, 3, and 4 happens again here. In addition: every phoneme unit contributes activation back to all the word units in the phonological lexicon to which it is connected. The connections are inhibitory when the word does not contain that phoneme, and so the activation contributed from phoneme unit to word unit is negative; the connections are excitatory when the word does contain that phoneme, and so the activation contributed from phoneme unit to word unit is positive.

And so it goes. As processing cycles progress, inhibitory and excitatory influences continue to flow upwards and downwards in the way described above until the reading-aloud response is ready. How is this readiness determined? As follows. In the description of processing cycles given above, the first cycle on which the phoneme system receives any activation is Cycle 4. At the end of cycle 4, some phoneme units will be activated, but extremely weakly. As processing continues, activation of some of the phoneme units will slowly rise. Quite often, early in processing, some of the phoneme units activated will be incorrect ones. But over time as phoneme activations continue to rise it is the correct phonemes that are the most activated. A reading response is considered to be ready when phonemes have reached a critical level of activation (set to.43 when the model is being used for simulating human reading aloud). The pronunciation generated by the model is taken to consist of the most highly activated phoneme within each of the eight sets of phoneme units (one set per position) that comprise the phoneme system. The processing cycle on which that state of affairs occurs is the DRC model's reading-aloud latency for the particular letter string that was input.

Processing along the nonlexical route does not begin to operate until cycle 10. Without this time lapse after the lexical route begins to operate, the model would have serious difficulty in reading aloud irregular words. When cycle 10 is reached, the nonlexical route translates the first letter of the string into its phoneme using the appropriate grapheme–phoneme rule, and contributes activation to the phoneme's unit in the phoneme system. This continues to occur for the next 16 processing cycles. The grapheme–phoneme conversion (GPC) system operates from left to right, so eventually will move on to consider the second letter in the string as well as the first. Every 17 cycles, the GPC system moves on to consider the next letter, translate it to a phoneme, and activate that phoneme in the phoneme system. So with the letter string DESK, the GPC system has no input until cycle 10, deals with just D until cycle 27, deals with just DE from cycle 28 to cycle 44, then DES until cycle 60, DESK until cycle 76 and so on.

Computations on the lexical and nonlexical route occur simultaneously – that is, information from the visual feature level is thought of as flowing simultaneously through the lexical and the nonlexical routes and converging on the phoneme system from these two sources. Whenever the input is an irregular word or a nonword, the two sources of activation conflict at the phoneme level. If the system is to produce correct pronunciations for irregular words and for nonwords, it will have to have a way of resolving these conflicts in favor of the correct pronunciation. Nevertheless, the model reads aloud irregular words and nonwords with high accuracy, so these conflicts are almost always resolved in a way that results in a correct pronunciation (via the interplay of inhibition and activation at various levels of the model). This depends on a judicious choice of values for the parameters of the model, such as the strengths of the inhibitory and the facilitatory connections between components of the model. If the lexical route is too strong relative to the nonlexical route, all words will be read correctly but there will be nonword reading errors. If the lexical route is too weak relative to the nonlexical route, all regular words and nonwords will be read correctly but there will be errors in reading irregular words. A delicate balance between the strengths of the two routes is needed if the model is to perform well with both nonwords and irregular words.

What the DRC Model Can Explain

One way in which Coltheart et al. (2001) evaluated the DRC model was to compare its reaction times to particular sets of stimuli to the reaction times of human readers when they are reading aloud the same stimuli. Do variables that affect human reading-aloud reaction times also affect DRC's reading-aloud reaction times? Many examples where this was so were reported by Coltheart et al. (2001). For both human readers and the DRC model:

(a) High-frequency words are read aloud faster than low-frequency words.

(b) Words are read aloud faster than nonwords.

(c) Regular words are read aloud faster than irregular words.

(d) The size of this regularity advantage is larger for low-frequency words than for high-frequency words.

(e) The later in an irregular word its irregular grapheme–phoneme correspondence is, the less the cost incurred by its irregularity. So CHEF (position 1 irregularity) is worse than SHOE (position 2 irregularity), which is worse than CROW (position 3 irregularity).

(f) Pseudohomophones (nonwords that are pronounced exactly like real English words, such as brane) are read aloud faster than non-pseudohomophonic nonwords (such as brene).

(g) Pseudohomophones derived from high-frequency words (e.g., hazz) are read aloud faster than pseudohomophones derived from low-frequency words (e.g., glew).

(h) The number of orthographic neighbors a non-pseudohomophonic nonword has (i.e., the number of words that differ from it by just one letter), the faster it is read aloud.

(i) The number of orthographic neighbors a pseudohomophone has does not influence how fast it is read aloud.

(j) The more letters in a nonword there are the slower it is read aloud; but number of letters has little or no effect on reading aloud for real words.

The DRC model was also used to simulate acquired dyslexias. Surface dyslexia was simulated by slowing down rate of access to the orthographic lexicon: this lesioned DRC made regularization errors with irregular words, more so when they were low in frequency, just as is seen in surface dyslexia, whereas its reading aloud of regular words and nonwords remained normal, as in the pure cases of surface dyslexia (Behrmann & Bub, 1992; McCarthy & Warrington, 1986). Phonological dyslexia was simulated by slowing down the operation of the nonlexical route: this lesioned DRC still read words correctly, but misread nonwords, especially if they were nonpseudohomophones, as in the case of phonological dyslexia.

Thus, the DRC model can explain an impressively large number of findings from studies of normal and disordered reading, far more than any other computational model of reading. Nevertheless, Coltheart et al. (2001) drew attention to a number of limitations of the current implementation of the DRC model: its procedure for performing the lexical decision task was crude, it was not applicable to the pronunciation of polysyllabic words or nonwords, it did not offer any account of one popular paradigm for studying reading (masked priming), the difference between word and nonword reading RTs by the model was probably implausibly large, the amount of variance of word reading RTs that the model could account for, though always significant, was disappointingly low, and the implemented model has nothing to say about semantics. A new version of the DRC model that will correct these and other shortcomings of the existing model is under development.

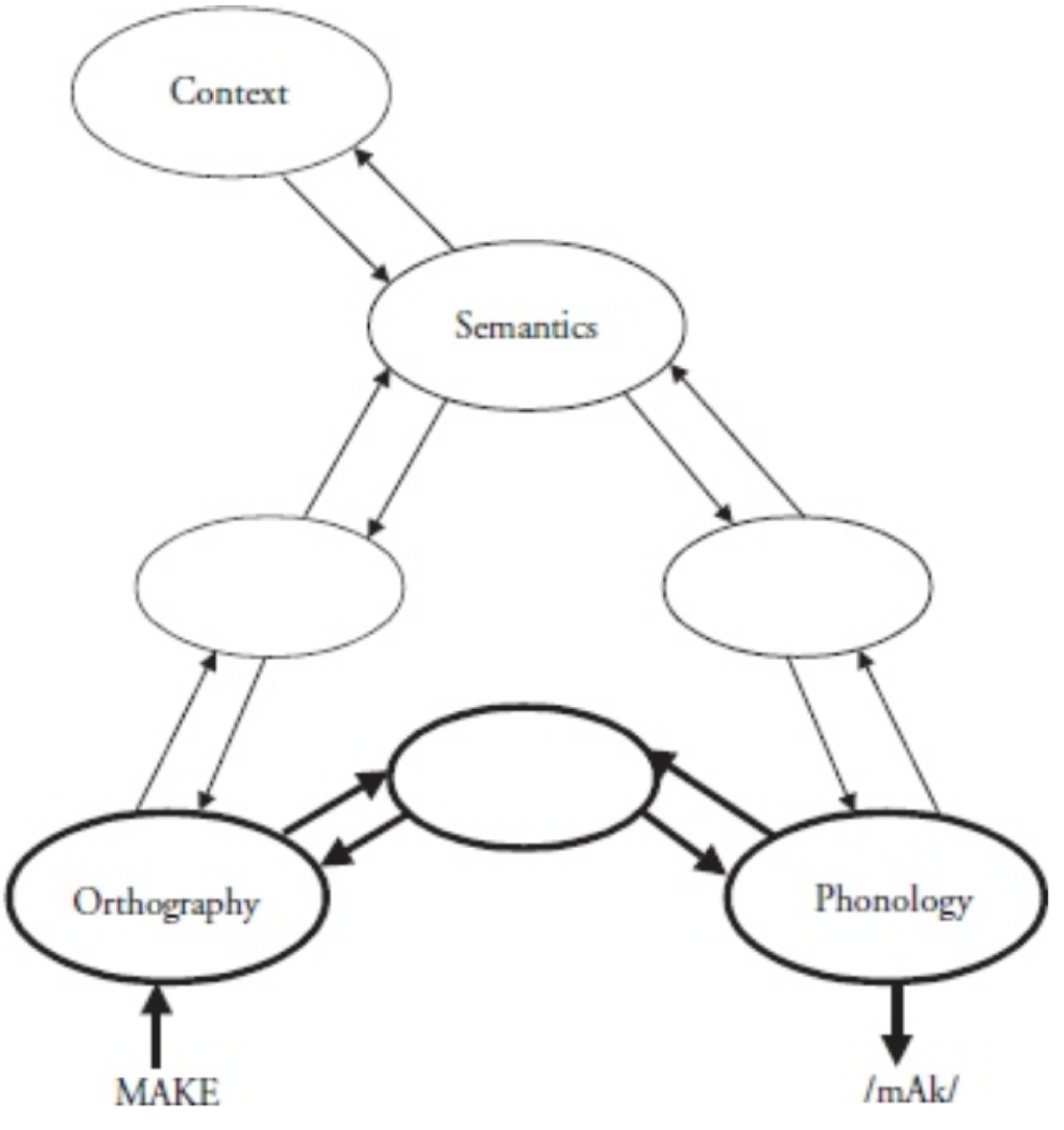
Connectionist and nonconnectionist modeling

This chapter distinguishes between connectionist models of reading (such as the models of Seidenberg & McClelland, 1989, and Plaut, McClelland, Seidenberg, & Patterson, 1996) and nonconnectionist models of reading (such as the DRC model). The description of the DRC model in Coltheart et al. (2001) uses the term "connection" and the model in fact "contains" about 4.5 million connections, in the sense of the term "connection" used by Coltheart et al. (2001). However, in the DRC model, connections are just expository devices used for talking about how the modules of the model communicate with each other. One could expound this in other ways without using the term "connection." In contrast, in connectionist models, the connections are often thought of as neuron-like, the models are referred to as neural networks, and terms like "biologically inspired" or "neurally plausible" are often applied. Here a connection is something that is physically realizable as an individual object, in contrast to the DRC model in which there is no such sense to the term.

A second major difference between connectionist and nonconnectionist modeling, at least as those trades have been practiced up until now, is that connectionist models have typically been developed by applying a neural-net learning algorithm to a training set of stimuli, whereas the architectures of nonconnectionist models have typically been specified by the modeler on the basis of the empirical effects that the model is meant to explain.

The Seidenberg and McClelland (1989) connectionist computational model of reading is often presented as an alternative to the dual-route model. Indeed, claims such as "The dual-route model has been more recently questioned by a plethora of single-route computational models based on connectionist principles" (Damper & Marchand, 2000, p. 13) are common in the literature. But that was not the view of the authors themselves. They were clear about this: "Ours is a dual route model," they stated (Seidenberg & McClelland, 1989, p. 559).

Figure 1.3 The Seidenberg and McClelland (1989) model. (The implemented model is in boldface type.)



This is perfectly evident from their diagram of their model (Seidenberg & McClelland, 1989, figure 1, reproduced as figure 1.3 here): it explicitly represents two distinct routes from orthography to phonology, one direct and the other via meaning, and explicitly represents two distinct routes from orthography to semantics, one direct and the other via phonology. One of the two routes for reading aloud (the one via semantics) can only be used for reading words aloud; it would fail for nonwords. The other (nonsemantic) route for reading aloud is required if the stimulus is a nonword. This model has come to be called the triangle model, perhaps because of the reference in Seidenberg and McClelland (1989, p. 559) to "the third side of the triangle in Figure 1." More than one subsequent model has been referred to as the triangle model despite being different from Seidenberg and McClelland's model. So far there have been seven different triangle models, an issue discussed later in this chapter.

What is it that has led to this widespread misunderstanding? The answer is clear: a failure to distinguish between the following two claims:

(a) It is possible for a single processing system to correctly read aloud all irregular words and all nonwords.

(b) The human reading system possesses only one procedure for computing pronunciation from print.

Seidenberg and McClelland (1989) did make claim (a). But they did not make claim (b); indeed, as the quotation in the previous paragraph indicates, they repudiated claim (b). That is why theirs is a dual-route model of reading aloud.

This seminal model turned out not to be able to offer a good account of how people read nonwords aloud because its accuracy on this task was far less than the accuracy that human readers show (Besner, Twilley, McCann, & Seergobin, 1990). The suggestion (Seidenberg & McClelland, 1990, p. 448) that this was because the database of words on which the model was trained was too limited and did not contain enough information for nonword reading to be learned from it was shown to be incorrect by Coltheart et al. (1993). They developed a GPC rule-learning algorithm and applied it to the Seidenberg–McClelland training set. The rule set that this algorithm learned from that training set was then used with 133 nonwords from Glushko (1979). Whereas the Seidenberg and McClelland model scored only 68% correct on a subset of 52 of these nonwords, the DRC read 97.9% of these correctly. This shows that the information needed to learn to be an excellent nonword reader is actually present in the model's database, and so "the poor performance of the PDP model in reading nonwords is a defect not of the database but of the model itself" (Coltheart et al., 1993, p. 594). Hence, as noted by Plaut (1997, p. 769) and (Plaut et al., 1996, p. 63), the Seidenberg and McClelland model did not succeed in providing evidence that it is possible for a single processing system to correctly read aloud all irregular words and all nonwords.

Nevertheless, it might well be possible to devise a single processing procedure that can correctly read aloud all irregular words and all nonwords. Plaut et al. (1996) sought to devise such a procedure via training a connectionist network similar in overall architecture to that of the network of Seidenberg and McClelland shown in figure 1.3 (it was, for example, a dual-route model in just the same sense that Seidenberg and McClelland viewed their model as a dual-route model, though training was carried out on only one of the two routes), but differing from the Seidenberg and McClelland model in a number of ways, including in the forms of orthographic and phonological representations used in the network. Input units, which were distributed representations in the Seidenberg and McClelland model, became local representations (each representing a grapheme). Output units, which were distributed representations in the Seidenberg and McClelland model, became local representations (each representing a phoneme).

Plaut et al. (1996) actually presented three different though related models – that is, a second, third and fourth triangle model, the first triangle model being that of Seidenberg and McClelland (1989):

Model 1: purely feedforward, 105 grapheme units, 100 hidden units, 61 phoneme units.

Model 2: as for Model 1 but with feedback from phoneme units back to hidden units: an attractor network.

Model 3: as for Model 1 but adding (unimplemented) external input to the output units, so as to mimic what could happen if there were an implemented semantic system activated by orthography and in turn activating phonology. This approach, discussed further below, was pursued in an attempt to simulate acquired surface dyslexia.

How well do these models read nonwords? Model 1 (which after training scored 100% on reading the 2,972 nonhomographic words in the training set) did quite well on nonword reading (see table 3 of Plaut et al., 1996), almost as well as human readers. However it still fails with items like JINJE, the reason being that there is no word in the training corpus that ends with the final grapheme of this nonword. It follows that careful selection of nonwords which exploits such gaps in the training corpus would produce a set of nonwords on which the model would score at or close to zero. Human readers would be vastly superior to the model on such nonwords. Results with nonword reading by Model 2 were similar, though its nonword reading was slightly worse than that of Model 1. The JINJE problem remained.

Given this work by Plaut et al. (1996), what are we to say about the two claims mentioned above? These claims were:

(a) It is possible for a single processing system to correctly read aloud all irregular words and all nonwords.

(b) The human reading system possesses just one procedure for computing pronunciation from print.

Although nonword reading was better by the PMSP models than by the SM model, the PMSP models still do not read nonwords correctly in the sense of "as well as human readers do," since it is not difficult to devise nonwords that human readers read well and the PMSP models read wrongly: there is no sense in which reading JINJE to rhyme with "wine" (as the PMSP models do) could be regarded as correct. So claim (a) remains without support. And no current model of reading aloud makes claim (b). Hence at present it is reasonable to regard both claims as false.

However, the work on simulation of surface dyslexia using Model 3 has an interesting implication for these claims. Indeed, in general simulation of disordered rather than normal reading it has been particularly crucial in recent years for comparative evaluation of computational models of reading. Hence much of the following discussion of dualroute modeling will focus on the application of such models to the explanation of disordered reading.

Simulating disordered reading with the triangle models

Simulating acquired surface dyslexia. Acquired surface dyslexia (Marshall & Newcombe, 1973; Patterson, Marshall, & Coltheart, 1985) is a reading disorder, caused by brain damage, in which there is selective impairment of the ability to read irregular words aloud with relative sparing of regular word and nonword reading. Many cases are not normal at regular word and nonword reading; I will focus here, as did Plaut et al. (1996), on two particularly pure cases, KT

(McCarthy & Warrington, 1986) and MP (Behrmann & Bub, 1992). Both showed virtually normal accuracy in reading aloud regular words and nonwords, but were impaired at reading irregular words, especially when these were low in frequency (KT: high frequency 47%; low frequency 26%; MP: high frequency 93%; low frequency 73%).

Computational models are meant to be able to explain impaired reading as well as normal reading: that is, it should be possible to artificially lesion these models so that their patterns of preserved and impaired reading correctly match such patterns seen in various forms of acquired dyslexia. Plaut and colleagues therefore investigated whether there was any way of lesioning any of their three models that would lead to impaired irregular word reading with preserved regular word and nonword reading.

This was investigated by studying the effects of deleting various proportions of the connections in the implemented orthography-to-phonology pathway, or various proportions of the hidden units, in Model 2. This was not successful in simulating the more severe patient KT: any lesion that produced accuracies of around 26% for low-frequency irregular words also produced very poor performance with nonwords, whereas KT was perfect at reading nonwords. It was therefore not possible to simulate acquired surface dyslexia just with the implemented part of the model.

So Plaut et al. turned from Model 2 to Model 3, which has an unimplemented component (semantic input to the phonological output level). With sufficient training, Model 3 does well with irregular words, regular words, and nonwords. What is crucial here, though, is the competence of the implemented (orthography-to-phonology) part of Model 3. When it is trained without semantics (this is Model 1), it learns to read irregular words perfectly and nonwords very well. But this is not the case when it is trained with concurrent semantic input. Low-frequency irregular words are never learned perfectly by the direct orthography-to-phonology pathway here: for this pathway operating on its own, accuracy for low-frequency irregular words is about 70% after 400 epochs of training and then declines down to about 30% correct after 2,000 epochs. Performance with high-frequency irregular words is almost perfect at 400 epochs, but further training progressively worsens performance with these words, down to about 55% at epoch 2,000. Regular word and nonword performance is almost perfect at epoch 400 and remains at that level with further training to epoch 2,000.

If training is stopped at 400 epochs, and semantic input to the system is then deleted, performance is good with regular words, nonwords, and high-frequency irregular words, but somewhat impaired with low-frequency irregular words; that matches the surface dyslexic pattern shown by MP.

If training is stopped at 2,000 epochs, and semantic input to the system is then deleted, performance is good with regular words, and nonwords, impaired with high-frequency irregular words, and very poor with low-frequency irregular words; that matches the surface dyslexic pattern shown by KT.

The suggestion here is that the cause of acquired surface dyslexia is semantic damage, and that the more the patient had relied on semantic input for reading aloud premorbidly, the more severe the surface dyslexia will be when semantic damage occurs. The implication is that, even if it is possible for a single processing system to correctly read aloud all irregular words and all nonwords, most human readers do not possess such a system.

