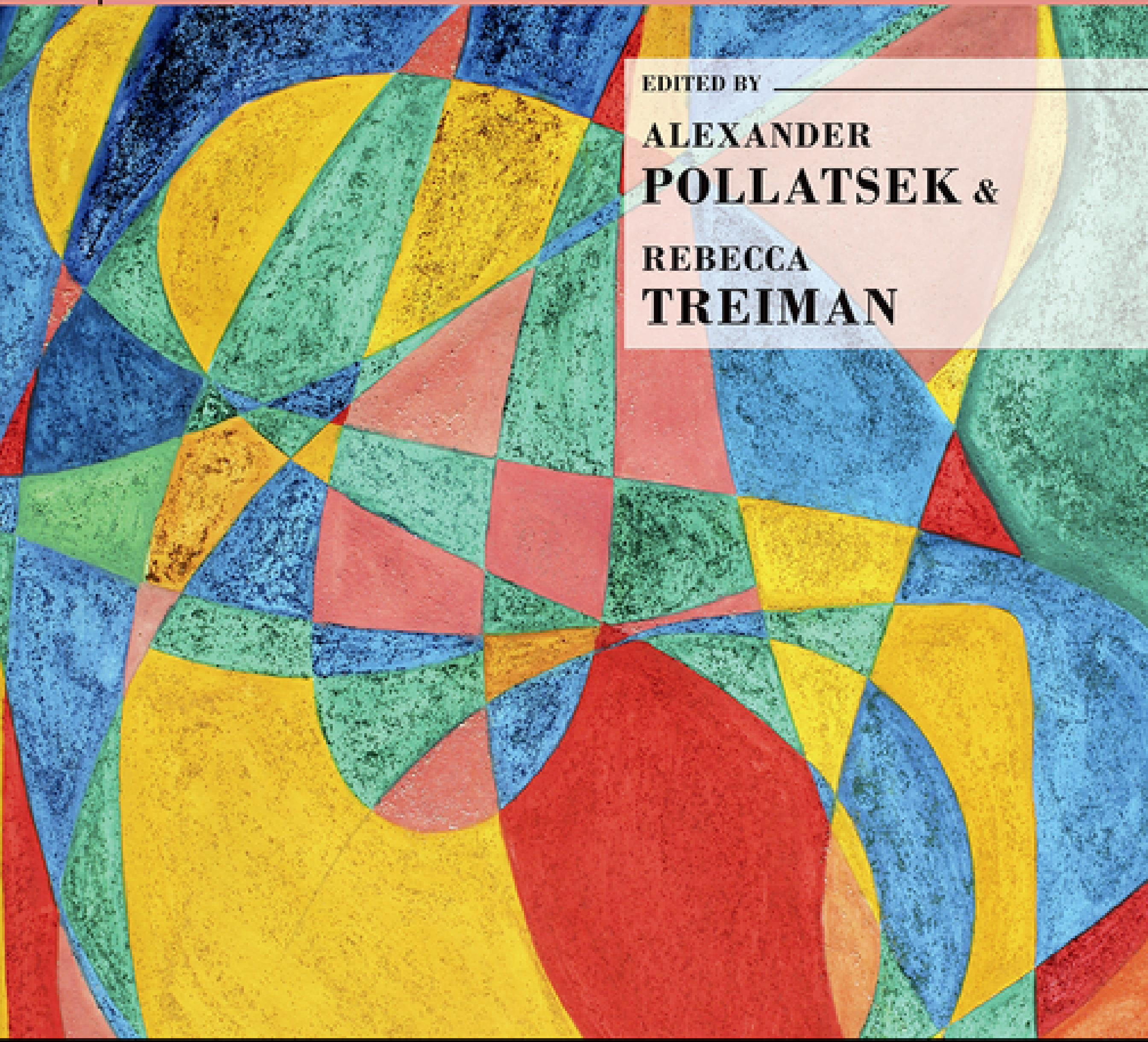
OXFORD LIBRARY OF PSYCHOLOGY



EDITED BY

ALEXANDER POLLATSEK &

REBECCA TREIMAN

The Oxford Handbook *of* READING

The Oxford Handbook of Reading

OXFORD LIBRARY OF PSYCHOLOGY

Editor in Chief PETER E. NATHAN

The Oxford Handbook of Reading

Edited by

Alexander Pollatsek

and

Rebecca Treiman

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trademark of Oxford University Press in the UK and certain other countries.

Published in the United States of America by
Oxford University Press
198 Madison Avenue, New York, NY 10016

Library of Congress Cataloging-in-Publication Data
The Oxford handbook of reading / edited by Alexander Pollatsek and Rebecca Treiman.
pages cm.—(Oxford library of psychology)
ISBN 978–0–19–932457–6 (alk. paper)
1. Reading—Psychological aspects. I. Pollatsek, Alexander, editor. II. Treiman, Rebecca, editor.
BF456.R2O94 2015
418′.4—dc23
2014039423

9 8 7 6 5 4 3 2 1
Printed in the United States of America
on acid-free paper

SHORT CONTENTS

OXFORD LIBRARY OF PSYCHOLOGY

The *Oxford Library of Psychology,* a landmark series of handbooks, is published by Oxford University Press, one of the world's oldest and most highly respected publishers, with a tradition of publishing significant books in psychology. The ambitious goal of the *Oxford Library of Psychology* is nothing less than to span a vibrant, wide-ranging field and, in so doing, to fill a clear market need.