Because there are patients with severe semantic damage who can read irregular words with normal accuracy (e.g., Cipolotti & Warrington, 1995; Lambon Ralph, Ellis, & Franklin, 1995; Schwartz, Saffran, & Marin, 1980a; see also Gerhand, 2001), Plaut et al. (1996, p. 99) had to suppose that some people learn to read without any support from semantics and so can read all irregular words without recourse to semantics. But in other work using the triangle models this supposition has been abandoned:

> It is important to note that, because this version of the triangle model assumes a causal relationship between semantic impairment and surface dyslexia, its adequacy is challenged by any observations of semantically impaired patients whose reading does not reveal a surface dyslexic pattern. (Fushimi et al., 2003, p. 1656)

> A degraded semantic system will inevitably impair the ability to "know" a letter string… as belonging to the repertoire of real words. (Rogers, Lambon Ralph, Hodges, & Patterson, 2004, p. 347)

According to Model 3 as it is applied to the analysis of surface dyslexia, intact human readers possess two routes from print to speech. Let's call these, theory-neutrally, Route A and Route B. Properties of these routes are:

(a) Route A can correctly read aloud all known words (regular or irregular) but cannot read nonwords aloud correctly.

(b) Route B can correctly read aloud all regular words and all nonwords, but will misread X% of irregular words.

This connectionist dual-route model of reading aloud differs from the nonconnectionist dual-route DRC model of reading aloud (Coltheart et al., 2001, discussed below) only with respect to the value of X. According to Plaut et al. (1996), premorbidly X can on rare occasions be zero (the patients referred to above who are normal at irregular word reading but have severe semantic impairments) but typically is not and can be at least as high as 64% (patient KT's overall error rate on irregular words). According to the DRC model, X is always 100%.

So, while it is of course logically possible that the system humans use for reading aloud has a single-route architecture, there are no theoretical proposals embodying such an architecture that can escape refutation from available data from studies of normal and impaired readers. All the models are dual-route models. Current and future theorizing is and will be about the details of what these two routes are actually like.

Simulating acquired phonological dyslexia

Harm and Seidenberg (2001) used another connectionist triangle model in work attempting to simulate acquired phonological dyslexia. In their view, this form of acquired dyslexia is always caused by a phonological impairment. Therefore, after training their model until it was performing well in reading words and nonwords, they lesioned the phonological component of the model by adding random noise each time the units in that component were being updated. This harmed nonword reading more than word reading and so simulated phonological dyslexia. However, this explanation of acquired phonological dyslexia predicts that cases of acquired phonological dyslexia without the presence of a phonological impairment will not be seen, and this prediction is incorrect. Dérouesné and Beauvois (1985), Bisiacchi, Cipolotti, and Denes (1989), and Caccappolo-van Vliet, Miozzo, & Stern (2004) have all reported cases of acquired phonological dyslexia with preserved phonological processing.

As we have seen, the development of connectionist triangle models of reading has been considerably influenced by attempts to simulate acquired dyslexia; and this approach has also been applied to the simulation of developmental dyslexia.

Simulating developmental dyslexia. Harm and Seidenberg (1999) developed a model in which to simulate developmental reading disorders. Their particular triangle model differed from all earlier triangle models in a number of ways:

(a) Learning in the phonological units was assisted by the presence of a set of cleanup units attached to the phonological units.

(b) The phonological units represented phonetic features, not phonemes.

(c) The orthographic units represented letters, not graphemes.

(d) Positional coding of orthography was relative to the vowel in the input string, rather than absolute.

After training, the model achieved satisfactory levels of performance in reading the irregular words in the training set, and also in reading nonwords (though again performance seemed slightly inferior to human nonword reading).

Harm and Seidenberg (1999) were specifically interested in attempting to simulate developmental dyslexia. Having shown that their triangle model was capable of learning to read adequately, they then investigated ways of impeding its learning that might result in either of two different subtypes of developmental dyslexia, one in which nonword reading is selectively affected (developmental phonological dyslexia) and another in which irregular word reading is selectively affected (developmental surface dyslexia; Harm and Seidenberg preferred the term “reading delay dyslexia” because they believed that the reading of children with developmental surface dyslexia is just like the reading of younger children who are learning to read normally).

Because Harm and Seidenberg (1999) believed that developmental phonological dyslexia is always caused by the child having a phonological processing deficit, their approach to simulating developmental phonological dyslexia involved lesioning their model's phonological system. This was done in two different ways:

(a) Mild phonological impairment: a slight degree of weight decay was imposed on the phonetic feature units throughout training.

(b) Moderate phonological impairment: in addition to the weight decay, the cleanup units were removed from the network, as were a random 50% of the interconnections between the phonetic feature units.

Both types of lesioning did impair the model's ability to learn to read nonwords. But when this impairment was more than mild, the ability of the model to learn to read words was also impaired. Hence what could not be simulated here was pure severe developmental phonological dyslexia (where "pure" means that word reading is in the normal range and "severe" means the impairment of nonword reading was more than mild). That raises the question: does one ever see pure severe developmental phonological dyslexia in human readers? A number of such cases have been reported (see e.g. Campbell & Butterworth, 1985; Funnell & Davison, 1989; Holmes & Standish, 1996; Howard & Best, 1996; Stothard et al., 1996). Hence these data from developmental cognitive neuropsychology provide a challenge for the Harm and Seidenberg (1999) connectionist model of reading.

Developmental surface dyslexia ("reading delay dyslexia") was simulated in the work of Harm and Seidenberg (1999) by reducing the number of hidden units in the network from 100 to 20, and also by reducing the network's learning rate. Both types of developmental damage to the network harmed the learning of irregular words more than the learning of nonwords; but in both cases the learning of nonwords suffered too. Thus it was not possible to simulate "pure" developmental surface dyslexia (i.e., impaired irregular word reading with *normal* nonword reading). However, pure developmental surface dyslexia is seen in human readers (Castles & Coltheart, 1996; Hanley & Gard, 1995; Goulandris & Snowling, 1991). Hence again these data from developmental cognitive neuropsychology do not provide support for the Harm & Seidenberg (1999) connec-tionist model of reading.

Conclusions

Reading theorists have reached unanimity concerning the existence in the human reading system of two separate procedures for reading aloud – that is, dual routes from print to speech. One of these processing routes is usable only when the stimulus to be read is a real word; it cannot read nonwords. The other route can read all nonwords and regular words; there is still some dispute concerning how well it reads irregular words.

These dual-route models differ in terms of whether they are connectionist models such as the triangle models or nonconnectionist models such as the DRC model. At present the data favor the nonconnectionist approach. The DRC model does a good job of simulating patterns of acquired dyslexia, which the connectionist models have not succeeded in doing. Nor have the connectionist models succeeded in accounting for developmental reading disorders, whereas the DRC model is compatible with everything we currently know about these disorders. Finally, none of the connectionist models can explain all of the phenomena from studies of normal reading listed above (see the section "What the DRC Model Can Explain"), whereas all of these can be simulated by the DRC model.

2

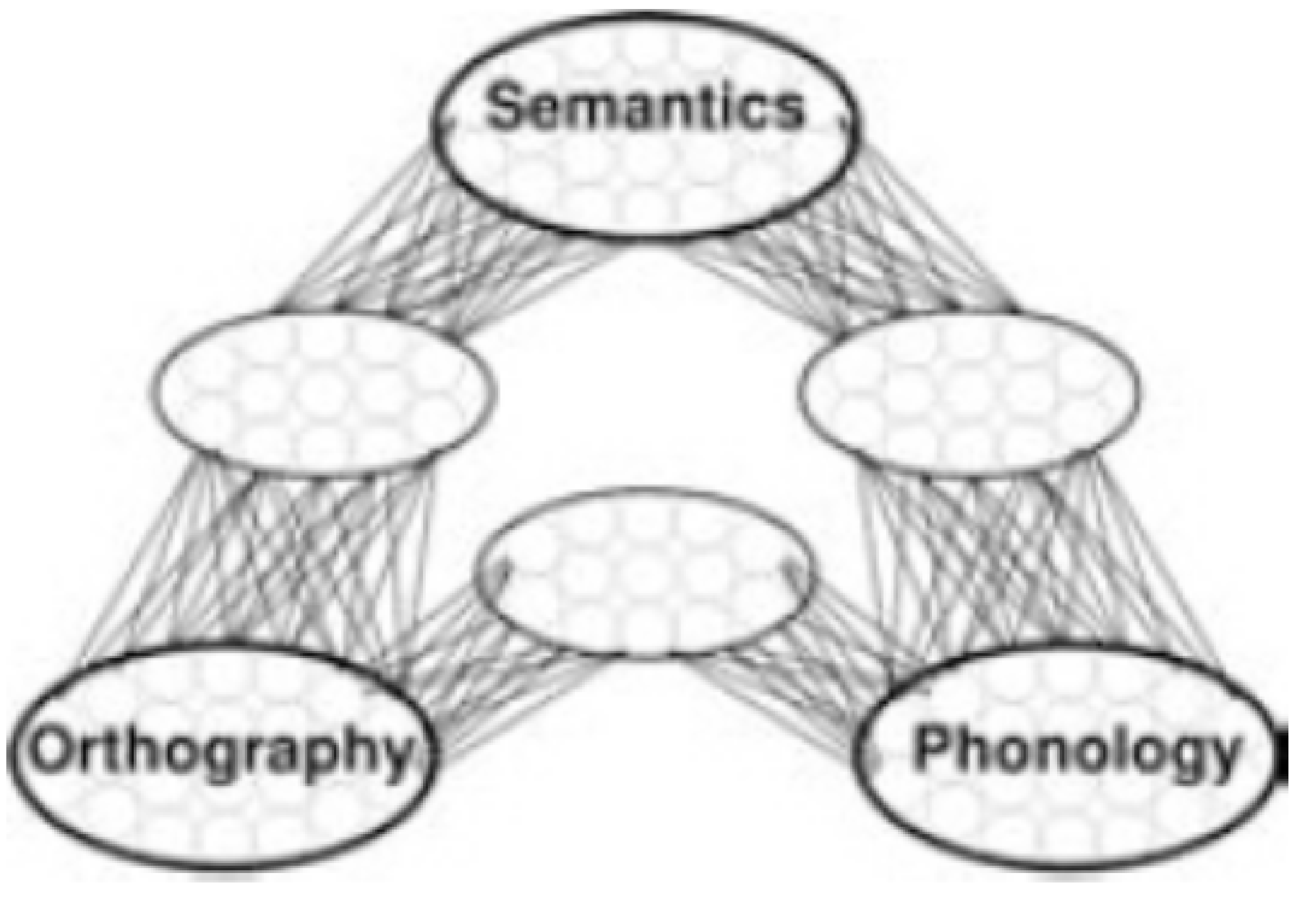
Connectionist Approaches to Reading

David C. Plaut

Reading is a highly complex task involving the rapid coordination of visual, phonological, semantic, and linguistic processes. Computational models have played a key role in the scientific study of reading. These models allow us to explore the implications of specific hypotheses concerning the representations and processes underlying reading acquisition and performance. A particular form of computational modeling, known as connectionist or neural network modeling, offers the further advantage of being explicit about how such mechanisms might be implemented in the brain.

In connectionist models, cognitive processes take the form of cooperative and competitive interactions among large numbers of simple neuron-like processing units. Typically, each unit has a real-valued activity level, roughly analogous to the firing rate of a neuron. Unit interactions are governed by weighted connections that encode the longterm knowledge of the system and are learned gradually through experience. Units are often organized into layers or groups; the activity of some groups of units encode the input to the system; the resulting activity of other groups of units encodes the system's response to that input. For example, one group might encode the written form (orthography) of a word, another might encode its spoken form (phonology), and a third might encode its meaning (semantics; see figure 2.1). The patterns of activity of the remaining groups of units – sometimes termed "hidden" units – constitute learned, internal representations that mediate between inputs and outputs. In this way, the connectionist approach attempts to capture the essential computational properties of the vast ensembles of real neuronal elements found in the brain using simulations of smaller networks of more abstract units. By linking neural computation to behavior, the framework enables developmental, cognitive, and neurobiological issues to be addressed within a single, integrated formalism. One very important advantage of connectionist models is that they deal explicitly with learning. Though many of these models have focused predominantly on simulating aspects of adult, rather than children's, reading, many of the models do explicitly consider the process of learning (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989). In essence, such models instantiate learning as a process as a slow incremental increase in knowledge, represented by increasingly strong and accurate connections between different units (e.g., the letters in printed words and the phonemes in spoken words to which they correspond).

Figure 2.1 A connectionist network that relates orthographic, phonological, and semantic information in word reading and other lexical tasks, based on the "triangle" framework (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989).



Another critical feature of many connectionist systems is that after learning they show the ability to generalize (e.g., to pronounce novel words which they have not been trained on). Finally, and related to this, such systems often show graceful degradation when damaged. Removing units or connections in such systems typically does not result in an all-or-none loss of knowledge; rather, damage results in a gradual degradation of performance. These three aspects of connectionist models have clear parallels in human reading behavior – children gradually learn to read more and more words in an incremental fashion over a long period, such learning brings with it the ability to generalize to novel items children have not been taught, and in cases of brain damage there are often graded declines in performance with inconsistent performance at different times. The fact that connectionist models display such parallels to human reading behavior has generated considerable excitement at the prospect that such models may offer new, explicit, and detailed accounts of how reading is implemented in the human brain.

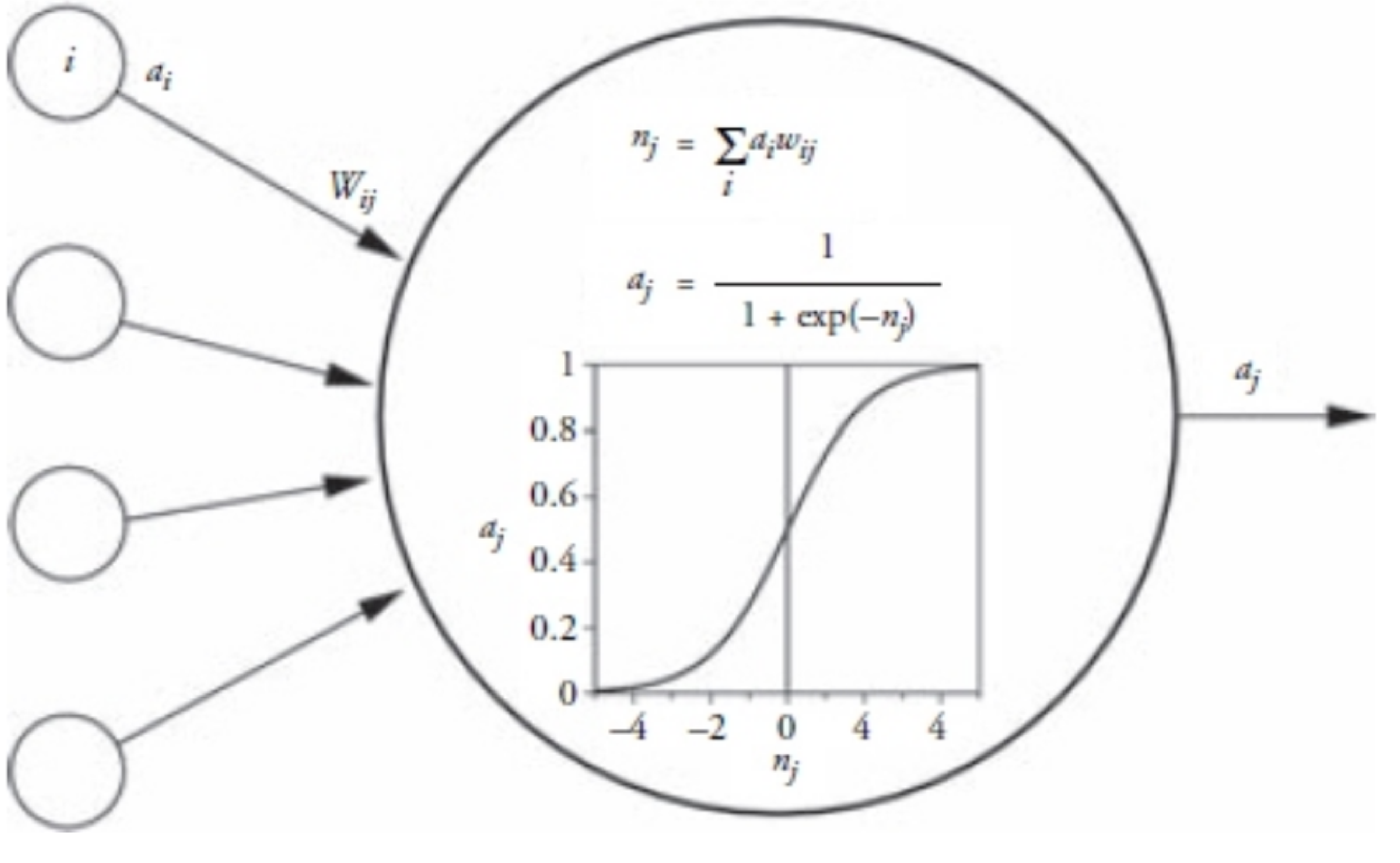
Principles of Connectionist Modeling

Before turning to how specific connectionist models have been applied to various reading-related phenomena, it will be helpful to consider the implications of the underlying computational principles more generally. These can be grouped into issues related to processing, representation, learning, and network architecture.

Processing

A standard connectionist unit integrates information from other units by first computing its *net input*, equal to a linear sum of positive-and negative-weighted activations from sending units, and then setting its own activation according to a nonlinear, monotonically increasing (sigmoid) function of this net input (see figure 2.2). In some networks, unit activations change gradually in response to input from other units instead of being recomputed from scratch each time.

Figure 2.2 The operation of a standard connectionist unit (indexed by *j*), which computes a net input *nj* as a weighted sum of activations a from other units (indexed by *i*), and then computes its own activation a as a smooth, nonlinear (sigmoid) function of its net input (where exp(.) is the exponential function).



Both the linear integration of net input and the nonlinear activation function play critical roles in shaping how connectionist networks behave. The fact that the net input to each unit is a simple weighted sum is at the heart of why networks exhibit similarity-based generalization to novel inputs (e.g., being able to pronounce a pseudoword like MAVE based on knowledge of words like GAVE, SAVE, MATE, etc.). If a unit is presented with a similar pattern of activity along its input lines, it will tend to produce a similar net input and, hence, a similar response. This fails to hold only if the weights for those inputs that differ between the patterns are very large, but such large weights develop during learning only when necessary (e.g., when handling exceptional cases; see the section on learning below).

If all processing in the network were strictly linear, however, the types of mappings it could learn would be severely limited (Minsky & Papert, 1969). The nonlinear activation function allows individual units – and hence the network as a whole – to preserve some types of similarity in its response while ignoring others. The sigmoid activation function asymptotes for large positive or negative net inputs, but produces roughly proportional responses for small and moderate net inputs (see figure 2.2). If networks start out with relatively small weights, most units' activations will fall in the linear range of the sigmoid function, and the network as a whole will give similar responses to similar inputs. However, when aspects of a task require responses that are not predicted by input similarity (e.g., pronouncing SEW like SO instead of SUE, or mapping CAP and CAT to completely different meanings), learning must develop sufficiently large weights to drive the relevant units into their nonlinear (asymptotic) range, where changes in net input have little if any effect on activation. In this way, a network can remain largely linear for systematic or "regular" aspects of a task, while simultaneously exhibiting nonlinear behavior for the unsystematic or "irregular" aspects.

Understanding how a connectionist network operates above the level of individual units requires consideration of how patterns of activity across the various groups of units interact and evolve over the course of processing a given input. A very useful concept in this regard is the notion of an *attractor*. At any given instant, the current pattern of activity over a group of units in the network (or over the network as a whole) can be represented in terms of the coordinates of a point in a multidimensional *state space* that has a dimension for each unit. As the pattern of activity changes during processing, the corresponding point in state space moves. In many networks, unit interactions eventually reach a state in which the activation of each unit is maximally consistent with those of other units and the pattern as a whole stops changing. The point in state space corresponding to this final pattern is called an attractor because interactions among units in the network cause nearby points (i.e., similar patterns) to be "pulled" towards the same final attractor point. (The region around an attractor that settles to it is called its *basin* of attraction.) The stability of attractor patterns gives networks a considerable degree of robustness to partially missing or noisy inputs, or to the effects of damage.

Representation

As described thus far, a typical connectionist network processes an input through unit interactions that cause the network to settle to an attractor, in which the resulting pattern of activity over output units corresponds to the network's response to the input. An issue of central relevance is the nature of the representations that participate in this process – the way that inputs, outputs, and groups of intermediate units encode information in terms of patterns of activity. Some connectionist models use *localist* representations, in which individual units stand for familiar entities such as letters, words, concepts, and propositions. Others use *distributed* representations, in which each such entity is represented by a particular pattern of activity over many units rather than by the activity of an single unit. Localist representations can be easier to think about and to manipulate directly (Page, 2000), but often permit too much flexibility to constrain theorizing sufficiently (Plaut & McClelland, 2000). By contrast, distributed representations are typically much more difficult to use and understand but can give rise to unanticipated emergent properties that contribute in important ways to the explanation of cognitive phenomena (see e.g. Hinton & Shallice, 1991).

Given that, as explained above, similar patterns tend to have similar consequences in connectionist networks, the key to the use of distributed representations is to assign patterns to entities in such a way that the similarity relations among patterns captures the underlying functional relationships among the entities they represent. For groups of units that must be interpreted directly (i.e., inputs and outputs), this is done based on independent empirical evidence concerning the relevant representational similarities. However, except for the simplest of tasks, it is impossible to perform the relevant mappings without additional intermediate units, and it is infeasible to specify appropriate connection weights for such units by hand. Accordingly, distributed connectionist networks almost invariably use learning to discover effective internal representations based on task demands.

Learning

The knowledge in a network consists of the entire set of weights on connections among units, because these weights govern how units interact and hence how the network responds to any given input. Accordingly, learning involves adjusting the weights in a way that generally benefits performance on one or more tasks (i.e., mapping from inputs to outputs).

Connectionist learning procedures fall into three broad classes based on how much performance feedback is available. At one extreme are *unsupervised* procedures, such as Hebbian learning (as it is typically applied; Hebb, 1949), that make no use of performance feedback and, instead, adjust connection weights to capture the statistical structure among activity patterns. At the other extreme are *supervised* procedures, such as back-propagation (Rumelhart, Hinton, & Williams, 1986), that assume the learning environment provides, for every trained input pattern, a fully specified "target" pattern that should be generated over the output units. Between these two extremes are *reinforcement* procedures, such as temporal difference methods (Sutton, 1988), that assume the environment provides potentially intermittent evaluative feedback that does not specify correct behavior but rather conveys the degree to which behavioral outcomes were good or bad.

When performance or evaluative feedback is available, it is relatively straightforward to use it to adapt connection weights to improve performance. If the activation of an output unit is too high, it can be reduced by decreasing positive incoming weights and the corresponding sending activations and by increasing (in magnitude) negative weights and sending activations (see the equations in figure 2.2); the reverse is true if output activation is too low. Changing the sending activations involves reapplying the same procedure to their incoming weights and incoming activations, and so on. Specific algorithms differ in how they compute feedback and how they distribute information on how to change weights.

Many applications of distributed connectionist modeling to cognitive phenomena use back-propagation despite its biological implausibility (Crick, 1989). This is partly because, unlike most alternatives, the procedure is effective at learning difficult mappings, including those with complex temporal characteristics (Williams & Peng, 1990). It is also the case that the time-course and ultimate outcome of learning with back-propagation is highly similar to the properties of more biologically plausible supervised procedures, such as Contrastive Hebbian learning (Ackley, Hinton, & Sejnowski, 1985; O'Reilly, 1996; Peterson & Anderson, 1987). Thus, one can interpret back-propagation as a computationally efficient means of learning internal representations in distributed connectionist networks in a way that approximates the properties of performance-driven learning in the brain.

Network architecture

The *architecture* of a network – the pattern of connectivity among and within groups of units representing different types of information – can have an important impact on the behavior of a connectionist model in its acquisition, skilled performance, and impairment following damage. The strong emphasis on learning in the development of connectionist models has led some researchers to conclude that the approach disavows any built-in structure within the cognitive system. A more accurate characterization would be that the effectiveness of learning in connectionist networks makes it possible to explore the degree to which built-in structure is necessary to account for some empirical phenomena. The modeling framework itself allows for the expression of a wide variety of network architectures, ranging from those with extensive built-in structure to those with minimal structure.

Connectionist models often contrast with alternative formulations in terms of the *kinds* of distinctions that are instantiated in the architecture of the system. A classic example is the traditional separation of rule-based and item-based mechanisms in “dual-route” theories of word reading (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) and inflectional morphology (Pinker, 1999). Because the processing mechanisms within a connectionist system are homogeneous – involving massively parallel unit interactions throughout – the underlying theories rarely isolate different types of *processing* into separate systems or pathways. Rather, architectural divisions typically reflect different types of *information* (e.g., orthographic, phonological, semantic). Given that such distinctions often correspond to modalities of input or output, they can be supported directly by data on neuroanatomic localization of the corresponding neural representations.

Realist Versus Fundamentalist Approaches

Before turning to an overview of connectionist models of reading, it is worth distinguishing two broad approaches to cognitive modeling, because they often have rather different goals. The *realist* approach tries to incorporate into a model as much detail as possible of what is known about the real system in the belief that complex interactions of these factors are necessary to capture the relevant phenomena. The *fundamentalist* approach, by contrast, holds that a model should, as much as possible, embody only those principles that are claimed to account for the relevant phenomenon and should abstract out extraneous details. In evaluating any given modeling effort, it is important to identify the specific goals of the work; some models are intended to provide comprehensive accounts of detailed behavioral data, whereas others are intended more as demonstrations of specific computational arguments. Often the most effective modeling approach over the long term is to begin with fundamentalist models to elucidate the key underlying principles, and then gradually move towards more realist models as the theoretical implications of additional details become understood.

Connectionist Modeling of Reading

Most connectionist models of reading have focused on single word processing as it is generally thought that, above the lexical level, written language engages largely the same mechanisms as spoken language. In the review that follows, these models are characterized in terms of whether their representations for words are localist (one unit per word) or distributed (alternative patterns of activity for each word) and whether they focus on the task of word recognition (deriving a lexical or semantic representation) or oral reading (deriving a pronunciation).

Localist models of word recognition

One of the earliest and arguably most influential connectionist models of reading is a localist, nonlearning model – the interactive activation and competition (IAC) model of letter and word perception (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). The model consists of three layers of units – letter feature units, letter units, and word units. The model was designed to recognize four-letter words, so there is a separate set of feature units and letter units for each of four letter positions. The activation of each unit can be thought of as reflecting the network's confidence in the hypothesis that the entity represented by the unit (e.g., a T in the first position, or the word TAKE) is part of the correct interpretation. The weights on connections between units reflect the degree to which one hypothesis is consistent or inconsistent with another. Within each level, units representing inconsistent hypotheses (e.g., a T versus a P in the first letter position, or the words TAKE and TRIP) have negative connections between them. Between levels, units representing consistent hypotheses (e.g., a top horizontal letter feature and the letter T, or a T in the first position and the word TAKE) have positive connections between them, whereas units representing inconsistent hypotheses (e.g., a P in the first position and the word TAKE) have negative connections between them. Connections throughout the system are bidirectional, allowing both top-down and bottom-up information to influence unit activations.

A primary goal of the model was to explain the *word superiority effect* (Reicher, 1969; Wheeler, 1970), in which the perception of a briefly presented letter is more accurate when it occurs in a word compared with when it occurs in a random consonant string or even in isolation (see Lupker, this volume). In the IA model, this effect arises due to partial activation of word units that provide top-down support for the letters they contain. The model was also able to explain the *pseudoword superiority effect* (e.g., Carr, Davidson, & Hawkins, 1978; McClelland & Johnston, 1977), in which letters occurring in pronounceable nonwords (e.g., MAVE) are perceived better than in consonant strings or in isolation (although not quite as well as in words). Although pseudowords are not fully consistent with any of the units at the word level in the model, they are partially consistent with many words. The presentation of a pseudoword typically generates weak activation of word units sharing three of its four letters; these units, in turn, conspire to provide top-down support for the letters in the pseudowords. In this way, the IA model provided an early demonstration of how even a localist model can generalize on the basis of similarity, through the use of what are essentially distributed representations for pseudowords.

In subsequent work, McClelland (1991) (see also Movellan & McClelland, 2001) elaborated the model to use units with an intrinsically noisy or stochastic activation function to bring the model in line with empirical evidence for statistical independence in how people integrate multiple sources of information (Massaro, 1988). More recently, Grainger and Jacobs (1996) generalized the interactive activation framework to address a broader range of tasks and issues related to word recognition.

Distributed models of word recognition

Mozer (1991) developed a connectionist model of object recognition and spatial attention, called MORSEL, that was applied to the specific task of recognizing words. In the model, an attentional system forms a spatially contiguous bubble of activation that serves to select a subset of the bottom-up letter feature information for further processing by a hierarchically organized object recognition system. Each layer in the recognition system (called BLIRNET) consists of units with spatially restricted receptive fields that form conjunctions of the simpler features in the previous layer. At the top of the system are position-independent units that respond to specific triples of letters (following Wickelgren's [1969] proposal for representing spoken words). In this way, words were represented by a pattern of activity over multiple letter triples (e.g., #HO, OUS, USE, SE#, for the word HOUSE) rather than by the activation of a single word unit (as in the IA model). Although there was no learning in the system, it was still successful at activating the correct set of letter triples for a fairly large vocabulary of words. When presented with multiple words, it usually selected and recognized one of them accurately but, like human subjects, would occasionally misrecognize the attended word due to letter migrations from the unattended word (Mozer, 1983). Moreover, when one side of the attentional mechanism was impaired, the damaged model exhibited all of the major characteristics of neglect dyslexia, the manifestation of hemispatial neglect with written words as stimuli (Mozer & Behrmann, 1990).

Although MORSEL used distributed word representations, it did not employ learning. Other distributed models have cast the problem of word recognition as mapping from the written forms of words to their meanings (rather than to higher-order orthographic representations, as in MORSEL), and have used learning to develop weights that accomplish this mapping. Note, however, that, apart from morphological relationships, the relationship between the surface forms of words and their meanings is largely arbitrary. In other words, similarity in form (e.g., CAT, CAP) is unrelated to similarity in meaning (e.g., CAT, DOG). This is the most difficult type of mapping for connectionist networks to learn, given their inherent bias towards preserving similarity. In fact, some researchers questioned whether it was even possible for distributed networks to accomplish this mapping without word-specific intermediate units. Kawamoto (1993) used a variant of Hebbian learning to train a distributed network to map among orthographic, phonological, and semantic representations (see also Van Orden, Pennington, & Stone, 1990). However, because the network lacked any hidden units, it could learn a vocabulary of only a few words. Nonetheless, Kawamoto was able to show that the model provided a natural account of a number of phenomena related to lexical semantic ambiguity resolution (see also Kawamoto, Kello, & Jones, 1994).

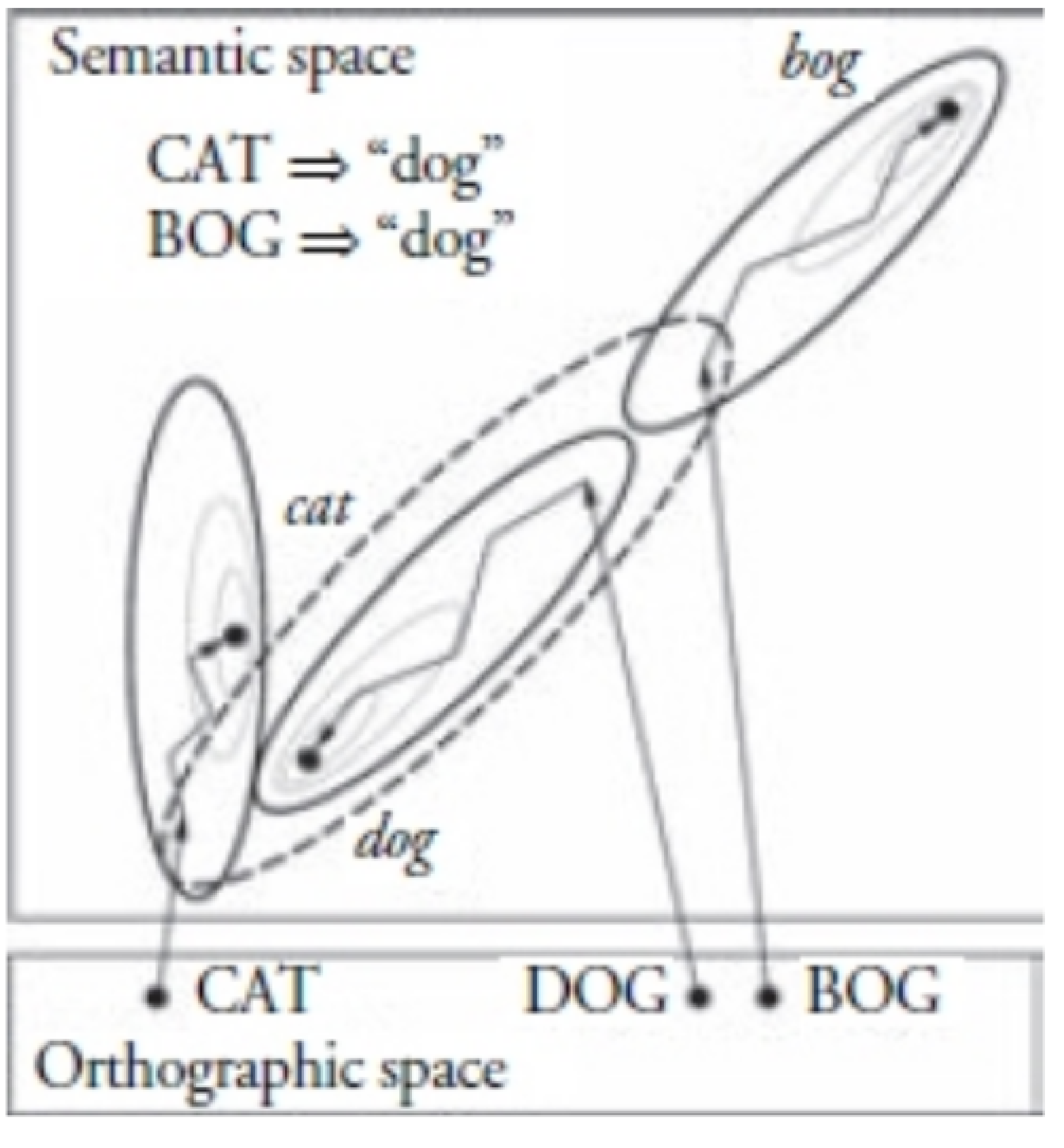
To address the more general challenge, Hinton and Sejnowski (1986) trained a Boltzmann Machine – a network of stochastic binary units – to map between orthography and semantics for a larger (although still small) set of words. Although training was difficult, the network was able to develop distributed representations over intermediate hidden units that accomplished the mapping. They also found that, with mild damage, the network occasionally responded to a word by giving another, semantically related word as a response (e.g., CAT read as DOG) – a *semantic error* reminiscent of those made by patients with *deep dyslexia* (Coltheart, Patterson, & Marshall, 1980).

Following Hinton and Sejnowski (1986), Hinton and Shallice (1991) used back-propagation to train a recurrent network with hidden units to map from orthography to semantics for 40 words falling into five concrete semantic categories. Orthographic representations were based on position-specific letter units; semantic representations consisted of subsets of 68 hand-specified semantic features that captured a variety of conceptual distinctions among word meanings. When the network was damaged by removing some units or connections, it no longer settled normally; the initial semantic activity caused by an input would occasionally fall within a neighboring attractor basin, giving rise to an error response. These errors were often semantically related to the stimulus because words with similar meanings correspond to nearby attractors in semantic space. Like deep dyslexic patients, the damaged network also produced errors with visual similarity to the stimulus (e.g., BOG read as DOG) and with both visual and semantic similarity (e.g., CAT read as RAT), due to its inherent bias towards similarity: visually similar words tend to produce similar initial semantic patterns, which can lead to a visual error if the basins are distorted by damage (see figure 2.3).

Plaut and Shallice (1993) extended these initial findings in a number of ways. They established the generality of the co-occurrence of error types across a wide range of simulations, showing that it does not depend on specific characteristics of the network architecture, the learning procedure, or the way responses are generated from semantic activity. They also showed that distributed attractor networks exhibited a number of other characteristics of deep dyslexia not considered by Hinton and Shallice (1991), including the occurrence of visual-then-semantic errors, greater confidence in visual as compared with semantic errors, and relatively preserved lexical decision with impaired naming. They also extended the approach to address effects of concreteness on word reading in deep dyslexia. They trained a network to pronounce a new set of words consisting of both concrete and abstract words. Concrete words were assigned far more semantic features than were abstract words, under the assumption that the semantic representations of concrete words are less dependent on the contexts in which they occur (Saffran, Bogyo, Schwartz, & Marin, 1980).

As a result, the network developed stronger attractors for concrete than abstract words during training, giving rise to better performance in reading concrete words under most types of damage, as observed in deep dyslexia. Surprisingly, severe damage to connections implementing the attractors at the semantic level produced the opposite pattern, in which the network read *abstract* words better than concrete words. This pattern of performance is reminiscent of CAV, the single, enigmatic patient with *concrete word dyslexia* (Warrington, 1981). The double dissociation between reading concrete versus abstract words in patients is often interpreted as implying that there are separate modules within the cognitive system for concrete and abstract words. The Plaut and Shallice simulation demonstrates that such a radical interpretation is unnecessary: the double dissociation can arise from damage to different parts of a distributed network which processes both types of items but develops somewhat different functional specializations through learning (see also Plaut, 1995a).

Figure 2.3 How damage to an attractor network can give rise to both semantic and visual errors. Points within each rectangular area correspond to specific patterns of activation over orthographic or semantic representations; neighboring points corresponding to similar (overlapping) patterns. The arrows reflect the way in which these patterns change over the course of processing. The solid ovals represent the basins of attraction in the normal network; the dashed ovals represent alterations of these basins due to damage (based on Hinton & Shallice, 1991).



"Dual-route" models of reading aloud

Much of the controversy surrounding theories of word reading centers not around how words are recognized and understood but how they are read aloud. In part, this is because, in contrast to the arbitrary nature of form-meaning mappings, the mapping between the written and spoken forms of words is highly systematic; words that are spelled similarly are typically also pronounced similarly. This property derives from the fact that written English follows an *alphabetic principle* in which parts of written forms (letters and multiletter graphemes like TH, PH) correspond to parts of spoken forms (phonemes). The sharp contrast between the systematic nature of pronunciation and the arbitrary nature of comprehension has led a number of researchers (e.g., Coltheart, 1978; Marshall & Newcombe, 1973) to propose separate pathways or "routes" for these two tasks, each employing very different computational mechanisms: a *sublexical* pathway employing grapheme–phoneme correspondence (GPC) rules for pronunciation, and a *lexical* pathway involving a word-specific lexical look-up procedure for comprehension (characterized much like the IA model in later formulations; see e.g. Coltheart et al., 2001; Coltheart, this volume). Complications arise, however, because the pronunciation task itself is not fully systematic; roughly 20% of English words are *irregular* in that they violate the GPC rules (e.g., SEW, PINT, YACHT). So-called "dual-route" theories propose that pronouncing such words also depends on the lexical pathway.

Although traditional dual-route models implement the sublexical pathway with symbolic rules (Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart et al., 2001), it is perfectly feasible to build a dual-route mechanism out of connectionist hardware. For example, Zorzi, Houghton, and Butterworth (1998) describe simulations in which direct connections from letter units to phoneme units support the pronunciation of regular words and nonwords, whereas a separate pathway, composed either of hidden units or localist word units, supports the pronunciation of irregular words (see also Ans, Carbonnel, & Valdois, 1998). Although the mechanisms employed for the two pathways are more homogeneous than in more traditional, rule-based implementations, the models nonetheless retain a categorical distinction between words that obey spelling-sound rules and words that violate them.

Distributed models of reading aloud

The first researchers to take on the challenge of training a single connectionist network to pronounce all English words were Sejnowski and Rosenberg (1987), who developed a system called NETtalk. Orthographic input was presented to NETtalk by sweeping a 7- letter window over a large text corpus (the Brown corpus; Kucera & Francis, 1967), successively centering the window on each letter in the text. For each letter position, the system was trained to generate the single phoneme corresponding to the central letter in the window. This allows each successive letter to be processed by the same set of units, so the knowledge extracted in processing letters in any position are available for processing letters in every other position. At the same time, the presence of other letters in the surrounding slots allows the network to be sensitive to the context in which letters occur. This is necessary not only for pronouncing exception words but also for handling multiletter graphemes (e.g., TH, PH, SH). For these, the system was trained to generate the appropriate phoneme for the first letter and then silence for the remaining letters. The alignment of phonemes to letters was specified by hand.