Encompassing a comprehensive set of handbooks, organized hierarchically, the *Library* incorporates volumes at different levels, each designed to meet a distinct need. At one level are a set of handbooks designed broadly to survey the major subfields of psychology; at another are numerous handbooks that cover important current focal research and scholarly areas of psychology in depth and detail. Planned as a reflection of the dynamism of psychology, the *Library* will grow and expand as psychology itself develops, thereby highlighting significant new research that will impact on the field. Adding to its accessibility and ease of use, the *Library* will be published in print and, later on, electronically.

The *Library* surveys psychology's principal subfields with a set of handbooks that capture the current status and future prospects of those major subdisciplines. The initial set includes handbooks of social and personality psychology, clinical psychology, counseling psychology, school psychology, educational psychology, industrial and organizational psychology, cognitive psychology, cognitive neuroscience, methods and measurements, history, neuropsychology, personality assessment, developmental psychology, and more. Each handbook undertakes to review one of psychology's major subdisciplines with breadth, comprehensiveness, and exemplary scholarship. In addition to these broadly-conceived volumes, the *Library* also includes a large number of handbooks designed to explore in depth more specialized areas of scholarship and research, such as stress, health and coping, anxiety and related disorders, cognitive development, or child and adolescent assessment. In contrast to the broad coverage of the subfield handbooks, each of these latter volumes focuses on an especially productive, more highly focused line of scholarship and research. Whether at the broadest or most specific level, however, all of the *Library* handbooks offer synthetic coverage that reviews and evaluates the relevant past and present research and anticipates research in the future. Each handbook in the *Library* includes introductory and concluding chapters written by its editor to provide a roadmap to the handbook's table of contents and to offer informed anticipations of significant future developments in that field.

An undertaking of this scope calls for handbook editors and chapter authors who are established scholars in the areas about which they write. Many of the nation's and world's most productive and best-respected psychologists have agreed to edit *Library* handbooks or write authoritative chapters in their areas of expertise.

For whom has the *Oxford Library of Psychology* been written? Because of its breadth, depth, and accessibility, the *Library* serves a diverse audience, including graduate students in psychology and their faculty mentors, scholars, researchers, and practitioners in psychology and related fields. Each will find in the *Library* the information they seek on the subfield or focal area of psychology in which they work or are interested.

Befitting its commitment to accessibility, each handbook includes a comprehensive index, as well as extensive references to help guide research. And because the *Library* was designed from its inception as an online as well as print resource, its structure and contents will be readily and rationally searchable online. Further, once the *Library* is released online, the handbooks will be regularly and thoroughly updated.

In summary, the *Oxford Library of Psychology* will grow organically to provide a thoroughly informed perspective on the field of psychology, one that reflects both psychology's dynamism and its increasing interdisciplinarity. Once published electronically, the *Library* is also destined to become a uniquely valuable interactive tool, with extended search and browsing capabilities, As you consult this handbook, we sincerely hope you will share our enthusiasm for the more than 500-year tradition of Oxford University Press for excellence, innovation, and quality, as exemplified by the *Oxford Library of Psychology*.

Peter E. Nathan
Editor-in-Chief
Oxford Library of Psychology

ABOUT THE EDITORS

Alexander Pollatsek is Professor Emeritus of Psychology at the University of Massachusetts, Amherst. His work in psychology has covered a broad range of topics, including object and scene perception and people's understanding of probability, statistics, and decision-making. "Dr. Pollatsek's primary interests, however, are in word processing, reading, and investigating how eye movements can be used to understand the reading process."

Rebecca Treiman is the Burke and Elizabeth High Baker Professor of Child Development Psychology at Washington University in St. Louis. Dr. Treiman's work focuses on the cognitive and linguistic processes involved in reading and spelling, particularly how children learn these skills. She has studied a variety of topics, including children's invented spellings, the characteristics of English and other writing systems, and the special difficulties faced by children with dyslexia and those who are deaf.

CONTRIBUTORS

Stephanie Al Otaiba
Professor and Patsy and Ray Caldwell Centennial Chair in Teaching and Learning
Annette Caldwell Simmons School of Education and Human Development
Southern Methodist University
Dallas, TX

Sally Andrews
Professor
School of Psychology
University of Sydney
Sydney NSW, Australia

David A. Balota
Professor
Department of Psychology
Washington University in St. Louis
St. Louis, MO

Molly S. Berenhaus
PhD Student
School of Psychology
University of Sussex
Falmer, UK

Kate Cain
Professor
Department of Psychology
Lancaster University
Lancaster, UK

Markéta Caravolas
Senior Lecturer
School of Psychology
Bangor University
Bangor, UK

Carol McDonald Connor
Professor
Department of Psychology
Arizona State University
Tempe, AZ

Anne E. Cook
Professor and Chair
Educational Psychology Department
University of Utah
Salt Lake City, UT

Holly K. Craig
Professor Emerita
School of Education
University of Michigan
Ann Arbor, MI

Anne E. Cunningham
Professor
Graduate School of Education
University of California, Berkeley
Berkeley, CA

Michael G. Cutter
PhD Student
School of Psychology
University of Southampton
Southampton, UK