Although impressive as a first attempt, the performance of NETtalk when judged in terms of entire words pronounced correctly was much poorer than skilled readers. In follow-up work, Bullinaria (1997) showed that performance in a NETtalk-like system could be improved dramatically by allowing the network to discover the best letter-phoneme alignment by itself. This was done by evaluating the network's output against all possible alignments, and training towards the one that yields the lowest overall error. This pressures the system to converge on alignments that are maximally consistent across the entire training corpus, yielding perfect performance on words and good generalization to pronounceable nonwords.

The need for strictly sequential processing on even the shortest words raises questions about the psychological plausibility of the NETtalk approach. One way to address this concern is to propose that skilled readers attempt to process as much of the input as they can in parallel, then redirect fixation and continue. In this view, unskilled reading may be strictly sequential, as in NETtalk, but as skill develops, it becomes much more parallel. To explore this possibility, Plaut (1999) trained a simple recurrent (sequential) network to produce sequences of single phonemes as output when given position-specific letters as input. The network was also trained to maintain a representation of its current position within the input string. When the network found a peripheral portion of the input difficult to pronounce, it used the position signal to refixate the input, shifting the peripheral portion to the point of fixation where the network had had more experience in generating pronunciations. In this way, the network could apply the knowledge tied to the units at the point of fixation to any difficult portion of the input. Early on in training, the network required multiple fixations to read words, but as the network became more competent it eventually read most words in a single fixation. The network could also read nonwords about as well as skilled readers, occasionally falling back on a refixation strategy for difficult nonwords. Finally, a peripheral impairment to the model reproduced the major characteristics of letter-by-letter reading in pure alexic patients (Behrmann, Plaut, & Nelson, 1998b). Specifically, when input letter activations were corrupted with noise, the model exhibited a clear effect of orthographic length in its number of fixations (a loose analog to naming latency), and this effect interacted with lexical frequency such that the increase was much greater for low- compared with high-frequency words.

An alternative approach to word reading, first articulated by Seidenberg and McClelland (1989), casts the problem as learning to map among orthographic, phonological and semantic representations for entire words in parallel (see figure 2.1). The approach does not deny the existence of sequential processes related to both visual input and articulatory output, but emphasizes the parallel interactions among more central types of lexical information. In support of this general "triangle" framework, Seidenberg and McClelland (1989) trained a connectionist network to map from the orthography of about 3000 monosyllabic English words – both regular and exception – to their phonology via a set of hidden units (i.e., the bottom portion of the framework in figure 2.1, referred to as the *phonological* pathway). The network was also trained to use the same internal representation to regenerate the orthographic input, providing a means for the network of distinguishing words from nonwords based on the accuracy of this reconstruction. Orthographic input was coded in terms of context-sensitive letter triples, much like the highest-level representations in MORSEL. Phonological output was coded in terms of triples of phonemic features. To determine the network's pronunciation of a given letter string, an external procedure constructed the most likely phoneme string given the feature triples generated by the network. This string was then compared with the actual pronunciation of the stimulus to determine whether the network made a correct or error response. After training, the network pronounced correctly 97.7% of the words, including most exception words. The network also exhibited the standard empirical pattern of an interaction of frequency and consistency in naming latency (Andrews, 1982; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Taraban & McClelland, 1987; Waters & Seidenberg, 1985) if its real-valued accuracy in generating a response is taken as a proxy for response time (under the assumption that an imprecise phonological representation would be less effective at driving an articulatory system). However, the model was much worse than skilled readers at pronouncing orthographically legal nonwords and at lexical decision under some conditions (Besner, Twilley, McCann, & Seergobin, 1990). Thus, although highly successful in many respects, the model failed to refute traditional claims that localist, word-specific representations and separate mechanisms are necessary to account for skilled reading.

Plaut et al. (1996) showed, however, that the limitations of the Seidenberg and McClelland model stem not from any general limitation in the abilities of connectionist networks, but from its use of poorly structured orthographic and phonological representations. The triples-based orthographic and phonological representations used by the original model fail to capture the relevant similarities among written and spoken forms of words adequately, essentially because the contribution that each grapheme and phoneme makes is overly sensitive to the surrounding context. When more appropriately structured representations are used – based on graphemes and phonemes and embodying phonotactic and graphotactic constraints – network implementations of the phonological pathway can learn to pronounce regular words, exception words, and nonwords as well as skilled readers. Furthermore, the networks also exhibit the empirical frequency-by-consistency interaction pattern, even when naming latencies are modeled directly by the settling time of a recurrent, attractor network.

Although Plaut et al. (1996) demonstrated that implementations of the phonological pathway on its own can learn to pronounce words and nonwords as well as skilled readers, a central aspect of their general theory is that skilled reading more typically requires the combined support of both the semantic and phonological pathways (see also Hillis & Caramazza, 1991; Van Orden & Goldinger, 1994), and that individuals may differ in the relative competence of each pathway (Plaut, 1997; Seidenberg, 1992). The division-of-labor between these pathways has important implications for understanding acquired surface dyslexia, a neuropsychological disorder in which patients pronounce regular words and nonwords normally but “regularize” exception words, particularly those of low frequency (e.g., SEW read as SUE; see Patterson, Coltheart, & Marshall, 1985). Plaut et al. (1996) explored the possibility that surface dyslexia might reflect the natural limitations of an intact phonological pathway that had learned to rely on semantic support that was reduced or eliminated by brain damage. They approximated the contribution that the semantic pathway would make to oral reading by providing phonological representations with external input that pushed them toward the correct pronunciation of each word during training. A semantic impairment was modeled by weakening this external input. Plaut and colleagues found that, indeed, a phonological pathway trained in the context of support from semantics exhibited the central phenomena of surface dyslexia following semantic damage: intact nonword reading and regularization of low-frequency exception words (see Lambon-Ralph & Patterson, this volume). Moreover, as explored in additional simulations (Plaut, 1997), individual differences in the severity of surface dyslexia can arise, not only from differences in the amount of semantic damage, but also from *premorbid* differences in the division of labor between the semantic and phonological pathways.

The relative strengths of these pathways, and the overall competence of the reading system, would be expected to be influenced by a wide variety of factors, including the nature of reading instruction, the sophistication of preliterate phonological representations, relative experience in reading aloud versus silently, the computational resources (e.g., numbers of units and connections) devoted to each pathway, and the reader’s more general skill levels in visual pattern recognition and in spoken word comprehension and production. On this view, the more severe surface dyslexic patients had greater premorbid reliance on the semantic pathway as a result of one or more of these factors.

A remaining limitation of the Seidenberg and McClelland model that was not addressed by Plaut et al. (1996) concerns the ability of a distributed network lacking word-specific representations to perform lexical decision accurately. The focus of work with the Seidenberg and McClelland model was on demonstrating that, under some conditions, lexical decisions can be performed on the basis of a measure of orthographic familiarity. Plaut (1997) demonstrated that lexical decisions can be made more accurately when based on a familiarity measure applied to semantics. A feedforward network was trained to map from the orthographic representations of the 2,998 monosyllabic words in the Plaut et al. (1996) corpus to their phonological representations and to artificially created semantic representations generated to cluster around prototype patterns over 200 semantic features. After training, the network was tested for its ability to perform lexical decision based on semantic *stress* – an information-theoretic measure of the degree to which the states of semantic units differed from rest. When tested on the pronounceable nonwords from Seidenberg, Plaut, Petersen, McClelland, and McRae (1994), there was very little overlap between the semantic stress values for nonwords and those for words: an optimal decision criterion yielded only 1% errors. Moreover, the distributions of stress values for words varied systematically as a function of their frequency. In a second test, the network produced reliably higher semantic stress values – and thus poorer discrimination from words – for the Seidenberg, Petersen, MacDonald, and Plaut (1996) pseudohomophones compared with their controls. Thus, the network exhibited accurate lexical decision performance overall, along with an advantage for higher-frequency words and a disadvantage for pseudohomophones, as found in empirical studies.

More recently, Harm and Seidenberg (2004) have developed a full implementation of the “triangle” framework (see figure 2.1) and used it to examine a number of issues related to the division-of-labor in the reading system. Although the focus of the work is on the comprehension of written words via the direct versus phonologically mediated pathways, the underlying principles apply equally well to the computation of phonology both directly or via semantics. First, to approximate preliterate language experience, the network was trained to map bidirectionally between phonology and semantics for 6,103 monosyllabic words (see also Harm & Seidenberg, 1999, for a computational examination of the relevance of preliterate experience to reading acquisition). The phonology of each word was encoded in terms of eight slots of 25 phonetic features, organized into a CCCVVCCC template. In constructing semantic representations, words were first categorized by their most frequent word class (Francis & Kucera, 1982). For uninflected nouns and verbs, semantic features were generated using the WordNet online semantic database (Miller, 1990). Adjectives, adverbs and closed-class words were hand-coded according to preexisting feature taxonomies (e.g., Frawley, 1992). Inflected words were assigned the features of their base forms plus specific inflectional features. In total, 1,989 semantic features were generated to encode word meanings, with words averaging 7.6 features each (range 1–37). Once the preliterate network was reasonably accurate at understanding and producing spoken words (86% and 90% correct, respectively), the network was then trained on the reading task. Orthography was encoded using letter units organized into vowel-centered slot-based representation (analogous to phonology). After extended training, the model succeeded in activating the correct semantic features for 97.3% of the words and the correct phonological features for 99.2% of the words.

The trained model exhibited the appropriate effects of word frequency, spelling-sound consistency, and imageability in pronouncing words, and was as accurate as skilled readers in pronouncing pseudowords. Harm and Seidenberg’s (2004) primary goal, however, was to address the longstanding debate on whether reading is necessarily phonologically mediated. An examination of the division-of-labor in activating meaning from print over the course of training indicated that the network relied heavily on phonological mediation (orthography-phonology-semantics) in the early stages of reading acquisition but gradually shifted towards increased reliance on the direct mapping (orthography-semantics) as reading skill improved. Even at the end of training, however, both pathways continue to make important contributions to performance. This is especially true for homophones (e.g., ATE, EIGHT), which cannot be comprehended solely by the mediated pathway. Harm and Seidenberg demonstrate that the model's performance with homophones matches the findings from a number of empirical studies (Jared & Seidenberg, 1991; Lesch & Pollatsek, 1993; Van Orden, 1987; see also Van Orden & Kloos, this volume).

Conclusion

Connectionist models instantiate a set of computational principles that are intended to approximate the core properties of neural computation. Early efforts to apply these models to reading employed localist representations for words and hand-specified connection weights. More recent efforts have focused on learning internal distributed representations that effectively mediate the interaction of orthographic, phonological, and semantic information. Because such systems lack word-specific representations and separate pathways for regular versus irregular items, they stand in sharp contrast to traditional dual-route theories of word reading. Existing models are still limited in the size and diversity of the vocabulary they handle and the range of empirical issues they address. Nonetheless, these systems illustrate how a common computational framework can provide insight into reading acquisition, normal skilled reading, patterns of reading impairment following brain damage, and even possible approaches to remediation of developmental (Harm, McCandliss, & Seidenberg, 2003) and acquired (Plaut, 1996) deficits.

Note

The preparation of this chapter was supported by NIH grant MH55628.

3

Visual Word Recognition: Theories and Findings

Stephen J. Lupker

The topic of "visual word recognition" may have the largest literature in Cognitive Psychology and, therefore, a chapter on the topic must be selective. This chapter will first place the relevant issues in a historical context and then review the basic visual word recognition phenomena within the context of current models. It will then be argued that any successful model of visual word recognition needs to incorporate the assumption of "interactivity," that is, that the various components of the visual word recognition system (i.e., orthographic, phonological, semantic) mutually activate and inhibit each other while a word is being processed (see also Van Orden & Kloos, this volume). (Hereafter, the term "word recognition" will be used as shorthand for the term "visual word recognition.")

What is "word recognition"? At least until the appearance of Seidenberg and McClelland's (1989) connectionist model of reading, word recognition was typically thought of as the process of going from a printed letter string to the selection of a single item stored in lexical memory. Lexical memory, or the "lexicon," is a mental dictionary containing entries for all the words a reader knows. Thus, word recognition was essentially synonymous with the terms "lexical access" or "lexical selection." Such a definition, of course, assumes that words are represented as lexical entries in memory. Seidenberg and McClelland's model explicitly denied the existence of such representations, arguing instead that representations were distributed across sets of simple subsymbolic processing units. To the extent that models of this sort have been successful, they have forced theorists to contemplate the possibility that some of the standard assumptions about the architecture of the word recognition system should be altered.

What appears to be an equally important aspect of Seidenberg and McClelland's (1989) model was that it contained a straightforward outline for how semantics should be integrated into the word recognition system. That is, semantic information was assumed to be represented no differently than other types of information (i.e., orthographic and phonological) and all of these mental representations were assumed to follow the same rules of activation. As such, this model represented what I would argue was the first complete model of word recognition. This is a crucial point because, as will be argued in this chapter (see also Balota, Ferraro, & Connor, 1991), any successful model of word recognition will need to have a mechanism for explaining the impact of semantics, both the impact of the semantic context within which a word is processed and the impact of the semantic attributes of the word itself (Whaley, 1978).

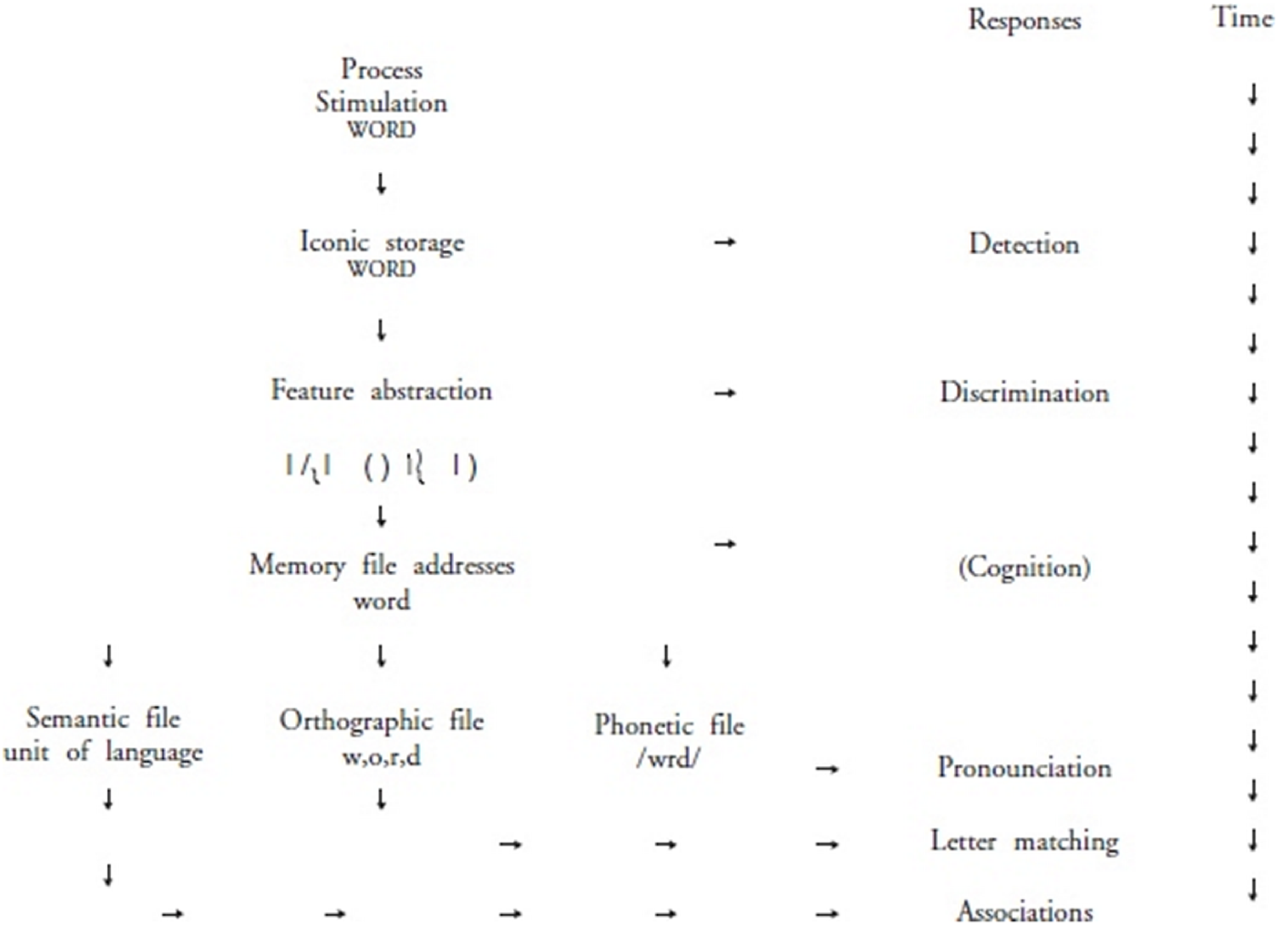
Historical Context

Most of the early models of word recognition (e.g., Gough, 1972; Massaro, 1975; Morton, 1969; Smith & Spoehr, 1974; Theios & Muise, 1977) relied on two assumptions. First, the human information processing system involves a series of processing stages that work in a serial, nonoverlapping fashion. Information only flows one way, that is, forward, through the system and, further, each stage is essentially completed before the next begins. The term "thresholded" is used to refer to the assumption that each stage must be completed before the next one can begin. The idea is that a stage is ready to pass information on to the next stage only when the activation at the initial stage reaches a threshold. In contrast, models proposing that information passes between stages as soon as information at one stage begins to be activated are referred to as "cascaded" (McClelland, 1979). The second assumption was that the word recognition system is a fairly autonomous system, that is, it works only with the information stored within it, in particular, the information that can be referred to as lexical information (Forster, 1981). (Theios & Muise's, 1977, model, contained in figure 3.1, is a typical example of this type of model.)

At the risk of overgeneralizing, these models proposed that there is initially a perceptually based process that leads to the activation of sublexical units (typically letter units). The activation of these sublexical units allows the formation of some sort of "prelexical" code. This code activates those word (i.e., lexical) units that are more or less consistent with it. Ultimately, one of these units is selected or accessed. Only at that point does meaning start to become activated. The specific assumption that meaning activation strictly follows lexical selection is referred to as the "form-first" assumption (Forster & Hector, 2002).

One major problem that the early models faced was explaining why there often seemed to be observable effects of "higher-level" information on "lower-level" processing. The classic example is the word superiority effect (Reicher, 1969; Wheeler, 1970). The word superiority effect refers to the fact that letters (i.e., lower-level information) are more accurately reported when presented in words than when presented in nonwords. The experimental task involves the rapid presentation of a letter string often followed by a mask in order to make perception difficult. One letter position is cued for report. To prevent guessing from differentially influencing responding, two alternatives are presented for the identity of the cued letter on each trial. If the letter string had been a word, both alternatives would create a word (e.g., if the word had been WORD and the final position had been cued for report, the alternatives might be D and K). If the letter string had been a nonword, both alternatives would create a nonword (e.g., VCRD with D and K as alternatives for the final position). The standard result is better performance in the word condition than in the nonword condition (e.g., Johnston & McClelland, 1973; Maris, 2002; Paap, Chun, & Vonnahme, 1999; Paap, Newsome, McDonald, & Schvaneveldt, 1982).

Figure 3.1 Model of word recognition (Theios & Muise, 1977).



The problem for models based on the principles of autonomy and thresholded processing is obvious. How can the existence of a mental representation for a word (e.g., a lexical unit) influence the processing of letter information if that mental representation itself is not accessed until the identity of the letter in question is known? Do changes have to be made to the functional architecture of the models to explain these findings, or is it only necessary to change a single assumption of the model? Alternatively, can these effects be explained in terms of some process (e.g., decision) not actually described by the model itself? It now seems clear that it was the impetus provided by these types of questions that led to the explosion in word recognition research witnessed since the early 1970s. (For a discussion of these issues in auditory word recognition, see Norris, Lewis, & Rubenstein, 1971a; Stanners & Forbach, 1973), the pseudohomophone effect (Coltheart et al., 1977; Dennis, Besner, & Davelaar, 1985), and the nonword neighborhood size effect (Coltheart et al., 1977). While an argument can be made that it is precisely when the word recognition system fails that we can learn most about it, these types of effects seem to have more to say about task specific processes, in this case, in lexical decision, than about the word recognition process per se.