S. Hélène Deacon
Professor
Department of Psychology and Neuroscience
Dalhousie University
Halifax, Nova Scotia, Canada

Denis Drieghe
Associate Professor
Department of Psychology
University of Southampton
Southampton, UK

Linnea C. Ehri
Distinguished Professor
PhD Program in Educational Psychology
CUNY Graduate Center
New York, NY

Ram Frost
Professor
Department of Psychology
Hebrew University
Jerusalem, Israel

Susan R. Goldman
Distinguished Professor
Department of Psychology
University of Illinois at Chicago
Chicago, IL

Jukka Hyönä
Professor
Department of Psychology
University of Turku
Turku, Finland

Debra Jared
Professor
Department of Psychology
University of Western Ontario
London, Ontario, Canada

Brett Kessler
Research Scientist
Department of Psychology
Washington University in St. Louis
St. Louis, MO

Sachiko Kinoshita
Associate Professor
ARC Centre of Excellence in Cognition and its Disorders and Department of Psychology
Macquarie University
North Ryde, NSW, Australia

Régine Kolinsky
Research Director
Center for Research in Cognition & Neurosciences
Université Libre de Bruxelles
Brussels, Belgium

Xingshan Li
Professor
Institute of Psychology, Chinese Academy of Sciences
Beijing, China

Simon P. Liversedge
Professor
Department of Psychology
University of Southampton
Southampton, UK

Jane V. Oakhill
Professor
School of Psychology
University of Sussex
Falmer, UK

Edward J. O'Brien
Professor
Department of Psychology
University of New Hampshire
Durham, NH

Colleen Ryan O'Donnell
Researcher
Institute of Human Development
University of California, Berkeley
Berkeley, CA

Bruce F. Pennington
John Evans Professor
Department of Psychology
University of Denver
Denver, CO

Manuel Perea
Professor
Departamento de Metodología
University of Valencia
Valencia, Spain

Robin L. Peterson
Assistant Clinical Professor
Department of Rehabilitation Medicine
Children's Hospital Colorado and University of Colorado School of Medicine
Aurora, CO

Alexander Pollatsek
Professor Emeritus
Department of Psychological and Brain Sciences
University of Massachusetts Amherst
Amherst, MA

Keith Rayner[†]
Atkinson Family Professor of Psychology
Department of Psychology
University of California, San Diego
La Jolla, CA

Erik D. Reichle
Professor
Centre for Vision and Cognition
Department of Psychology
University of Southampton
Southampton, UK

Eyal M. Reingold
Professor
Department of Psychology
University of Toronto at Mississauga
Mississauga, Ontario, Canada

[†] Keith Rayner died as the handbook was going to press. As this volume makes clear, he made many important contributions to the field of reading research.

Anna Samara
Research Associate
Division of Psychology and Language Sciences
University College London
London, UK

Elizabeth R. Schotter
Postdoctoral Scholar
Department of Psychology
University of California, San Diego
La Jolla, CA

Monique Sénéchal
Professor
Department of Psychology
Carleton University
Ottawa, Ontario, Canada

Heather Sheridan
Postdoctoral Researcher
Centre for Vision and Cognition
Department of Psychology
University of Southampton
Southampton, UK

Catherine E. Snow
Patricia Albjerg Graham Professor
Harvard Graduate School of Education
Cambridge, MA

Erin Sparks
PhD Student
Department of Psychology and Neuroscience
Dalhousie University
Halifax, Nova Scotia, Canada

Adrian Staub
Associate Professor
Department of Psychological and Brain Sciences
University of Massachusetts Amherst
Amherst, MA

Marcus Taft
Professor
School of Psychology
University of New South Wales
Sydney, NSW, Australia

Rebecca Treiman
Professor
Burke and Elizabeth High Baker
Department of Psychology
Washington University in St. Louis
St. Louis, MO

Anna M. Woollams
Senior Lecturer
Neuroscience and Aphasia Research Unit
School of Psychological Sciences
University of Manchester
Manchester, UK

Melvin J. Yap
Associate Professor
Department of Psychology
National University of Singapore
Singapore

Chuanli Zang
Associate Professor
Academy of Psychology and Behavior
Tianjin Normal University
Tianjin, China

CONTENTS

Part Three • Reading Sentences and Texts

Part Four • Reading and Spelling Development

Part Five • Reading Instruction

PART 1

Introduction

CHAPTER

1 The Oxford Handbook of Reading: Setting the Stage

Alexander Pollatsek *and* Rebecca Treiman

Abstract

This chapter serves as an introduction and outline for the remainder of the handbook. After first considering the importance of writing and reading in society, an overview of the history of reading research is provided. The chapter discusses the impact of the cognitive revolution, which began in the 1950s, on both basic reading research and research on the teaching of reading. The conclusion of the chapter summarizes the five sections of the handbook: Introduction, Word Identification, Reading Sentences and Texts, Reading and Spelling Development, and Reading Instruction. Previews are provided of the chapters in each of the sections, and the general themes and issues in the chapter are discussed.

Key Words: reading, writing, reading instruction, cognitive revolution, behaviorism, eye movements, response time, linguistics, cognitive neuroscience, dyslexia

In 1859, Abraham Lincoln gave a speech on the inventions that were most significant in human history. After considering such things as clothing, stone tools, and wheels, he settled on the idea that "writing—the art of communicating thoughts to the mind, through the eye—is the great invention of the world" (Fehrenbacher, 1989, p. 7). (Lincoln did admit that the noninvented way of communicating thoughts to the mind through the ear and gestures was a superior way of communication for many purposes.) Writing is great, Lincoln stated, "in the astonishing range of analysis and combination which necessarily underlies the most crude and general conception of it—great, very great, in enabling us to converse with the dead, the absent, and the unborn, at all distances of time and space" (p. 7). Lincoln went on to claim that without writing, civilization as we know it and democracy itself would be impossible. When people could write, and when people could read writing, important observations were written down, taken in, and reflected upon; "the seeds of invention were more permanently preserved, and more widely sown" (p. 9).

Of course, writing would be useless if no one could read it. Lincoln speculated that there must have been special circumstances where the inventors of writing could both develop a system of symbols and be sure there was a group of people who would be able to learn to decipher them. He also speculated about the circumstances that would have led to the development of the alphabet and made what would perhaps now be a controversial assertion that an alphabetic system is superior to writing systems like Chinese, which he probably misunderstood as being simple pictographic systems. Lincoln also discussed the invention of the printing press in his speech. Indeed, the combination of an alphabetic writing system and printing was an important part of bringing reading material to the masses.

Given the importance of writing and reading for modern societies, a large body of research has examined how people read and how they learn to read. Our handbook is an attempt to convey the state of the art in this exciting and important area of research. We begin this introduction with a historical overview of reading research. Systematic research on reading does not reach quite as far back as Lincoln, but it has been going on for over 100 years. We then focus on theories and controversies about how reading should be taught. After this, we describe the sections of the handbook and the chapters that they contain.

Early Research on Reading

Perhaps the culmination of the early research on reading was the publication of Huey's *The Psychology and Pedagogy of Reading* in 1908, which was reprinted in 1968. The topics covered in the book were quite similar to many of those in the present handbook and other current work on the psychology of reading: word recognition processes, inner speech, reading rate. Moreover, even though many of the findings discussed in Huey's book were obtained using archaic equipment, most have been shown to be basically valid. Thus, the conclusions drawn in Huey's book about reading are still worth considering. Another, more specialized, branch of research from this era is research on how much information can be extracted from the page from a single *fixation*; that is, a single glance. This research (e.g., Cattell, 1886; Erdmann & Dodge, 1898) was done using a device called a *tachistoscope*, which could present a visual display briefly for very precise amounts of time. Now this research can be done much more easily with computers and, in the context of people reading text, using eye-movement recording.

Psychological research in this vein more or less died out during the period when behaviorism dominated psychology, between about 1920 and 1960. The behaviorist philosophy, most prominently advocated by B. F. Skinner, promoted investigation of stimuli—the input to the organism—and behaviors—the output. To look at anything in between, including a hypothetical process such as recognizing a word, was thought to be misguided. With respect to reading research, this line of work perhaps culminated with the publication of Skinner's book *Verbal Behavior* (1957). In this book, consistent with behavioral psychology, reading and its development were analyzed largely through readers' overt behaviors and whether these behaviors were rewarded in some fashion or not. Consistent with the title of his book, Skinner also discussed the development of speaking.

The Cognitive Revolution and Reading Research

A major breakthrough in studying the reading process started around the late 1950s, coinciding with what is often called the *cognitive revolution*. Three works were particularly influential in this revolution. Two were intended for academic audiences—one as an advanced textbook and the other as a research monograph—and we will return to them shortly. The third appeared in a biweekly magazine, *The New York Review of Books* (Chomsky, 1971). Its title, "The Case Against B. F. Skinner," was a full-scale attack on a behavioristic approach to studying language. This attack came from someone who was and still is widely viewed as one of the world's most eminent linguists. The article did not provide any experimental data, but did convince many people of the shortcomings of the conceptual tools of the behavioral approach in dealing with understanding syntax and higher-order language structures and the limitations of the research that the behaviorists were providing on reading.

The other two works, by Broadbent (1958) and Neisser (1967), were part textbooks and part research monographs. As such, they reviewed quite a bit of the research in cognitive psychology that had emerged at the point that the books were written. Indeed, Neisser's book is usually credited with either inventing or at least publicizing the term "cognitive psychology." Neither book is specifically about reading, although both essentially start out by discussing how visual stimuli are encoded and recoded into other forms, primarily auditory. A central idea promoted in both books is that stimuli enter the organism (the discussion is almost completely about humans), but then invariably get recoded. This recoding is represented schematically by boxes and arrows, where a box represents a place in the mind, and presumably also in the brain, where information is represented in a particular format. For example, visual information about a letter, such as ‹s›, might come into the eye but get transformed into a form representing sound information, such as that it makes a hissing type of sound. These transformations from the sensory organs to deeper layers of the processing system are represented by arrows in the diagrams in the models of cognitive theories. These models usually had several layers of coding systems or boxes between the sensory organ and the point when one gets to something of interest to reading researchers, such as a word detector. The cognitive theories tried, if possible, to be consistent with what is known about the brain and the nervous system, but they did not limit themselves to that.