The Models

Search models

The bin model. Search models best represent the way in which one can build a model based on the assumption of thresholded, autonomous processing. According to search models, readers recognize a word by comparing a prelexical code against a set of lexical codes until a match is obtained. The search is not through all of lexical memory but rather, some process designates a section of lexical memory as the optimal search area and the search is confined there. The model that best exemplifies this idea is Forster's bin model (1976; 1989).

According to Forster's (1976) model, the lexical system involves three peripheral access files and a master file, each containing information about all the words in our lexicon. The three peripheral files are orthographically-, phonologically- and semantically-based and each serves as a means of getting to word entries in the master file where all the information about the word is contained. It is relevant to visual word recognition to focus on the orthographic file in which each word in our lexicon contains an entry (this is also true for the other two peripheral files). In each entry in the orthographic file are two things, an "orthographic access code," which is a description of the orthographic properties of the word, and a pointer to the location for that word in the master file.

When a word is viewed, a perceptual process turns that word into a prelexical code that is format compatible with the access codes in the orthographic file. The orthographic file is then searched by comparing the prelexical code with the orthographic access codes. As noted, this search is constrained to a section of the orthographic file. In particular, the orthographic file is organized into bins that contain similar orthographic access codes. So, for example, the words CAT and CAN would probably be in the same bin. In essence, the search is constrained to the bin that is most likely to contain the word being viewed.

The idea of bins may be better understood by drawing a partial parallel to looking up a word in a dictionary. When using a dictionary, one checks the words at the top of each page and only looks at the individual items on the page if it is a page that the word is likely to be on (e.g., the word COMET is virtually certain to be on the page with the heading COMBO–COMFORT). Each bin is like a page in the dictionary and the reader goes directly to the bin most likely to contain the word being viewed. The parallel is not perfect, however, because the words in the bin are not ordered alphabetically, as they are on a dictionary page, but in descending order of frequency. Thus, the entries in the bin are searched in descending order of frequency.

If the search through the designated bin turns up a close match with one of the entries, the location of this entry is flagged while the search continues, looking for other close matches. If a match is close enough, the entry is opened and the pointer to the master file is used to access the word's entry in that file. This process engages a second analysis, referred to as "post-access check," which compares the properties of the stimulus with the properties of the word in the master file. If this comparison is successful, the word has been successfully recognized. Note also that if none of the words in the bin are successfully recognized in the initial search, close matches that had been flagged but not had their entries opened are then evaluated (Forster, 1989; Forster, Mohan, & Hector, 2003).

In terms of the four basic phenomena, the model has no difficulty explaining the frequency effect and the masked repetition priming effect. The more rapid processing of high-frequency words follows directly from the fact that the bins are searched in descending order of frequency. Masked repetition priming arises because the prime begins the word recognition process and, if the target is a repetition of the prime, its processing has a head start. In particular, it is assumed that the prime begins to open the correct entry in the orthographic file. Thus, the entry opening time for the target is shortened, producing more rapid processing. In contrast, the model does not have any obvious way of explaining the word superiority effect.

The other phenomenon, semantic priming, can be explained in terms of cross-referencing in the master file, at least according to the original version of the model (Forster, 1976). Entries for semantically related words are directly linked in the master file. Thus, after the prime DOG has been recognized, the CAT entry in the master file can be easily accessed. As a result, the post-access check of the properties for CAT against the properties of the stimulus can be started without bothering with the search process.

This proposal concerning the (limited) impact of semantics on the word recognition process has a number of implications. One is that, because semantically primed words do not engage the search process, there should be no frequency effect for those words. In fact, Becker (1979) has demonstrated that the frequency effect is smaller when words are semantically primed. A second implication is that semantic priming effects should only exist when the prime's entry is successfully accessed in the master file. Thus, semantic priming effects from primes that are masked in order to avoid recognition (e.g., Carr, McCauley, Sperber, & Parmelee, 1982; Fischler & Goodman, 1978; Hines, Czerwinski, Sawyer, & Dwyer, 1986; Marcel, 1983) are problematic for the model. Finally, because the only impact of semantics on lexical processing is due to the structure of the master file, the model cannot explain any effects of semantics on word recognition with the exception of semantic priming effects. As will be discussed subsequently, there are a number of such effects.

The activation-verification model. Paap et al.'s (1982) activation-verification model (see also, Paap, Johansen, Chun, & Vonnahme, 2000) is also a search model; however, it differs from Forster's (1976) in that, although it is an autonomous model, it invokes cascaded processing. In the model, first letter units and then word units are activated in a serial, but cascaded, fashion (so that information passes through the system before initial processing is complete). Letter activation occurs in position-specific channels and is conceptualized as a feature matching process. Thus, there is some probability that an incorrect but featurally similar letter will be activated at each letter position. Activity at the letter level continuously feeds into the lexicon with the activation of any lexical unit being a function of the activity levels of that word's constituent letters.

It is the activity levels in the lexicon that determine which set of word candidates is selected for further processing. The nature of that further processing is crucially dependent on whether the reader has also been able to establish a "refined perceptual representation of the word" (Paap et al., 1982, p. 574), which is the situation in normal reading. In this case, the set of candidates is serially verified against the perceptual representation (this is the search process). If there is a sufficient match between a candidate and the perceptual representation at any point, the candidate is accepted and the verification process is terminated. As in Forster's (1976) model, the verification process is frequency-based (higher-frequency words are verified first). Further, if there is a semantic context (i.e., a prime), words semantically related to the prime will enter the candidate set and be verified first. If a refined perceptual representation cannot be established, as in perceptual identification tasks, it is not possible to carry out the verification process. Thus, a probabilistic selection is made from among the candidates based on the activation levels of those candidates.

In terms of the four basic phenomena, this model has had its greatest success explaining the word superiority effect (see Paap et al., 1982). The model can also explain both frequency effects and semantic priming effects. Frequency effects arise due to the serial, frequency-based verification process, whereas semantic priming effects are due to the inclusion of semantically related words in the candidate set. Masked repetition priming effects are more problematic for the model. Presenting a masked prime that is identical to the target will activate target representations at both the letter level and the lexical level. The most important determinant of processing speed, however, is the search process, which only begins once the candidate set has been established. The prime's activation of the target's representations may increase the probability that the target will be in the candidate set; however, it should not change its position in the search order. Hence, unless additional assumptions are added, the prediction would be that a masked repetition prime would not have any effect on target processing.

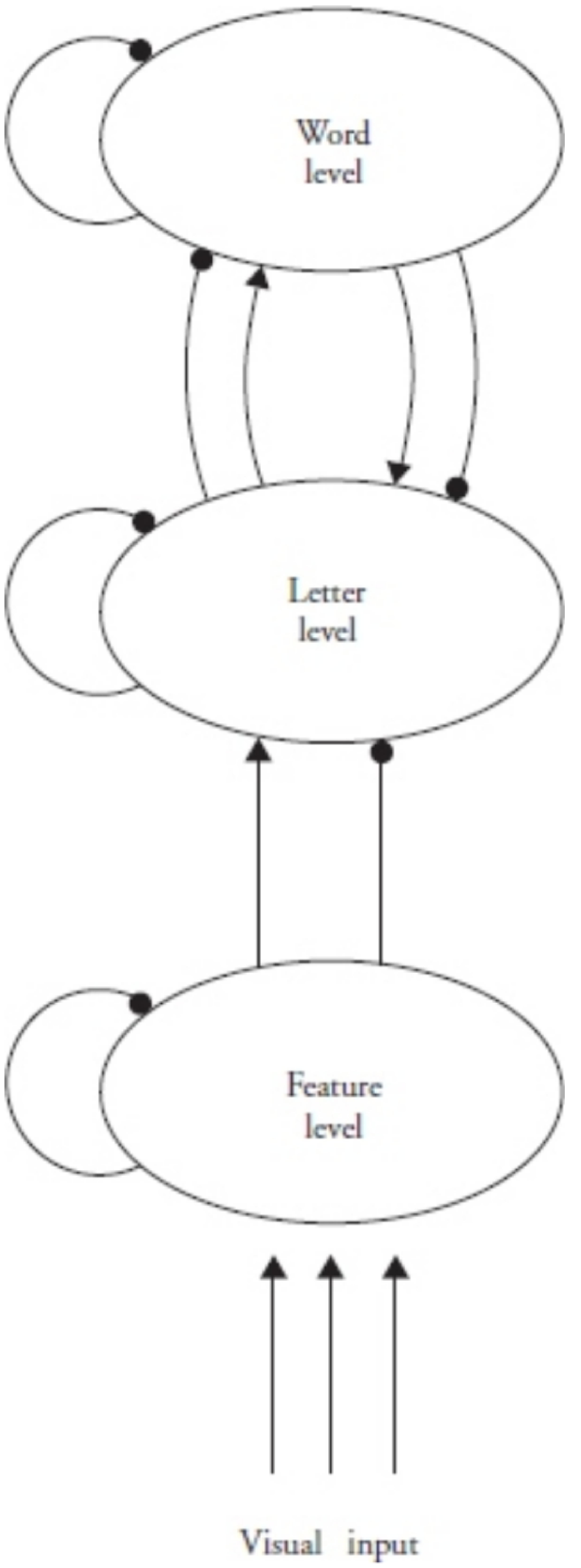
Two additional points should be made about the model. First, in data-limited tasks, tasks in which it is hard to establish a refined perceptual representation, the verification process cannot be carried out. Thus, there is no mechanism for producing a frequency effect. Indeed, there does not appear to be much evidence for frequency effects in word superiority effect experiments (Manelis, 1977; Paap & Johansen, 1994; Paap & Newsome, 1980). Second, as is true of Forster's model (1976), this model has no means of explaining any semantic effects other than semantic priming effects.

Activation models

The interactive activation model. Activation models represent the other end of the continuum from the search models in terms of cascaded and autonomous processing. The preeminent activation model is McClelland and Rumelhart's (1981) interactive activation model. This model represents the first real implementation of activation and inhibition processes. It also forms the core of a number of other models in the literature (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Grainger and Jacobs, 1996).

The interactive activation model was specifically intended to be a model that would explain the effects of higher-level information on lower-level processing, in particular, the word superiority effect. In the model, there are three levels of representation: feature, letter, and word. When processing begins, there is a continuous flow of activation upstream from feature-level representations to letter-level representations to word-level representations, as well as downstream from word-level representations back to lower-level representations ("feedback activation"). There is also a flow of inhibition between representations at the same level. Lexical selection is achieved when the activation in a lexical representation exceeds a threshold. (See figure 3.2 for a graphic description of the interactive activation model.).

Figure 3.2 Interactive activation model (McClelland & Rumelhart, 1981).



As McClelland and Rumelhart (1981) argue (see also Rumelhart & McClelland, 1982), this type of system can readily account for the impact of higher-level representations on lower-level representations and, hence, it can explain the word superiority effect. It can also explain frequency effects due to the fact that the resting level activations of word-level representations are frequency dependent. Thus, once activated, representations for high-frequency words will reach their activation threshold more quickly than representations for low-frequency words. Masked repetition priming effects would be explained in terms of the residual the triangle representation (see figure 3.3). The appropriate connections between sets of units have to be learned, just as a young reader must learn to read. Within the model, learning is essentially an error correction process. When presented with a word, the units at all levels begin to activate (and inhibit) each other, resulting in a pattern of activation across all the units. These activation patterns, which initially will be quite inaccurate, are compared with the correct patterns and then weights between units are adjusted in order to make processing more accurate the next time. This process continues with each new exposure to the word. As a result, over time, activation in one set of units comes to produce the appropriate activation in the units in the other pools (e.g., orthographic processing of the visually presented word CAT allows the activation of the phonological units for the phoneme sequence [kat]). In addition, as shown in figure 3.3, these models also incorporate "hidden units." These units help define the relationships between units in the main pools (for an explanation of why hidden units are necessary see Hinton, McClelland, & Rumelhart, 1986).

The nature and number of representations in each domain (orthographic, phonological, semantic) are often model specific. For example, in the Seidenberg and McClelland (1989) model, the orthographic units are not intended to represent psychologically real concepts; nonetheless, the pattern across units does give rise to sets of letter triples (e.g., MAK). Due to this fact, although the model had considerable success in explaining word recognition data, it also had some serious limitations, particularly in its ability to name nonwords (Besner, Twilley, McCann, & Seergobin, 1990; Fera & Besner, 1992; Coltheart, Curtis, Atkins, & Haller, 1993). In contrast, the orthographic units in Plaut, McClelland, Seidenberg, and Patterson's (1996) model directly represent either letters or letter combinations, allowing some of the problems noted by Besner and colleagues to be fixed. Semantic units, which, until recently (e.g., Harm & Seidenberg, 2004), have played a smaller role in the models' development are typically assumed to represent semantic features (Hinton & Shallice, 1991; Plaut et al., 1996; Plaut & Shallice, 1993), even though the features themselves are often not specified (e.g., Masson, 1991; although see Cree, McRae, & McNorgan, 1999; Harm & Seidenberg, 2004; and McRae, Seidenberg, & de Sa, 1997).

In terms of the four basic phenomena, the model can clearly account for frequency effects. Indeed, the frequency of exposure to a word is the main determinant of the values of the connection weights for the word's units. Further, because the model is an activation model, it should also be able to explain masked repetition priming. That is, the briefly presented masked prime would activate some of the prime's units, allowing for more rapid processing of the target if that target is a repetition of the prime. With respect to semantic priming, modeling work by Masson (1991, 1995), Cree et al. (1999), Plaut (1995b), Plaut and Booth (2000), and McRae et al. (1997) has demonstrated that these types of models can produce semantic priming effects, at least when the concepts share semantic features. In particular, according to these models, the processing of related targets (e.g., DOG following the prime CAT) is easier because the two concepts share semantic units. As a result, some of the semantic units for the target will have already been activated when the target appears, speeding target processing. Notice that the interactivity inherent in these models, specifically, the feedback processes presumed to occur between the semantic units and the "lower-level" orthographic and phonological units, plays essentially no role in this explanation. As will be discussed below, explanations in which these feedback processes do play an important role may provide an even better way of explaining semantic priming effects.

Explaining the word superiority effect is more of a challenge for the model. It would seem like the feedback processes at work in the model should, as with the interactive activation model, produce a word superiority effect. A key distinction here, however, is that it is the feedback from word units in the interactive activation model that produces the word superiority effect. Those units do not exist in PDP models. However, as Plaut et al. (1996) note, units in network models tend to organize themselves into stable patterns called "attractors." These patterns of units that group together to become attractor units may function somewhat similarly to word units, providing the necessary feedback. Thus, it is possible that, with the correct assumptions, the model could also explain the word superiority effect.

The various models discussed above all have their strengths and weaknesses based on their ability to account for the basic phenomena. At present, there is no clear winner. The proposition to be argued in the remainder of this chapter, however, is that other evidence indicates that any successful model of word recognition will need to assume that there is an interactive flow of activation among the processing structures (see also Stone & Van Orden, 1994; Taft & van Graan, 1998; Van Orden & Goldinger, 1994). That is, not only does activation flow forward from activated units to other sets of units, but also, once those units start to become activated, they send activation back to the appropriate units at other levels (see Van Orden & Kloos, this volume). The term "feedback" is used to refer to the flow of activation back to units that were initially activated (i.e., the orthographic units when the word CAT is read). As noted, the interactivity notion is embodied to various degrees by the activation models and is a direct contradiction of the autonomy assumptions that tend to characterize the search models. The framework for this discussion will be the triangle framework, which is most clearly an attribute of the PDP models. In theory, it would be possible to refer instead to a model like Coltheart et al.'s (2001), as it also has the triangle structure embedded within it (i.e., with its orthographic input lexicon, phonological output lexicon, and semantic system). However, the units in the first two of these systems are lexical, while some of the effects to be discussed below (e.g., Stone, Van Hoy, & Van Orden, 1997) are based on sublexical units, making it more difficult to see how those effects fit within this framework.

The Orthographic-Phonological Interaction

Feedback from phonology to orthography

In visual word recognition tasks, the units initially activated are the orthographic units. Thus, evidence for feedback activation would come from experiments demonstrating that phonological activation affects the activation in those orthographic units. For example, Stone et al. (1997) and Ziegler, Montant, and Jacobs (1997) have shown that words that have multiple possible mappings from phonology to orthography (i.e., words like GAIN, which could have been spelled GANE) produce longer lexical decision latencies than words like TENT which have only one possible mapping (see also Perry, 2003). Words like GAIN are referred to as "feedback inconsistent" because the mapping from phonological units to orthographic units are one-to-many. Words having one-to-one mappings, between phonology and orthography, like TENT, are referred to as "feedback consistent." The explanation for these findings is that, with inconsistent words, feedback slows the activation of the correct orthographic code because at least some of that feedback is misdirected to the incorrect orthographic code (e.g., ANE), creating competition.

The effects reported by Stone et al. (1997) and Ziegler et al. (1997) were not large and their reliability has been challenged by Peereman, Content, and Bonin (1998). One could argue, however, that the reason the effects were small was because the manipulations were weak. The feedback directed to the incorrect orthography (i.e., ANE) does not activate a strong competitor for GAIN because neither GANE nor ANE are words and, hence, neither is strongly represented within the orthographic units. The use of homophones allows for a much stronger manipulation. Homophones are words that have different spellings but the same pronunciation (e.g., PAIN and PANE). According to a feedback account, if either is presented visually, the activation of /pAn/ would lead to activation being fed back to the orthographic codes for both PAIN and PANE. The result should be strong competition, leading to a delay in responding. That is, there should be a homophone disadvantage in tasks based on orthographic processing (e.g., lexical decision).

The available data are firmly supportive of this prediction. Rubenstein et al. (1971a) were the first to report a homophone disadvantage in lexical decision in comparison to a control condition. While there was considerable controversy about this finding (e.g., Clark, 1973) and some failures to replicate (e.g., Coltheart et al., 1977), more recently the pattern has emerged very clearly (Davelaar, Coltheart, Besner, & Jonnasson, 1978; Pexman & Lupker, 1999; Pexman, Lupker, & Jared, 2001; Pexman, Lupker, & Reggin, 2002).

Early accounts of homophone effects were based on the idea that visual word recognition was phonologically mediated. Part of the reason was that, originally, these effects were only found when processing the lower-frequency member of the homophone pair (e.g., PANE). The explanation was that both PAIN and PANE activated the phonological code /pAn/ and it then led to the selection of the lexical unit for the higher-frequency member of the homophone pair (i.e., PAIN). Further processing allowed the discrepancy to be noted, at which point the lexical selection process was restarted. The result, of course, was longer latencies for low-frequency homophones. The Pexman et al. (2001) paper is especially important in this regard. Here the nonwords in the lexical decision task were pseudohomophones (nonwords that sound like words when pronounced; e.g., BRANE). These nonwords produced longer word latencies and not only did the homophone effect increase for the low-frequency words, but there was also a significant homophone effect for the high-frequency words. This should never happen if the homophone effect were due to selecting the higher-frequency member of the pair first in lexical search because that event should not be altered by changing the type of nonword being used. In contrast, this result is quite consistent with the claim that these effects are feedback effects.

Feedback from orthography to phonology

The key issue for word recognition is the impact of phonology on orthographic processing. However, for completeness, it is important to discuss the impact of orthography on phonological processing. Research directly relevant to this issue is fairly extensive (e.g., Borowsky, Owen, & Fonos, 1999; Dijkstra, Roelofs, & Fieuws, 1995; Ziegler, Muneaux, & Grainger, 2003); I will focus on two of the earlier papers. A key finding suggesting that orthographic feedback has an impact on speech perception was reported by Seidenberg and Tanenhaus (1979) (see also Donnenwerth-Nolan, Tanenhaus, & Seidenberg, 1981). Seidenberg and Tanenhaus presented participants with a cue word (typically auditorily) followed by auditorily presented target words. The subject's task was to respond as soon as one of the target words rhymed with the prime. Although accurate performance in this task must be based on an evaluation of the phonological code, there was a clear impact of the orthographic relationship between the cue and target. That is, participants were much faster to respond when the two words were spelled the same (e.g., GREED–DEED) than when they were not (e.g., BEAD–DEED).

This effect indicates that orthographic information is automatically activated when a spoken word is heard. It also indicates that orthographic information plays a role in the processing of subsequently presented spoken words. Although it might be the case that participants evaluate the spelling of the words in these experiments in spite of the fact that they are explicitly told to do something else, a more reasonable explanation is that orthographic information was automatically activated when the cue word was processed, and it fed back to the phonological codes for similarly spelled words. Thus, those words were more activated and, hence, easier to process.