Perhaps the most crucial aspect of the cognitive revolution was to go beyond thinking about what one was observing as behavior to trying to analyze it as the product of a series of hypothetical mental stages or boxes between the stimulus input and the response. While the early cognitive research usually included only rough arguments tying these black boxes to brain areas or other parts of the nervous system, considerable work is now being done in a field called *cognitive neuroscience* to try to localize these hypothetical functions in the nervous system. This field uses brain imaging techniques, and it also examines electrophysiological data (measurements of electrical activity on the scalp in order to make inferences about activity in the brain) and data from patients with damage to particular portions of their brains. Of particular interest in the field of reading are *acquired dyslexics*, adults who have lost some or all of their reading skills as a result of brain damage.

The cognitive revolution led to many new ingenious new techniques for studying reading and word identification. For example, researchers developed the *lexical decision task*, in which people decide as rapidly as possible whether a sequence of letters is a word, and *priming* techniques, which compare performance on a word when the same word or a similar word was presented just previously and when a different word was presented. Participants' time to perform such tasks is measured to millisecond accuracy. Indeed, *reaction time* (RT) measures have played a central role in cognitive research on reading and other topics. Another new method paired the eye-movement techniques that had been used at the turn of the twentieth century with computer systems and video devices that can display text at precise time intervals in order to be able to study how people read text on a moment-to-moment basis. Researchers in the cognitive tradition have developed detailed models of the normal reading process in skilled adult readers in several languages. They have also examined the cognitive processes of children learning to read, both normally developing children and children with *developmental dyslexia* (serious and specific difficulties in learning to read). While all the authors of in this volume may not label themselves cognitive psychologists or cognitive scientists, we think that all of their work has been strongly influenced by the techniques and the ways of asking questions that the cognitive revolution brought to studying human psychology. Some of the authors have also been strongly influenced by cognitive neuroscience, and their chapters include evidence from this tradition as well.

The Teaching of Reading

The cognitive revolution sparked research on many topics, including attention and memory, but one of the most actively researched topics was reading. Part of the reason for this was the recognition that basic research on reading and reading development has the potential to better society by providing a research base for reading instruction. Rather than being based on tradition or ideology, decisions about reading instruction could be based on research findings. Research on reading had been underway before the advent of the cognitive revolution, but that development brought a new rigor to the field and a new emphasis on what was happening in the minds of individual learners.

Different methods of instruction have been used ever since writing was invented, and the choice of method depends in part on the nature of the writing system. For alphabetic writing systems, *phonics* approaches emphasize teaching children to decode words using connections between letters and sounds. For example, English-speaking children may have a lesson about the vowel group *ee* in which they work with words such as *see, need*, and *feet*. Phonics instruction is designed to systematically cover the important rules of a writing system. In the *whole-word* method, children are encouraged to memorize the pronunciations of entire words rather than to attempt to sound them out. This resolves the problem that would otherwise occur with a word like *been*, which in many dialects of English does not fit the rule about the pronunciation of *ee* that is taught in phonics. Another tension in reading instruction concerns the balance between different levels of language. According to the *whole-language* approach, which was popular in the United States during the 1980s and 1990s (Goodman, 1986; Smith, 1971), children should be exposed to good literature from the beginning and should focus on the meaning of what they read rather than on individual words or parts of words. In whole-language instruction classrooms children are encouraged to use context, background knowledge, and the pictures in a book to guess the identity of new words and to confirm their guesses. Phonics cues are a last resort, to be used only when children cannot identify a word based on other clues.

In the United States, several influential reports (National Reading Panel, 2000; Snow, Burns, & Griffin, 1998) have reviewed the research on early reading and reading instruction. One of the findings of these reports was that systematic phonics instruction helps children learn to read and should be an essential part of an early reading curriculum. This recommendation has been influential, and approaches that include phonics instruction are on the rise. There are still many questions about the nature of that instruction, however. For example, which rules should be taught, and in what order? How should exceptions to taught rules be dealt with? Moreover, phonics is not the only important part of an early literacy curriculum. Children who have learned to decode words well may have difficulty understanding what they read for other reasons. Basic research on these and other topics has an important role to play in shaping education.

Overview of the Handbook

We have divided this handbook into five sections. The Introduction section, of which this chapter is a part, sets the stage for the remaining chapters. It provides an overview of writing systems followed by an overview of research on how individual words are identified, and it closes with an introduction to how eye movements are used to study reading of sentences and larger units of text. In the next section, Word Identification, the focus is on adult skilled readers identifying individual words in isolation. Data from some other groups, including adult readers with brain damage, are also included in some of the chapters. Some of the chapters in this section include comparisons across languages as well. Most of the chapters of the third section, Reading Sentences and Texts, also focus on skilled readers, and most of them use data from the eye-movement techniques introduced in the final chapter of the Introduction. The issues discussed in this section range from the role of phonology in reading to processing of discourse and modeling of the pattern of eye movements in reading. There are again some comparisons across languages, including between English and Chinese. The fourth section, Reading and Spelling Development, is primarily about how literacy skills develop in children. The chapters in this section consider the development of word identification, spelling, and comprehension. In addition, one chapter examines how learning to read and spell influences other linguistic and cognitive abilities. This section of the handbook also includes discussion of cross-linguistic similarities and differences in literacy development and of children with developmental dyslexia. Although some of the chapters in the Reading and Spelling Development section address implications for instruction, the final section of the handbook covers this topic in more detail. It considers research and policy involving populations ranging from preschool age to adolescence, and it addresses what teachers should know and do to meet the needs of various groups.

Introduction

Chapter 2, by Brett Kessler and Rebecca Treiman, starts with writing, which as Lincoln noted is the starting point for reading. The chapter explores the diversity of modern writing systems, and it emphasizes that all writing systems, including those that are sometimes considered pictographic, are to a large extent representations of the linguistic units of a particular language. The chapter ends by discussing the implications of the design of writing systems for reading and learning to read. Chapter 3, by Melvin J. Yap and David A. Balota, gives an overview of the research on one of the most fundamental questions in reading research: How are written words deciphered? This question has been studied both by having people read words in isolation and by examining reading of connected discourse. The advantage of the former methodology is that it allows greater control of the materials and the situation, and the advantage of the latter methodology is that it is closer to a normal reading situation. The research in Chapter 3 relies more heavily on the former type of methodology, and it examines such issues as how the frequency and length of words in the language affect the time to identify and name them. Chapter 4, by Elizabeth R. Schotter and Keith Rayner, is an introduction to how eye-tracking technology has allowed experimenters to achieve a detailed record of the reader's progress through the text and to draw sophisticated conclusions about the cognitive operations of the reader. This includes inferences about how people encode words and how much information they obtain from a single glance at the page.

Word Identification

The focus of the chapters in this section of the handbook is on how individual printed words are identified. Chapter 5, by Sachiko Kinoshita, begins by discussing how the frequency of a word is related to the ability to identify it. Kinoshita presents a theoretical framework based on Bayesian decision-making in which this relationship can be viewed. Using this framework, the remainder of the chapter discusses how readers identify and code the order of letters in a word. The issue of *neighborhood* *effects*—how readers' ability to identify a word is altered by factors like the number of words that are spelled similarly —is followed up in Chapters 6 and 7. The focus of Chapter 6, by Manuel Perea, is on how the concept of a neighborhood should be defined. Focusing largely on alphabetic writing systems, Perea provides a historical overview of how the concept of neighborhoods evolved and shows how the definitions have become more sophisticated. The chapter also discusses whether the major effect of neighborhoods on word identification is facilitative or inhibitory. Chapter 7, by Ram Frost, considers whether the effects found in the bulk of the neighborhood research, which has largely been conducted in Indo-European languages with alphabetic writing systems, hold true across all languages. Using Hebrew as a comparison language, Frost concludes that even some simple effects, such as letter-position effects in word recognition, show profound cross-linguistic differences. Chapters 8 and 9 are both about *multimorphemic* words, namely, words that contain more than one unit of meaning (*morpheme*). Chapter 8, by Marcus Taft, focuses on research using methods in which single words are presented in isolation. Taft presents a detailed model for the encoding of multimorphemic words in reading, and he also discusses how this model links with models of the production of multimorphemic words in speech. The research on multimorphemic words reviewed in Chapter 9, by Jukka Hyönä, primarily uses methods where people read sentences and where eye-movement measures are the primary indices of performance. This chapter focuses on compound words. Much of the research reported is in Finnish and German, where there is no space between the elements of a compound word. Although the chapters by Taft and Hyönä converge on some basic findings, they indicate that there is still much to be learned about the processing of multimorphemic words, an area that has not been as heavily researched as the processing of single-morpheme words.

The section on words closes with three chapters examining word identification in different groups of readers. Most studies of word identification in skilled readers present data that are averaged across individuals, but Sally Andrews argues in Chapter 10 that there are systematic individual differences in word identification processes within groups of relatively skilled readers. The chapter reviews these differences and challenges the assumption that there is a single way in which all skilled readers identify words. Chapter 11, by Anna M. Woollams, reviews data from adult patients with acquired dyslexia who have very specific word-recognition and-production problems. She then presents and discusses the viability of several current theories of word recognition of normal readers in the light of these data. Chapter 12, by Debra Jared, is about readers who are bilingual. It examines how knowledge of one language and reading in one language influences reading in the other. Jared reviews evidence from skilled readers and also from developing readers, a topic that is more fully covered in a later section of the handbook. The emphasis in Jared's chapter, as in the other chapters in this section, is on the identification of individual words.

Reading Sentences and Texts

The chapters in this section of the handbook discuss normal adult readers negotiating their way through text, and most heavily rely on the eye-movement technology explained in detail in Chapter 4. The first three chapters in this section, 13, 14, and 15, are about basic issues in the reading process. Chapter 13, by Alexander Pollatsek, is about the role of sound or phonology in silent reading. This topic is also discussed in some of the models of word identification that are presented in Chapter 3 and in several of the chapters in the Word Identification section of the handbook. However, Chapter 13 also goes beyond the role of phonology in word encoding and discusses its possible broader role in reading, such as serving a memory function as well. Chapter 14, by Adrian Staub, also goes beyond how individual words are decoded and discusses how skilled readers process syntax. One of the key issues in this literature is how processing of syntax and meaning relate to each other. For example, do readers necessarily have to process the syntax of the part of the sentence read so far before deciding on the meaning of the fragment? Staub's chapter also reviews some electrophysiological measures of these phenomena. Chapter 15, by Edward J. O'Brien and Anne E. Cook, discusses larger structures of text, the entire story or discourse, and attempts to model the reader's mental representation of what is being read. The key question that this chapter attempts to answer is whether the reader is actively constructing a model of what the writer is about to discourse about or whether the process is more passive, such that discourse structures are activated from long-term memory by the text.