A second finding suggesting the impact of orthographic feedback on speech perception was reported by Ziegler and Ferrand (1998). As those authors note, effects like those reported by Seidenberg and Tanenhaus's (1979) derive from the processing of an initial stimulus. Many strategies are available to participants in such a situation, allowing a number of alternative explanations for the findings. Ziegler and Ferrand investigated an on-line effect using an auditory lexical decision task. The key variable was whether the target word had only one or multiple possible spellings. That is, as before, because the word TENT has only one way that it could possibly be spelled, its phonology-orthography mapping is referred to as "consistent" while the word GAIN, which could have been spelled GANE, has an "inconsistent" phonology-orthography mapping. The results showed more rapid latencies for consistent words than for inconsistent words. The explanation offered is that words like GAIN, when presented auditorily, activate incorrect orthographies (e.g., GANE) reducing the support the phonological code /gAn/ receives through feedback activation. Thus, its activation is slowed.

Interactions with Semantics

Facilitative effects of feedback

To provide a satisfactory explanation of any effect, there needs to be at least an implicit assumption about how an experimental task is performed. More specifically, it is necessary to take a position on what units are important in each task. The following discussion will focus on interactions involving semantic and both orthographic and phonological units. We have argued that the interaction between semantics and orthography manifests itself in effects in lexical decision (Hino & Lupker, 1996; Pexman & Lupker, 1999), whereas the interaction between semantics and phonology manifests itself in effects in naming (e.g., Hino & Lupker, 1996; Pexman, Lupker, & Reggin, 2002). In short, the process of making a lexical decision is driven mainly by activity within the orthographic units, while the naming task is mainly based on activity within the phonological units.

The first effect to be discussed is the ambiguity effect in lexical decision. The standard finding is that words with more than one meaning (e.g., BANK) have shorter latencies than words with a single meaning (e.g., EVENT – Borowsky & Masson, 1996; Hino & Lupker, 1996; Hino, Lupker, & Pexman, 2002; Hino, In the other case, the argument has been that a single phonological representation feeds activation to two sets of orthographic units, producing competition and, hence, a processing cost. Both of these predictions are based on the nature of the links between units and the presumed nature of the processing required for the task (e.g., an evaluation of orthographic codes in lexical decision). However, there is nothing special about these particular links. Any linkages between units allow for an analysis and predictions.

The type of relationship between orthography and phonology that produces a homophone effect (i.e., two sets of orthographic units are linked to one set of phonological units) has a parallel in the relationship between orthography and semantics and in the relationship between phonology and semantics. In particular, when considering words that have synonyms, there are two sets of orthographic (phonological) units being mapped into one set of semantic units. Thus, when that set of semantic units is activated, it will feed activation back not only to the correct set of orthographic (phonological) units but also to the set of orthographic (phonological) units appropriate to the synonym, producing a competition. The prediction is that there should be a processing cost in both lexical decision and naming. The results in both Dutch (Pecher, 2001) and Japanese Katakana (Hino et al., 2002) support this prediction.

Two Other Emerging Issues

Representing ambiguous words

As Joordens and Besner (1994) note, models based on distributed representations make a clear prediction about the semantic processing of ambiguous words. They predict that there will be competition at the semantic level between the sets of units for the different meanings, producing a processing cost. Thus, ambiguous words should be more difficult to process than unambiguous words, completely the opposite of what is typically reported in lexical decision and naming experiments (Borowsky & Masson, 1996; Gottlob et al., 1999; Hino & Lupker, 1996; Hino et al., 1998; Hino et al., 2002; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas et al., 1988; Lichacz et al., 1999; Millis & Button, 1989; Pexman & Lupker, 1999; Rubenstein et al., 1970; Rubenstein et al., 1971b). When considering just these two tasks, a feedback explanation within a PDP framework gets around the problem if one assumes that lexical decision responses are based on activity at the orthographic level and naming responses are based on activity at the phonological level. Thus, semantic level competition has little impact in either task (see also Borowsky & Masson, 1996, and Kawamoto, Farrar, & Kello, 1994, for other ways of addressing this problem). The question still lingers, however, as to whether there is any behavioural evidence for this rather key prediction of PDP models.

There are now three sets of results in the literature supporting this prediction. First, Rayner and colleagues (e.g., Duffy, Morris, & Rayner, 1988; Rayner & Duffy, 1986) reported that in some circumstances, ambiguous words receive longer fixations. Second, Gottlob et al. (1999) and Piercey and Joordens (2000) have reported an ambiguity disadvantage in a relatedness-judgment task in that subjects found it more difficult to decide that a word (e.g., MONEY) was related to an ambiguous word (e.g., BANK) than to an unambiguous word (e.g., DOLLAR). Finally, Hino et al. (2002) have shown that ambiguous words take longer to classify (on negative trials) in a semantic categorization task (i.e., BANK – is it a living thing?).

On closer inspection, however, this evidence appears to be quite weak. As Duffy et al. (1988) note, their results are perfectly compatible with a decision-based explanation. That is, it is possible that all meanings of the ambiguous word are activated simultaneously (and without competition) and that the delay in gaze duration is due to the time taken to select the intended meaning for the sentence. In a similar vein, the effects reported by Gottlob et al. (1999) and Piercey and Joordens (2000) can be explained by assuming that all meanings of the ambiguous words are activated without competition but there is then a competition between response tendencies. When the stimulus is MONEY–BANK, the financial meaning of BANK produces a drive to respond “yes” (the correct response), while the river meaning of BANK produces a drive to respond “no.” Indeed, recent work by Pexman, Hino, and Lupker (2004) has shown that when there is no response competition, there is no effect. That is, when the correct response is “no” (e.g., TREE–BANK vs. TREE–DOLLAR), there is no ambiguity disadvantage. Finally, Hino et al.’s (2002) effect in the semantic categorization task has been shown to be category dependent. That is, it arises when the category is broad (e.g., living things) but not when the category is narrow (e.g., animals, vegetables) (Forster, 1999; Hino, Lupker, & Pexman, 2001). These results also point toward a decision-based, rather than a semantic-processing, explanation.

One issue that might be relevant here is that there are essentially two types of ambiguous words (see Klein & Murphy, 2001, 2002). One type is words that have completely unrelated meanings; for example, BANK. The fact that this word has at least two meanings – ‘a place where you keep your money and the edge of a river’ is an accident of history. These types of words are called *homonyms* and, presumably, the various meanings are represented separately in semantic memory. There are also words that have multiple senses that have evolved from a single meaning; for example, the word BELT. The fact that this word means ‘the strip of material that you have around your waist’, ‘a thin area of land’, ‘a hard blow’, and ‘a drink of an alcoholic beverage’ is not an accident of history. These senses all evolved from the basic meaning of BELT. These types of words are called *polysemous* and the different senses may not be represented separately in semantic memory. If so, there could be different processing implications for the two word types; in particular, only homonyms may produce any real competition during semantic processing.

Klein and Murphy (2001, 2002) examined the question of how different senses of a word are represented in memory using their "sensicality judgement" task. In this task subjects see adjective-noun pairs, like SHREDDED PAPER, and their task is to decide whether this combination of words makes sense. Klein and Murphy reported that participants were much faster and more accurate in responding to the second of two adjacent word pairs when the two pairs tapped the same sense of the polysemous noun (e.g., WRAPPING PAPER followed by SHREDDED PAPER) than when the two pairs tapped different senses of the polysemous noun (e.g., DAILY PAPER followed by SHREDDED PAPER). A similar pattern emerged when they considered homonyms. That is, participants were faster and more accurate in responding to the second word pair when the pairs tapped the same meaning of the noun (e.g., COMMERCIAL BANK followed by SAVINGS BANK) than when the pairs tapped different meanings of the noun (e.g., CREEK BANK followed by SAVINGS BANK). Based on these results, Klein and Murphy suggested that the separate senses of polysemous words appear to be represented separately in semantic memory and, in fact, they are represented essentially the same way that the separate meanings of homonyms are.

In contrast, Azuma and Van Orden (1997) and Rodd, Gaskell, and Marslen-Wilson (2002) have argued that there is an important distinction between homonyms and polysemous words that has processing implications for lexical decision tasks. Rodd et al. selected a set of words in which the number of senses (few vs. many) varied orthogonally with the number of meanings (one vs. more than one). Although words with multiple senses showed shorter latencies than words with few senses (in line with previous research), words with multiple meanings produced slightly longer latencies than words with one meaning, although this effect was quite small and nonsignificant. Further, these results only emerged when the nonwords were pseudohomophones (as did the relevant results in Azuma and Van Orden's, 1997, experiments). If a number of assumptions are made about how participants perform the lexical decision task, Rodd et al.'s inhibitory effect of multiple meanings supports the PDP model prediction about the semantic processing of ambiguous words. However, their findings have not yet proved replicable (Hino & Lupker, 2003; Hino et al., 2001; Pexman, Hino, & Lupker, 2002) and there appears to be no other example of an ambiguity disadvantage of this sort ever reported in the literature.

Prelexical coding

So far, little has been said about the nature of the prelexical code that is used to access the core components of the word recognition system. Indeed, research on this issue is sparse. In all of the models that exist as simulations, assumptions about this code have had to be made. Even in those situations, however, the assumptions have been driven mainly by modeling convenience. Nonetheless, as Andrews (1997) has argued, these assumptions are important. For example, most models are based on the idea that this code allows at least partial activation of all word units in which the orthography is similar to what is contained in this code. In order to produce legitimate simulations of the word recognition process, it is necessary to specify which set of words is activated in this fashion and to what degree. A second, more concrete, example of why this is important is the fact that it was a change in the assumptions about the orthographic codes that the prelexical code contacts that allowed Plaut et al.'s (1996) model to account for the nonword naming data that Seidenberg and McClelland's (1989) model could not. The nature of the prelexical code is, of course, constrained by the nature of the orthographic code because the only purpose of the former is to activate the latter.

Most of the standard models of word recognition (e.g., Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981; Paap et al., 1982) assume a "channel specific" coding scheme for the prelexical code. That is, each letter in a word is immediately assigned to a channel and then identified within that channel. So, when SALT is being read, there would be some activation of word units for HALT, MALT, WALT, SILT, and SALE, those words overlapping in three of the four channels, but much less, if any, activation of SENT, SLAT and SAT. However, extant evidence, suggests that this assumption is incorrect. Humphreys, Evett, and Quinlan (1990), for example, have shown that shorter letter strings can prime longer words (e.g., oitk–WHITE) in a masked prime, perceptual identification task (see also Perea & Carreiras, 1998; and de Moor & Brysbaert, 2000). Such effects have forced researchers, more recently, to adopt "relative-position" coding schemes (e.g., Coltheart et al., 2001); however, the problem does not appear to be fully solved even with those schemes. For example, data suggest that letter strings containing transposed letters (i.e., SALT–SLAT) are actually more similar to one another than letter strings matching in N-1 positions (e.g., SALT–HALT). For example, transposed letter nonwords (e.g., JUGDE) are harder to reject than one-letter different nonwords (e.g., JUDPE) or control nonwords (e.g., SLINT) in lexical decision tasks (Andrews, 1996; Chambers, 1979; Holmes & Ng, 1993). In addition, Forster et al. (1987) showed that masked priming effects for transposed letter primes (e.g., anwser–ANSWER) were as large as those for repetition primes (e.g., answer–ANSWER) and larger than those for one-letter different primes (e.g., antwer–ANSWER). Finally, Perea and Lupker (2003) have reported significant facilitation in a masked semantic priming experiment with transposed letter primes (e.g., jugde–COURT) and little, if any, priming with one letter different primes (e.g., judpe–COURT) (although see Bourassa & Besner, 1998, for a demonstration that these latter effects can become significant if the experiment has enough power). Taken together, these results suggest that, for example, the letter string JUGDE has more potential to activate the lexical structures for the word JUDGE than the nonword JUDPE does, a conclusion that is quite inconsistent with the assumptions of virtually all of the models of word recognition discussed in this chapter. Future empirical work should be directed at a better understanding of the nature of these prelexical codes (see e.g. Davis, 1999, Ratcliff, 1981, and Whitney, 2001, for some possibilities) and that knowledge should then be used to modify current models.

Parting Thoughts

Since the early 1970s, tremendous strides have been made in terms of understanding the visual word recognition process. A major trend that has emerged during this time period has been a movement toward word recognition models that assume considerable interactivity among the various types of lexical and semantic structures. This is not to suggest that the more autonomous conceptualizations of word recognition, such as those described in the search models, can never make a comeback. Nor is it to deny that certain local components of the word recognition system may work on more autonomous principles (e.g., the establishment of prelexical codes; see Norris et al., 2000). The picture is far from complete and future work is likely to capitalize on insights not only from experimental cognitive psychology but also from the neuroscientific study of reading development and reading disorders.

Note

Much of the author's research discussed in this chapter was supported by Natural Sciences and Engineering Research Council of Canada Grant A6333. I would like to thank Penny Pexman, Yasushi Hino, Manuel Perea, and Sachiko Kinoshita for their splendid contributions to that research and for their comments on earlier drafts of the chapter. I would also like to thank Ken Forster and Ken Paap for their helpful comments on sections of this chapter in which their models were discussed.

The two processes, grapheme–phoneme rules and direct access, should distinguish between skilled and unskilled readers. An unskilled reader should identify words by applying grapheme–phoneme rules, like a child who has just learned to read. Skilled readers should bypass grapheme–phoneme rules as direct access becomes available. Skilled readers should only use grapheme–phoneme rules when they confront an unfamiliar word such as *pharisee* (Doctor & Coltheart, 1980).

In this dual-process view, learning to read depends crucially on learning the alphabetic principle, of which grapheme–phoneme rules is one hypothetical approximation. Skilled reading, in contrast, is predicted to occur without mediating phonology. Evidence from homophone errors and priming studies does not corroborate this clear-cut prediction, but neither is the prediction ruled out, as explained next.

Homophone errors

In a semantic categorization task, a homophone target such as *break* is sometimes miscategorized as a *part of a car* (Van Orden, 1987). This error stems from the fact that *break* shares identical phonology with *brake*. At face value, such homophone errors clearly demonstrate mediating phonology; words are confused because they share the same phonology.

Dual-process theory predicted that skilled readers would read familiar words via direct access. Yet skilled readers make homophone errors to homophone words irrespective of their familiarity. Homophone errors are no less likely when frequently read homophones like *break* appear as targets than when targets are relatively unfamiliar homophones, like *peek* for the category *part of a mountain*. Insensitivity to word familiarity would appear to falsify the direct-access hypothesis of skilled reading, but the story is not that simple.

If a categorization task includes more broadly specified categories such as *object*, then familiar homophone targets such as *break* produce no more errors than control items (Jared & Seidenberg, 1991; Van Orden, Holden, Podgornik, & Aitchison, 1999). Familiarity now matters. Semantic categorization to broadly specified categories produces homophone errors to low frequency homophones like *peek* but not to high frequency homophones like *break*. This finding is inconsistent with the previous results, but consistent with the direct-access hypothesis. Direct-access should be available for words that are frequently read and direct access would preclude homophone errors.

On the basis of homophone errors' absence, one could argue that the readers are not using phonology (Jared & Seidenberg, 1991). That is, the null effect of homophone phonology, when familiar homophones are judged against broadly specified categories, could imply that no phonology link exists in this case. If so, then the link between phonology and comprehension of printed words is at best partial and certainly not obligatory for skilled readers.

This logic may seem too arbitrary or simplistic. Too many reasonable alternatives present themselves for how a change in task demands may eliminate a phonology effect but not eliminate phonology (Bosman & de Groot, 1996; Lesch & Pollatsek, 1993). For example, the effects of phonology as a cause of homophone word comprehension may be concealed in contexts where performance rises to ceiling, as it usually does to highly familiar words (Lukatela & Turvey, 1994a; Van Orden et al., 1999). Perhaps overly familiar homophone words are coded too efficiently to reveal a phonology effect under the conditions of the broadly specified categories (cf. Unsworth & Pexman, 2003). Or perhaps phonology interacts in complex ways with task demands and other sources of information and the question is altogether too simply framed (Van Orden, Holden, & Turvey, 2003).

Clearly, the evidence provided by homophone errors leaves the causal status of phonology as a mediator between print and meaning undecided. As a consequence, homophone errors do not answer the question of phonology and reading to the satisfaction of all reading scientists. Special circumstances of task demands are required to produce homophone errors to familiar homophones. But a phonology effect based on special circumstance is not persuasive; it will not dissuade scientists who trust the direct-access hypothesis. In the same vein, special circumstances of task demands are required to make homophone errors go missing, and the consequent null phonology effect will not dissuade scientists who trust that reading includes mediating phonology.

Priming studies

The direct-access hypothesis has a constant traveling companion: the assumption that mediating phonology is delayed with respect to direct access (as Frost, 1998, points out; however, see Paap, Noel, & Johansen, 1992). According to this assumption, word identification via an assembly process of grapheme–phoneme rules takes more time than the direct visual associations of direct lexical access. Thus, for example, skilled readers may not base their response in a lexical decision task on phonology representations because direct access recognizes a familiar word, as a word, prior to assembly of phonology. Perhaps studies that address the delayed-phonology hypothesis may decide the status of phonology in reading.

How soon after seeing a printed word does phonology become available? One way to answer this question is with a combination of backward masking and priming. Masking concerns the length of time that items are visible. Priming concerns how one letter string may affect another, how a prime such as *REEZ* may affect identification of a target such as *rose*, for instance. The target word *rose* appears for a fraction of a second before it is replaced by the prime *REEZ. REEZ* serves as a mask of *rose* because it limits the amount of time available to derive *rose* phonology, and it serves as a prime because it shares partial phonology with *rose*, the consonants /r..z/. The prime *REEZ* itself is also briefly presented before being replaced by a visual pattern mask such as #####. The pattern mask ends visibility of *REEZ.* Backward masking strictly limits the time that *rose* and *REEZ* are visible, which limits the time available to derive phonology. If phonology becomes available rapidly, then the interaction of *REEZ* and *rose* phonology should benefit identification of *rose*, compared to a control condition.

The backward priming paradigm revealed that phonology is available very soon after seeing a word. Berent and Perfetti (1995) demonstrated that consonant phonology of pseudoword primes such as *REEZ* is available 20–40 ms after the pseudoword becomes visible (cf. Lee, Rayner, & Pollatsek, 2001; Perry & Ziegler, 2002; but cf. Lukatela & Turvey, 2000). Colombo, Zorzi, Cubelli, and Brivio (2003) established that both consonant and vowel phonology of printed Italian are available under the same conditions of brief visibility. And Lukatela, Frost, and Turvey (1998) demonstrated that the phonology of pseudohomophones such as *KLIP* is available within a 29ms window of visibility (see also Berent & Van Orden, 2000, 2003; Lee, Rayner, & Pollatsek, 1999; Perfetti, Bell, & Delaney, 1988; Rayner, Sereno, Lesch, & Pollatsek, 1995; Xu & Perfetti, 1999).

Rapidly available phonology is at least consistent with the possibility that phonology is a mediating cause in word comprehension (Frost, 1998; see also Frost, this volume).

Yet it is one thing to demonstrate that phonology is rapidly available and another thing to demonstrate that phonology has priority over direct access. Ziegler, Ferrand, Jacobs, Rey, and Grainger (2000) conducted an incremental priming study to explore the latter issue. Incremental priming allows a continuous manipulation of how one letter-string may affect another, how a prime nonword may affect a target word, for instance. The beauty of incremental priming is its precise control over when primes are available. The duration or intensity of priming words can be changed incrementally from a range in which primes do not benefit the identification of target words to a range in which they do. This adds a dimension of control that is missing in most other priming studies (Jacobs, Grainger, & Ferrand, 1995).

Ziegler and his colleagues examined the relative priority of phonology versus direct access using forward masking and they conducted the experiment in French. Forward masking rearranges the order of events compared to backward masking. A forward-masking trial briefly presents a mask (#####), which is quickly replaced by a prime, which is in turn replaced quickly by a target word. The target word remains visible until the participant responds.

The experimental manipulation consisted of three priming conditions that differ in similarity between prime and target. In one condition the primes were similar to targets in spelling and identical to targets in French phonology (e.g., pseudohomophone *nert* for target word *NERF).* This condition was called the O+P+ [O plus, P plus] condition, O+ implying similar orthography between prime and target, and P+ implying similar phonology. Their second condition O−P+ [O minus, P plus]presented primes that were dissimilar in spelling but identical in phonology (e.g., pseudohomophone *nair* for target NERF). And their third condition O+P− presented primes that were similar to targets in spelling but dissimilar in phonology (e.g., nonword *narf* for target NERF).

In all three prime conditions, lexical decisions to targets showed facilitation from priming compared to a no-prime control condition. A facilitation effect equals the degree to which the prime reduces the latency of the target “word” decision-time, compared to a baseline. In the facilitation calculus of Ziegler et al. (2000), the slower response time in the O−P+ condition minus the faster response time in the O+P+ condition estimates the facilitation effect of spelling similarity – the direct access effect. Likewise, O+P− minus O+P+ estimates facilitation due to similar phonology – the mediating phonology effect.

With a prime duration of 29 ms, the facilitation calculus revealed a greater magnitude of facilitation due to similar spelling compared to facilitation due to similar phonology. Similar spelling outdid similar phonology and so direct access must have priority over mediated access from phonology (see also Ferrand & Grainger, 1992, 1993, 1994). Other studies using masking and priming paradigms have found comparable patterns that sometimes include null effects of similar phonology, which seems to reinforce the case for priority of direct access (Brysbaert & Praet, 1992; Davis, Castles, & Iakovidis, 1998; Shen & Forster, 1999; Verstaen, Humphreys, Olson, & D’Ydewalle, 1995). Again however the story is not so simple; the pattern of facilitation changes if a different task is used.

In a comparable word naming study, Ziegler et al. (2000) observed results that contradict the pattern from lexical decision. In word naming, similar phonology appears to outdo similar spelling at all prime durations and the pattern becomes statistically reliable at a prime duration of 42 ms (see also Montant & Ziegler, 2001). In this case, the results suggest that mediated access from phonology has priority over direct access. So which task demands are most comparable to the demands of natural skilled reading – forward masking or backward masking, 29 ms or 42 ms, lexical decision or naming, or none of the above? The story only gets murkier. How one may interpret Ziegler et al.’s (2000) lexical decision results rests on debatable assumptions about similarity and activation, and a possible confound, which is discussed next.

The facilitation logic depends on whether similarity has been straightforwardly added in or subtracted out of relations between primes and targets. This may not be the case for the O+P− lexical decision primes. The O+P− condition was supposed to entail a reduction in similar phonology between primes and targets compared to the O+P+ condition. The contrast between the conditions was meant to isolate the facilitation due to the more similar phonology of the O+P+ condition: O+P− response times minus O+P+ response times estimated facilitation due to more similar phonology. However, in the O+P− “priming condition the consonantal skeleton is typically maintained”; for example *n..rf − N..RF* (Ziegler et al., 2000, p. 687). This creates a confound whereby O+P− and O+P+ priming can be almost identical at very short prime durations. As a consequence, the magnitude of facilitation due to similar phonology is systematically underestimated.

Consonant phonology is more quickly available than vowel phonology in languages with predominantly ambiguous vowel spellings (Berent & Perfetti, 1995; Lee et al., 2001; Perry & Ziegler, 2002). The earliest moments of activation emphasize reliable correspondences between consonant spelling and phonology, and the O+P− primes share these reliable correspondences with their targets. This means that O+P− primes are comparable to the O+P+ primes in their potential for facilitation in the earliest moments of activation. As a consequence, phonology priming is underestimated at the shortest prime durations, such as the 29 ms duration. The contrasted conditions O+P− versus O+P+ could only be expected to diverge at longer prime durations, as vowel phonology comes into play. This confound undermines the contrast in the 29 ms condition that seemed to favor similar spelling over similar phonology. The confound renders the lexical decision outcome equivocal; it no longer favors direct access.

Priming manipulations that contrast degrees of similarity are often problematic. How does one discount the similarity in phonology that is inherent when items are similar in spelling? Some accounts claim that the first instants of word comprehension include multiply active patterns of phonology that, over time, settle into a single pattern (e.g., Kawamoto & Zemblidge, 1992; Van Orden et al., 1990; Van Orden & Goldinger, 1994). Consequently, items such as *plaid* and *plain* would activate virtually identical “clouds” of phonology in the first milliseconds, but they are not identical in spelling and do not settle into the same phonology. Other accounts assume that the first milliseconds of word comprehension include incompletely specified phonology. This assumption also allows that similar spellings may activate identical phonology at the outset of word comprehension (e.g., Berent & Perfetti, 1995; Frost, 1998).

Finally, how does one insure that similarity along a phonology dimension is ever comparable in magnitude to similarity along a spelling dimension? Do null effects of similar phonology stem from weak manipulations of similarity? Sometimes yes; other times nobody knows (Frost, Ahissar, Gotesman, & Tayeb, 2003). Again, how one interprets the idiosyncratic task conditions that produce the evidence determines how one interprets the evidence, and there are inestimable degrees of freedom for interpretation of task demands. Like homophone errors, the evidence from masked priming studies leaves the causal status of phonology representations in skilled reading undecided. Some special task demands yield reliable effects of phonology variables, and others do not.

Giving Up Ether

Ideally, robust phonology effects would be found in all laboratory reading contexts. Ideally, laboratory methods should reveal a blueprint of reading that is independent of the laboratory tools used in the investigation. With respect to this ideal, a phonology effect that cuts across all reading contexts would satisfy the requirements (Jacobs & Grainger, 1994). Mediating phonology would then become an accepted component in the architecture of word comprehension. But performance of an interactive system, rather than with respect to isolated causal factors. One prominent variable is ambiguity.

Notice how many ways the same ambiguous phoneme /eI/ can be spelled in *Kay, weigh, made*, and *pail*, or how many ways the same ambiguous vowel spelling *ai* can be pronounced in *plaid, raid, said*, and *aisle.* Such ambiguity has consequences for performance. For instance, if a presented spelling can be pronounced in more than one way, then it yields a slower naming time compared to an unambiguous spelling, all other things equal.

Ambiguity is not your standard causal factor. Ambiguity effects cannot be localized in spelling or phonology taken separately. Ambiguity is only defined in a relation between the two. Thus empirical tests for ambiguity effects are tests about how phonology is related to other aspects of language, such as spelling. The next sections of this chapter describe empirical findings that demonstrate ambiguity effects.

Simulations of interactive processes

Before turning to the experiments, briefly consider some previous simulations of interactive processes. Simulations of interactive processes among spelling, phonology, and semantics have changed the way scientists look at the structure of language (e.g., Grossberg & Stone, 1986; Jacobs, Rey, Ziegler, & Grainger, 1998; Kawamoto & Zemblidge, 1992; Masson, 1995; McClelland & Rumelhart, 1981; Plaut, McClelland, Seidenberg, & Patterson, 1996). They have focused scientists' attention on ambiguity and the statistical structure of language (Plaut et al., 1996; Saffran, 2003; Van Orden et al., 1990), and they introduced the possibility of feedback in word comprehension.

Consider the ambiguous spelling of the homograph word *wind. Wind* has two legitimate pronunciations: it can rhyme with *pinned* or *find.* In an interactive model, spelling nodes representing *wind*'s spelling activate nodes that represent the two pronunciations of *wind*, and these two pronunciations both feed back activation to their common spelling. This creates two competing feedback loops, which characterizes how ambiguity is expressed in an interactive model. Ambiguity breeds competition between multiple potential outcomes, which takes time to resolve (see also Lupker, this volume).

Kawamoto and Zemblidge (1992) simulated the competition between homograph pronunciations as it unfolds across a naming trial. The model included feedforward and feedback connections among letter, phoneme, and semantic node families. The connections modulate node activity very roughly as synapses may modulate the activity of neurons. In the Kawamoto and Zemblidge model, connections were excitatory between node families but mostly inhibitory within node families. For instance, letter nodes excite phoneme nodes and phoneme nodes excite letter nodes, but competing phoneme nodes inhibit each other. Consequently, phoneme nodes compete directly with other phoneme nodes and indirectly with letter or semantic nodes. A phoneme node competes indirectly by activating some particular letter or semantic node that can compete directly. Thus every node interacts with every other node, either directly or indirectly.

Simulations have been successful as guides for how to look at language. They are less successful as models of actual psychological processes. Despite highly unintuitive and yet reliable predictions, actual simulations are perpetually challenged by the details of human performance (e.g., Spieler & Balota, 1997; Treiman, Kessler, & Bick, 2003). It is the assumptions behind the simulations that seem to capture a reliable picture of language, but painted in somewhat broad strokes.

Ambiguity at the scale of whole words

Homographs like *wind* have a dominant pronunciation (the more regular pronunciation that rhymes with *pinned*) and a subordinate pronunciation (the less regular pronunciation that rhymes with *find*). In an actual word naming experiment, some readers will produce the dominant pronunciation and some will produce the subordinate. Also, when the dominant pronunciation is produced, it yields faster naming times, on average, than the subordinate pronunciation. One way to think about this pattern is that the two pronunciations compete in the course of a word naming trial prior to an observed pronunciation.

In a simulated naming trial, *wind*'s subordinate pronunciation is less strongly activated, at least initially, but nevertheless can win the competition. To do so it must accrue sufficient activation, within the time course of the trial, to overcome activation of the dominant pronunciation. This implies an on-line qualitative change from dominant to subordinate phonology. The qualitative change occurs at an exchange point in what is called a bifurcation. Kawamoto and Zemblidge (1992) simulated the bifurcation of a homograph pronunciation, from statistically dominant to subordinate, as a transcritical bifurcation.

The dominant pronunciation of *wind* has a stronger feedback loop between letter and phoneme nodes, a stronger and more stable local attractor. The subordinate pronunciation has the weaker or less stable attractor between letter and phoneme nodes, but has the more stable attractor between phoneme and semantic nodes. The feedback loop between phoneme and semantic nodes takes some time to grow in strength and lend sufficient support to *wind*'s subordinate pronunciation. Enough support makes *wind*'s subordinate pronunciation a winner. This outcome occurs when a reader or model is sufficiently more familiar with the subordinate pronunciation's semantic variants, which counters the inherent disadvantage of the subordinate pronunciation's less-regular relation between spelling and phonology.

Initially *wind* activates the two pronunciation patterns, and the dominant pattern is initially favored. However, slowly accruing activation in a semantic and phoneme feedback loop lends increasing support to the subordinate pronunciation. Within the time of a naming trial, activation in the phoneme-semantic feedback loop grows to a sufficient degree that it turns the tide in the competition. The tide turns at the bifurcation point. Within the time between the appearance of *wind* and a pronunciation, semantic-phoneme activation and the subordinate's letter-phoneme activation overtake the otherwise dominant pronunciation. At the bifurcation point, semantic-phoneme feedback puts *wind*'s subordinate pronunciation over the top, and the dominant pronunciation exchanges stability with the less-regular subordinate pronunciation. Subsequently, the model produces the subordinate pronunciation.

So why do some readers produce the dominant pronunciation and others the subordinate? Different readers, or models, may sample language differently. Each reader has a unique history of covariation among words' spellings, phonology, and semantics. Pronunciations can have strong or weak ties to semantics based on different readers' different familiarity with different words. At any particular time, some readers will quickly produce the dominant more regular pronunciation, and other readers, sufficiently more familiar with subordinate variants, will more slowly produce the subordinate pronunciation.

Wind's homograph spelling is one ambiguous spelling, one pocket of ambiguity, within a reader's accumulated sample of English. Yet *wind* is only ambiguous if that reader's history includes samples of both interpretations of *wind*. A reader's sample of a language delimits the potential for ambiguous or unambiguous relations. The aggregate statistical pattern of relations that makes up a reader's language is specific to the reader's history and changes throughout a lifetime of reading.

Multiple scales of ambiguity

Homograph ambiguity exists at multiple scales. In a homograph, every letter has associations with different pronunciations. For example, the homograph *wind* is ambiguous at a micro scale because its grapheme *i* is ambiguous. This ambiguity is amplified at a meso scale of *wind*'s ambiguous spelling-body *-ind*, and is further enlarged at a macroscale of the ambiguous whole word. In this way of thinking, local ambiguity is infectious, in a manner of speaking. A local ambiguity, like *wind*'s ambiguous grapheme *i*, infects every larger scale of spelling that has a history of multiple pronunciations.

Words infected with more ambiguity have slower naming times. Compare the homograph *wind* with the word *pint. Pints* spelling is also ambiguous but not to the same degree as *wind*. *Pint* is infected with ambiguity up to the scale of its spelling-body *-int*, but *pint* does not entail whole-word ambiguity. The difference explains why homograph pronunciations are slower than pronunciations to ambiguous control items that are not homographs (Gottlob, Goldinger, Stone, & Van Orden, 1999; but cf. Hino, Lupker, & Pexman, 2002). This outcome would be observed even if every letter of *wind*, taken one at a time, were no more ambiguous than the individual letters of *pint. Wind* has a slower naming time even when contrasted with precisely constructed *mint* and *pint* controls equated for spelling body ambiguity (Holden, 2002).

Connectionist models track in the same matrix all the scales at which spelling relates to phonology. They illustrate how all these relations can be co-instantiated; different levels of representation are not necessary for the different scales to be effective. Models with recurrent feedback connections, in addition, track multiple scale relations in all directions. Stronger feedback loops like those of dominant relations correspond to relatively more stable attractors in the network. One can find dominant and subordinate relations at each scale, which means that dominant and subordinate relations may be nested across scales. In other words, there are relations within relations, attractors within attractors.

Now everything is in place to discuss feedforward ambiguity effects, and then feedback ambiguity effects, that have been demonstrated empirically. Ambiguity effects can be identified at the scale of spelling-bodies and graphemes and feedback effects can be identified at all the same scales.

Feedforward ambiguity at the scale of spelling-bodies

The more regular dominant pronunciation of the spelling body *-int* rhymes with *mint* (consider *lint, tint*, and *hint).* The subordinate pronunciation of *-int* rhymes with *pint. Pint* takes longer to name than *mint* because *pint*'s rime is the subordinate pronunciation of the body *-int.* When *pint* is the word to be named, a mispronunciation of *-int* to rhyme with *mint* strongly competes with *pint*'s correct pronunciation. This competition is so close that a mispronunciation of *pint* can be elicited even from skilled readers. For example, participants can be trained to respond rapidly, in time with a beat, in a word naming task, but in doing so they commit errors of pronunciation including the kind of error in which *pint* is mispronounced to rhyme with *mint* (Kello & Plaut, 2000). More slowly emerging semantic features must combine with *pint*'s correct pronunciation to counter the dominant rhyme with *mint* (Farrar & Van Orden, 2001). In this case, *pint*'s rhyme with *mint* is not a word and would not have coherent semantic associations (cf. Lesch & Pollatsek, 1998, however).

When *mint* is the word to be named, the subordinate mispronunciation that would rhyme with *pint* competes with *mint*'s correct pronunciation. Just as for homographs, the two pronunciations compete in the course of a naming trial prior to an observed pronunciation, and the competition takes time to resolve. Thus naming times to *mint* should be slower than to words with unambiguous body-rime relations. Compare the spelling body *-int* with *-uck*, the spelling body of *duck. Duck's* spelling body is unambiguous; it supports only one pronunciation (consider *luck, buck, muck*, and *puck).* The /uk/ rime also reliably covaries with the *-uck* body. Together they form an invariant relation between body and rime, and rime and body. Indeed, words like *mint* are more slowly named than words like *duck* (Glushko, 1979). A word like *mint* is more widely infected with ambiguity than a word like *duck*.

Feedforward ambiguity at the scale of graphemes

Pockets of more or less ambiguity are also found at the microscale of graphemes and phonemes (compare Zorzi, Houghton, & Butterworth, 1998). English vowel spellings are almost always ambiguous. But some English consonants have invariant relations with phonology. The consonant grapheme *d*, at the beginning of a word, is always associated with the phoneme /d/, and the /d/ phoneme is always spelled *d*. Overall, in English, consonant spellings covary more reliably with their pronunciations than do vowel spellings. Consequently, in English, consonant phonology is resolved earlier than vowel phonology. For example, the relative ambiguity of consonant and vowel spellings predicts when their phonology will become available in masked priming experiments: consonant phonology coheres before vowel phonology (Berent & Perfetti, 1995; Lee et al., 2001; Perry & Ziegler, 2002).

For a visually presented word, the mapping from spelling to phonology is the feedforward relation and the mapping from phonology to spelling is the feedback relation. For auditory presentations this is reversed. The mapping from phonology to spelling is feedforward and the mapping from spelling to phonology is feedback. The previous examples all concerned ambiguity from spelling to phonology. Ambiguity effects also generalize to the inverted mapping from phonology to spelling, as feedback in visually presented homophones for instance.

Feedback ambiguity at the scale of whole words

Consider the homophone phonology /braik/ and the corresponding spellings *break* and *brake.* Just as the homograph *wind* supports two pronunciations, the homophone /braik/ supports two spellings. Homophone words produce slower visual lexical decision times than control words that are not homophones (Ferrand & Grainger, 2003; Pexman, Lupker, & Jared, 2001; Pexman, Lupker, & Reggin, 2002). Homophones have slower lexical decision times even when contrasted with precisely constructed controls equated for rime-body ambiguity – feedback effects accrue across scales (Holden, 2002).

Notice that homophone effects in visual lexical decision are unintuitive. From the traditional view, activation should always flow forward from a cause to an effect, as from spelling to phonology in a visual lexical decision task. It should not matter for visual lexical decisions that *break*'s pronunciation /braik/ may have more than one spelling, unless there exists feedback from phonology to spelling. Consequently, slower visual lexical decision times to homophone words imply feedback from phonology to spelling.

effects mentioned so far. After all they concern relations with words' meanings and it is the pursuit of meaning that drives word comprehension in reading.

Letter perception

A briefly presented pseudohomophone such as *brane* can induce the false impression that a pre-designated letter *i* was seen (Ziegler, Van Orden, & Jacobs, 1997). Participants report that an *i* appeared in the presented spelling *brane*, but only if the letter is contained in *brane's* sound-alike base-word *brain.* The flip side of this effect is also observed. Pseudohomophones such as *taip* may induce the false impression that a pre-designated letter *i* did not appear, but only if the letter is missing from *taip's* sound-alike base-word *tape.* These phenomena were first demonstrated in German (Ziegler & Jacobs, 1995), then later in English (Ziegler et al., 1997) and French (Lange, 2002). Such phenomena appear quirky within a conventional framework where they may suggest postlexical inferences about which letters were seen. They are expected, however, if feedback from base-word phonology activates *brain*'s letters or inhibits letters that are not present in the base-word *tape.*

Perceived lexicality

Relations between spelling and phonology are sources of perceived lexical structure (Vanhoy & Van Orden, 2001). Wordlike body-rimes actually add "word-ness" to letter strings that are not words. For example, it is widely reported that lexical decisions to pseudohomophones such as *jale* are slower and more likely to end in a false 'word' response than are control items. Also, correct 'word' responses to actual words are slower when pseudohomophones appear as foils.

It is not simply that pseudohomophones mimic word phonology; it also matters that they are composed of body-rime relations like those found in actual words. *Jale* is constructed on an extant body-rime that appears in the words *bale, sale*, and *tale.* The pseudohomophone *stahp*, which sounds like *stop* in American English, is constructed on a novel body-rime that does not appear in an actual word. In lexical decision, pseudohomophones like *jale* produce reliable pseudohomophone effects; pseudohomophones like *stahp* do not.

Natural variation across languages

Each language presents a unique compilation of ambiguity that will be uniquely sampled by each reader. Hebrew includes mostly homographs, and Chinese includes very many homophones. Dutch, Spanish, German, and Italian minimize or eliminate ambiguity between phonology and spelling by staying closer to a system of grapheme–phoneme rules. French is more like English. French has ambiguities at multiple scales of correspondence between phonology and spelling. Serbo-Croatian has two alphabets that sometimes contradict each other in their relation to phonology, and other times not. Clearly, the consequences of ambiguity for complex interactions must be worked out carefully one language at a time (e.g., Frost, this volume; Colombo et al., 2003; Frost, Katz, & Bentin, 1987; Goswami, Ziegler, Dalton, & Schneider, 2003; Lukatela & Turvey, 1998; Ziegler, Perry, Jacobs, & Braun, 2001; Ziegler et al., 2000; and many other publications not cited here). Different languages exaggerate or reduce different sources of ambiguity and all sources interact in performance (Bosman & Van Orden, 1997; Lukatela & Turvey, 1998; Van Orden & Goldinger, 1994).