The following chapters in this section are focused on the interplay between the linguistic information being taken in by readers and how they progress through the text by moving their eyes. These chapters consider almost exclusively skilled adult readers. Chapter 16, by Xingshan Li, Chuanli Zang, Simon P. Liversedge, and Alexander Pollatsek, discusses this question in Chinese, where the answer is by no means obvious because Chinese characters are not grouped into perceptual units that correspond to words. The chapter provides evidence that the word (in addition to the character) is an important cognitive unit in Chinese reading and, among other things, guides how far the eyes move on a moment-to-moment basis. Chapter 17, by Michael G. Cutter, Denis Drieghe, and Simon P. Liversedge, focuses on issues that are important to understanding almost everything about the reading process: when the eyes move from one fixation to the next, what information is preserved from one fixation to the next, and how the information that is preserved combines with the information that is seen on the next fixation. Chapter 18, by Eyal M. Reingold, Heather Sheridan, and Erik D. Reichle, is a general discussion of models of eye control in reading, focusing on the durations of the fixations. It uses sophisticated modeling techniques to determine whether models of eye movements in reading can adequately explain that variables such as word frequency can affect readers' fixation durations as quickly as they do. The section Reading Sentences and Texts closes with Chapter 19, by Erik D. Reichle and Heather Sheridan. This chapter presents a detailed model of cognitive processes and the eye-control system that hopefully is a reasonable approximation of how this complex system works when skilled readers read text.

Reading and Spelling Development

Learning to read individual words is a central part of reading development, and the opening chapter of this section of the handbook, by Linnea C. Ehri, considers how children learn to do this. In Chapter 20, Ehri portrays the development of word-reading skill for learners of English and other alphabetic writing systems as a sequence of four phases. Each phase is characterized by use of a particular type of connection between a word's spelling and its pronunciation in memory. Learning to read words benefits from precise knowledge about their spellings, as Ehri discusses, and learning to produce spellings is an important part of literacy development as well. Chapter 21, by S. Hélène Deacon and Erin Sparks, discusses how children learn to spell words. The authors review research on this topic and evaluate it in light of different theories about the process. The emphasis, as in the chapter by Ehri, is on the learning of alphabetic writing systems. Broadening the picture, Markéta Caravolas and Anna Samara present information in Chapter 22 about spelling and reading development in several other types of writing systems as well. Whereas the earlier chapters in the Reading and Spelling Development section emphasize patterns that hold across children, the chapter by Caravolas and Samara focuses on differences across children. The authors present a model of the cognitive and linguistic skills that help make learning to read and spell words easier for some children than for others. In Chapter 23, by Jane V. Oakhill, Molly S. Berenhaus, and Kate Cain, the emphasis shifts from individual words to texts. Oakhill and colleagues outline the development of processes that are related to reading comprehension during the early school years. They also discuss individual differences in comprehension and children who have specific difficulties with this aspect of reading. Serious difficulty in learning to read is the focus of Chapter 24, by Bruce F. Pennington and Robin L. Peterson. These authors discuss how developmental dyslexia emerges in children, considering the roles of genes and the environment and the ways in which they interact. The final chapter in the section on development, Chapter 25 by Régine Kolinsky, looks at development from a different perspective by asking what skills and abilities are affected when a person develops the ability to read. Using data from adults who are illiterate and from people who learned to read as adults rather than as children, Kolinsky argues that learning to read influences the processing of spoken language and that it has influences outside the domain of language as well.

Reading Instruction

Some of the chapters in the Reading and Spelling Development section of the handbook include material discussing the implications of research for reading instruction, but the chapters in the final section of the handbook deal more directly with issues related to instruction and instructional policy. In literate societies learning to read often begins at home, when parents provide informal or even formal teaching about letters, the alphabet, or words. In addition, many parents read books to their children. Chapter 26, by Monique Sénéchal, discusses the literacy-related experiences that young children have at home and how they relate to later reading outcomes.

Once formal literacy instruction begins, around the age of six in many cultures, many questions arise about how to teach decoding and comprehension and what approaches work best for students who come to language different from that used in the classroom or students who encounter difficulty in learning to read for other reasons. In Chapter 27, Carol McDonald Connor and Stephanie Al Otaiba review research and policy pertaining to these and other issues in primary grade reading instruction in the United States. The particular case of African American students in the United States is the focus of Chapter 28, by Holly K. Craig. This chapter discusses the dialect that many of these children use, African American English, their developing ability to switch between African American English and Standard American English, and the relationships between African American English and reading achievement. The chapter also discusses how teachers deal with African American English. What teachers know about language and reading comes to the fore in Chapter 29, by Anne E. Cunningham and Colleen Ryan O'Donnell. The authors argue that teachers need a specific body of disciplinary and pedagogical knowledge in order to teach reading effectively and that they are not always provided with the opportunities they need to acquire this knowledge. This chapter focuses on teachers of young children, consistent with the emphasis on primary grade students in most of the chapters in the Reading Instruction section of the handbook. Chapter 30, by Susan R. Goldman and Catherine E. Snow, turns to the new challenges that adolescents face. Reading should become a tool for acquiring information, understanding points of view different from one's own, critiquing arguments, and reasoning, and literacy instruction for older students must be designed to make this possible.