Is dual-process theory false?

The spectrum of ambiguity effects and feedback effects that experiments demonstrate would not likely have been anticipated with dual-process theory as the guide. However, this does not mean that dual-process theory is false. Findings that contradict dual-process theory simply reveal that grapheme–phoneme rules were not the best compass to discover salient structure between phonology and spelling (Paap et al., 1992). The theory itself can be reconstituted indefinitely to absorb new contradictory findings (e.g., Coltheart, Curtis, Atkins, & Haller, 1993; Norris, 1994; Zorzi et al., 1998). Ad hoc changes create alternative ways to see the contradictory data and can be useful for that fact (Feyerabend, 1993). Nonetheless, it is a bit hard to imagine how scientists in the exclusive pursuit of mechanistic causal chains would have stumbled on these effects. The discovery of feedback effects as predicted by feedback models is a remarkable discovery of basic reading science with profound implications for all cognitive science.

What is the nature of response time?

The last point is a caveat that concerns how one should look at the data from all these experiments. The previous discussion has emphasized mean effects, differences between average response times or accuracy, as did almost all of the cited authors. This will prove in time to have been misleading. Ambiguity effects are not so simply expressed; they do not simply reflect shifts in average response times. Rather, they largely reflect increases in the proportion of very slow responses. They reflect redistribution of response times and changes in the shapes of response time distributions (Holden, 2002). This general observation about effects and response times is not new to reading science (Andrews & Heathcote, 2001; Balota & Spieler, 1999), but its implications have not been widely acknowledged.

Redistributions of response times often appear as changes in so-called power laws – equations in which the probability of a particular response-time is a function of the response-time itself (Holden, 2002; Van Orden, Moreno, & Holden, 2003). Power laws may suggest a complex interdependence in which the processes that compose a system change each other as they interact (Jensen, 1998). Consequently, co-instantiated relations between phonology and spelling, for example, become causally entwined and interdependent (Van Orden & Holden, 2002; Van Orden et al., 2003). It is the nature of living systems that they comprise entwined processes and do not reduce to causal elements (e.g., Rosen, 2000; Wilson, 2003).

Power law behavior could imply a radical suggestion that separate representations of phonemes and letters, for example, need not be posited. Relations between a word's spelling and its phonology, its body and rime, and its graphemes and phonemes become mutually reinforcing relations with neither being causally prior to the other. Yet there remains a useful way to think about cause in the sense of a basis or foundation for reading. Unless a child becomes attuned to the alphabetic principle in relations between spelling and phonology, learning to read does not occur or occurs with great difficulty (Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001). In this sense of cause, the alphabetic principle has a causal priority in the development of skilled readers.

Summary and Conclusions

The first half of this chapter ended on the horns of the dilemma concerning phonology and skilled reading. Over 100 years of reading research failed to decide whether skilled reading involves mediating phonology, or whether it does not. The question of mediating phonology hinges on the discovery of a task independent phonology effect for skilled readers reading familiar words. This discovery could possibly situate phonology in the architecture of word comprehension, part of cognition's larger absolute frame of reference. However, despite the plausibility that such a phonology effect could exist, all phonology factors, like all other word factors, change the pattern of their effects across the variety of task conditions.

The second half of this chapter reviewed ambiguity effects at multiple scales of relations between spelling and phonology. The reviewed findings present snapshots of a complex structure that relates phonology and spelling. In the contemporary picture of English, this relation appears as a context sensitive, bidirectional, statistical structure that changes on multiple scales and in each instance of reading – a statistical structure in perpetual motion, one might say. The complex structure of ambiguity effects intertwines written and spoken English in feedback. Some prominent intertwined relations are readily discernible, relations like those between bodies and rimes, or graphemes and phonemes. Nonetheless, the intention is not to propose a pretty hierarchy, and it would soon sprout weeds in any case. Letters and groups of letters change their relation to phonemes and groups of phonemes according to the contexts in which they appear.

Feedback models of interacting processes predict ambiguity and feedback effects. Context sensitivity within these models is useful to explain the context sensitivity of relations between spelling and phonology. It is a natural extension of this view to expect context sensitivity at all levels of a system, including sensitivity to the laboratory contexts of task demands. Until now context sensitivity has been a reason not to take some other scientist's data as conclusive. Now context sensitivity is the likely key to understand reading, the paradigmatic cognitive performance.

Note

We acknowledge support from the National Science Foundation. The views expressed in this chapter are the authors' and do not necessarily represent those of the National Science Foundation.

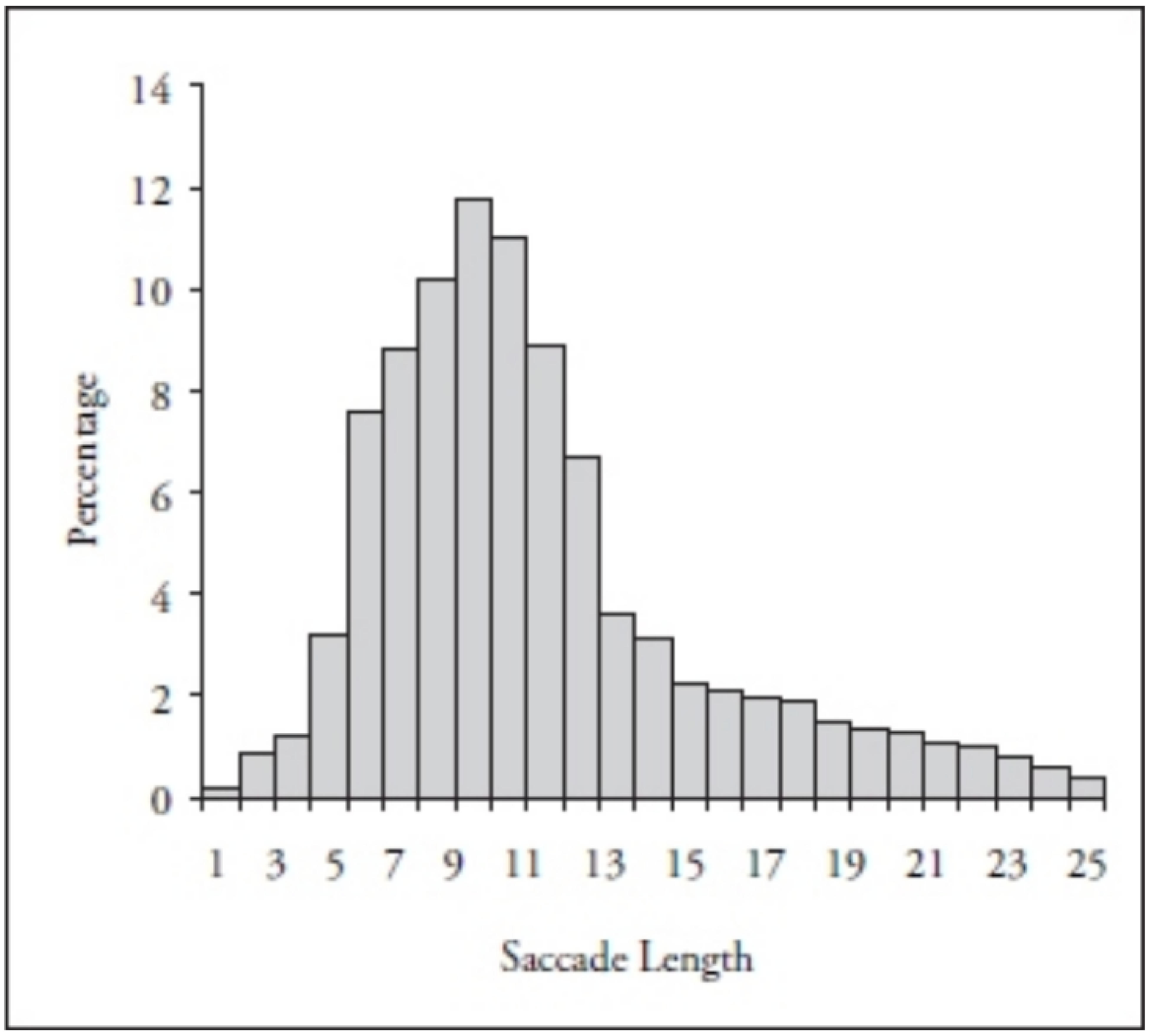
5

Eye Movements During Reading

Keith Rayner, Barbara J. Juhasz, and Alexander Pollatsek

The study of eye movements has a long and rich history in reading research. Indeed, some of the earliest experimental studies of skilled reading involved measuring eye movements (see Huey, 1908). Since 1975, there has been an increasing awareness that eye movements provide very important information about the moment to moment processing that occurs during reading (Rayner, 1978, 1998). In this chapter, we first provide background information about the basic characteristics of eye movements during reading and how they are affected by reading skill. Then we review research on (1) the perceptual span during reading, (2) how much readers benefit from a preview of words to the right of the fixated word during reading *(preview benefit*), and (3) the control of eye movements during reading. Much of the research on eye movements during reading has focused on these issues. Following our discussion of these important issues, we discuss recent trends regarding eye movements and reading. We conclude with a discussion of models of eye movement control in reading.

We will begin by making two important points with respect to eye movement research. First, there are two types of research with respect to eye movements and reading (see Rayner, 1995; Rayner & Liversedge, 2004, for discussion). Some researchers are primarily interested in eye movements per se and use the task of reading as a way to study the oculomotor system. At the other extreme are researchers who use eye movements as a tool to study some aspect of language processing. This group tends not to be interested in the details of eye movements per se. From our perspective, it is important to have some understanding of research from both approaches because low-level oculomotor variables impinge on higher-order processing and vice versa (Rayner & Liversedge, 2004). Second, although a great deal of data have been collected regarding eye movements in reading, perhaps the most important recent trend is the development of sophisticated models of eye movement control that simulate reading performance. We will discuss this trend later.



One final point that we will make is that the nature of the writing system influences eye movements. Above, we discussed the characteristics of eye movements in alphabetic writing systems (particularly English). With nonalphabetic systems (such as Chinese or Japanese), eye movements are clearly affected. Thus, fixation durations in Chinese and Japanese tend to be longer and saccades tend to be shorter in terms of number of characters; however, a character in Chinese or Japanese conveys more information than a letter in an alphabetic system. Hebrew, which is more densely packed than English (though not as densely packed as Chinese or Japanese) because most vowels are omitted, yields shorter saccades than English.

Table 5.1 Developmental Characteristics of Eye Movements During Reading (Adapted from Rayner, 1998)

	Grade level						
	1	2	3	4	5	6	Adult
Fixation Duration (ms)	355	306	286	266	255	249	233
Fixations per 100 words	191	151	131	121	117	106	94
Regression Frequency (%)	28	26	25	26	26	22	14

Reading Skill and Eye Movements

For over 80 years (Buswell, 1922), it has been known that reading skill influences eye movements. Skilled readers make shorter fixations, longer saccades, and fewer regressions than less skilled readers (Rayner, 1978, 1998). Furthermore, there are marked developmental trends in eye movements: as reading skill increases, fixation durations decrease, saccade lengths increase, and the frequency of regressions decreases. Table 5.1 shows a summary of important eye movement measures from beginning reading to sixth-grade level, with adult data for comparison. Here, it can be seen that there is a steady decrease in the average fixation duration and the number of fixations per 100 words as reading skill increases. The most marked changes occur between beginning reading and about third- or fourth-grade level. By the time children have had four years of reading experience, their eye movement behavior is not too different from adults. The exception is that the frequency of regressions is larger for a sixth-grader than an adult reader.

To date, there have been little data examining the effect of aging on eye movements. The studies that do exist (Kliegl, Grabner, Rolfs, & Engbert, 2004; Solan, Feldman, & Tujak, 1995) indicate that older readers (i.e., approximately 70 years old) have slightly longer fixations on average than younger readers, and they also make more fixations and more regressions than their younger counterparts. But Kliegl et al. (2004) concluded that the similarities that existed in the eye movement patterns of the young and older adult readers were much more impressive than the differences between them.

One area that has been highly controversial concerns the eye movements of poor readers and readers with dyslexia. Obviously, disabled readers make longer fixations, shorter saccades, more fixations, and more regressions than normal readers. Given this, it has sometimes been suggested that faulty eye movements cause poor reading and dyslexia. We will not review the research in this area (see Rayner, 1998, for a complete summary), but the best evidence indicates that eye movements rarely are the cause of reading disability. Rather, less fluent eye movements reflect the difficulties that disabled readers are having understanding the text they are reading.

There may be differences in eye movement characteristics in readers with dyslexia as a function of their writing system. Specifically, studies with Italian dyslexic readers (De Luca, Borrelli, Judica, Spinelli, & Zoccolotti, 2002; De Luca, Di Pace, Judica, Spinelli, & Zoccolotti, 1999) and German dyslexic readers (Hutzler & Wimmer, 2004) suggest some differences. Specifically, the Italian readers had moderately increased fixation durations, but not a lot of regressions. However, they made lots of fixations and short sac-cades. The German dyslexics, like the Italian dyslexics, made fewer regressions than are typically seen in readers of English, but had very long fixation durations. The lower incidence of regressions by the Italian and German dyslexics may be due to the fact that the orthography is more regular than English. Hutzler and Wimmer (2004) suggested that the longer fixation durations of the German dyslexics might be due to the greater syllabic complexity of German.

Eye Movements and Measures of Processing Time in Reading

As noted above, eye movements have become recognized as one of the best ways to study moment-to-moment language processing (Rayner & Liversedge, 2004). Thus, eye movement data are widely used to study topics such as lexical ambiguity resolution (Binder, 2003; Duffy, Morris, & Rayner, 1988), phonological coding (Jared, Levy, & Rayner, 1999; Pollatsek, Lesch, Morris, & Rayner, 1992; Rayner, Pollatsek, & Binder, 1998), morphological processing (Andrews, Miller, & Rayner, 2004; Hyönä & Pollatsek, 1998; Juhasz, Starr, Inhoff, & Placke, 2003; Niswander, Pollatsek, & Rayner, 2000; Pollatsek, Hyönä, & Bertram, 2000), syntactic ambiguity and parsing (Binder, Duffy, & Rayner, 2001; Clifton et al., 2003; Frazier & Rayner, 1982), and discourse processing (Cook & Myers, 2004; Garrod & Terras, 2000; O'Brien, Shank, Myers, & Rayner, 1988). We will not attempt to review the results of these studies here. Some of these studies rely on examining the eye movement measures on a single target word, whereas others rely on examining eye movements in a larger segment of text. In this section, we will review the primary measures that eye movement recordings provide for researchers who study moment-to-moment language processing activities.

A major issue concerns how to summarize the eye movement record to obtain the best measure of processing time for a given region of text. When the unit of analysis is the word, certain measures are typically focused on, whereas when the unit of analysis is larger than a single word, other measures are employed.

With respect to the word as the unit of analysis, if readers always made only one fixation on a word there would be little difficulty choosing the most appropriate measure of processing time: the fixation duration on the word would obviously be the best measure of the time to process a word. However, many words are skipped: *content words* (nouns, verbs, adjectives, and adverbs) are fixated about 85% of the time while *function words* (prepositions, conjunctions, articles, and pronouns) are fixated about 35% of the time. One reason function words are skipped more than content words is that they tend to be short and there is clear relationship between the probability of fixating a word and its length (Rayner & McConkie, 1976). Another problem in interpreting eye movements is that many words are fixated more than once (or refixated). The problem of multiple fixations has led to alternative (highly correlated) measures. The mean fixation duration is inadequate because it underestimates the time the eyes are on a word (i.e., a 250 ms fixation and a 200 ms fixation would yield a mean of 225 ms when the eyes were actually on the word for 450 ms). The strategy of only including words that received just one fixation *(single fixation duration)* is also problematic because too many data might be eliminated. Thus, the two most frequently used measures are the *first fixation duration* and the *gaze duration* on a word. First fixation duration is the duration of the first fixation on a word regardless of whether it is the only fixation or the first of multiple fixations on a word. Gaze duration is the sum of all fixations on a word prior to an eye movement to another word. A fourth measure, the *total fixation duration* on the word reflects the sum of all fixations on the target word (including any regressions back to it). The first three measures therefore reflect the first pass processing time for a word (and are often assumed to reflect lexical access processes, as we shall discuss later in conjunction with models of eye movement control). The latter measure reflects both initial and later processing activities.

Arguments over which measure is best to use as an index of processing time partly depends on what is being examined, but the problem of assessing the average time spent processing a word is not trivial. There are three components of the problem. First, words are clearly processed when they are not fixated (Rayner, 1998). Second, readers begin processing a word before they fixate on it (which is referred to as *parafoveal preview benefit).* Third, there are *spillover effects* (Rayner & Duffy, 1986; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989) as the processing of a word is not always completed by the time the eyes move; that is, processing of a word can "spill over" onto the next word and influence how long it is fixated.

Should preview benefit and spillover time be added to the time actually spent on a word? This gets complicated and can cause frustration for researchers. Given these points it is clear that any single measure of processing time for a word is a pale reflection of the reality of the true processing associated with that word. Thus, the strategy of analyzing large amounts of text with a single measure of processing time is likely to be of limited value. A strategy that most researchers have adopted is to select target words for careful analysis and then examine many different measures. By doing so, it is possible to draw reasonable inferences about the reading time for a target word.

When the unit of analysis is larger than a single word, *first pass time* is generally used as the primary measure. The first pass time is the sum of all fixations in a region prior to moving forward in the text. It is important, when analyzing larger regions to distinguish between first pass and second pass (i.e., rereading) times. There has been some controversy regarding how to best analyze a region when readers make regressions (Altmann, 1994; Rayner & Sereno, 1994a; 1994b). For example, Rayner and Sereno (1994b) noted that when readers enter a region and then quickly make a regression out of that region, the first pass time is very short in comparison to when the reader does not regress. It appears that the most appropriate way to deal with this issue is to use regression-path duration or go-past analyses (Konieczny, Hemforth, Scheepers, & Strube, 1997; Liversedge, Patterson, & Pickering, 1998; Rayner & Duffy, 1986).

With this analysis, reading time is the sum of all fixations starting with the first fixation in a region and ending with the first forward saccade past the region under consideration. Liversedge et al. (1998) and Rayner & Liversedge (2004) discuss various issues related to categorizing eye movements spatially (i.e., grouping fixations that are all on the same region of text such as gaze duration or first pass) versus temporally (i.e., grouping a temporally contiguous set such as regression-path or go-past measures) and how to deal with regions that vary in length.

Basic Issues Regarding Eye Movements in Reading

The perceptual span during reading

How much information can a reader process in each fixation? Experiments that have used eye-contingent display techniques provide rather definitive answers to this question. Before discussing these results, however, we note that the main reason that readers make saccades is due to acuity limitations. While acuity in the central 2° of the visual field (the fovea) is very good, acuity drops off markedly in the parafovea, which comprises 5° on either side of the fixation, and is poor in the peripheral region, which encompasses the remaining information on a line of text. Thus, the purpose of eye movements in reading is to place the to-be-processed text in the fovea, where it can be most easily identified.

There are three main types of eye-contingent display paradigms (each has several variants), which have been useful for answering many questions (see figure 5.3). The *moving window technique*, was first used by McConkie and Rayner (1975). In this paradigm, a portion of text (defined by the experimenter) centered on a reader's fixation appears as it normally should. All of the text outside of this "window" is replaced by something meaningless (such as random letters or xs). This window moves as the reader's eyes move, so that each time there is a new fixation, there is a region of normal text surrounded by meaningless text (see figure 5.3). The theory behind this technique is that if the window is large enough, reading will not be affected. The second type of eye contingent display technique, called the *foveal mask technique*, is the inverse of the moving window (Rayner & Bertera, 1979). With this paradigm, letters around the fixation point are replaced by *x*s or a masking pattern. Finally, the most utilized type of eye contingent display technique is the *boundary technique* (Rayner, 1975), where there is an invisible boundary specified by the experimenter in the text. When a reader's eye crosses this boundary, the word or letter string to the right of this boundary that was originally displayed is replaced by a target word. Importantly, this change occurs during a saccade, so that it is not noticeable to the reader.