Summary

The study of reading and its development has benefited greatly from use of a cognitive approach, and the research has some important implications for how reading should be taught. In editing this handbook, we have tried to gather a collection of chapters that are accessible to a range of readers and that present the diversity of the field. By surveying the achievements of the field, and by raising open questions, we hope that the handbook will spur understanding and promote additional research and application of research.

References

Broadbent, D. E. (1958). *Perception and communication*. London, England: Pergamon.

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind, 11,* 63–65.

Chomsky, N. (1971). The case against B. F. Skinner. *The New York Review of Books, 17,* 18–24.

Erdmann, B., & Dodge, R. (1898). *Psychologische Untersuchungen über das Lesen auf experimenteller Grundlage* [Psychological research on reading on an experimental basis]. Halle, Germany: Niemeyer.

Fehrenbacher, D. E. (Ed.). (1989). *Lincoln: Speeches and writings (1859–1865)*. New York, NY: Library of America.

Goodman, K. (1986). *What's whole in whole language?* Portsmouth, NH: Heinemann.

Huey, E. B. (1968). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press (Original work published 1908).

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institutes of Health.

Neisser, U. (1967). *Cognitive psychology*. New York, NY: Appleton-Century-Crofts.

Skinner, B. F. (1957). *Verbal behavior*. Acton, MA: Copley.

Smith, F. (1971). *Understanding reading*. New York, NY: Holt, Rinehart, & Winston.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

CHAPTER

2 Writing Systems: Their Properties and Implications for Reading

Brett Kessler *and* Rebecca Treiman

Abstract

An understanding of the nature of writing is an important foundation for studies of how people read and how they learn to read. This chapter discusses the characteristics of modern writing systems with a view toward providing that foundation. It considers both the appearance of writing systems and how they function. All writing represents the words of a language according to a set of rules. However, important properties of a language often go unrepresented in writing. Change and variation in the spoken language result in complex links to speech. Redundancies in language and writing mean that readers can often get by without taking in all of the visual information. These redundancies also mean that readers must often supplement the visual information that they do take in with knowledge about the language and about the world.

Key Words: writing system, script, alphabet, syllabary, logography, semasiography, glottography, underrepresentation, conservatism, graphotactics

The goal of this chapter is to examine the characteristics of writing systems that are in use today and to consider the implications of these characteristics for how people read. As we will see, a broad understanding of writing systems and how they work can place some important constraints on our conceptualization of the nature of the reading process. It can also constrain our theories about how children learn to read and about how they should be taught to do so.

Figure 2.1 shows examples of some of the writing systems that are used by modern societies. Each sentence expresses the sentiment 'I can eat glass' (Mollick, 1996). The visual similarity of Hebrew and Yiddish reflects the fact that those languages use the same *script*, or set of signs. Their similarity in appearance, or *outer form*, disguises the fact that the two languages are very different from each other and that their writing systems work differently in some respects. For example, in Hebrew, as we discuss later, many vowels are left unwritten, whereas that is not generally the case in Yiddish. Conversely, differences in outer form can disguise important similarities in how writing systems function, or their *inner structure* (see Gelb, 1952, for discussion of outer form and inner structure). As we will see, all systems of writing, even those written with different scripts, such as Classical Mongolian, Japanese, Hebrew, and Hindi, share some important properties, both properties related to their outer form and properties related to their inner structure. We discuss these commonalities in the first section of the chapter, focusing on those of most potential relevance for reading. That section is followed by an examination of some of the major differences across writing systems of the world. The final section of the chapter lays out several implications of the characteristics of writing systems for how people read and for the learning and teaching of reading.

Shared Properties Across Writing Systems

In this section we discuss some of the more important properties that are shared by modern

(1) 

(2) 私はガラスを食べられます。

(3) אני יכול לאכול זכוכית.

(4) איך קען עסן גלאָז.

(5) मैं काँच खा सकता हूँ।

Fig. 2.1 Writing system sampler. (1) Classical Mongolian. (2) Japanese. (3) Hebrew. (4) Yiddish. (5) Hindi.

writing systems. In order to keep the discussion uncluttered, we will occasionally skip over counterexamples from special-purpose writing systems, such as the tactile systems used by blind people, or from writing systems that are no longer in use. Our focus is on properties that are common across current writing systems that are used for purposes of general literacy.

Writing Is Visual

Writing normally takes the form of visible marks on the surface of a relatively permanent object. The fact that writing is designed to be taken in by the eye leads to some characteristics that are essentially universal in modern writing. We discuss these properties of writing's outer form in what follows.

Visual processing, whether of writing, spoons, or coins, requires a certain amount of time and attention. Perceivers must determine the number of visual objects that are present, the location of the objects relative to one another, and the identity of the objects. They must often act before they have had time to perform a full analysis of the incoming information. For example, readers may not take in all of the visual information that is available to them, perhaps not fixating on a word at all or not processing all of the letters in a word on which they do fixate. Readers may make decisions about which words are present before they have fully processed the information they have taken in. Some universal tendencies of writing systems have developed to make reading possible despite people's limitations in vision and attention, to allow them to do a reasonable job even when going quickly.

One property of modern writing systems that aids visual processing is that there is a reasonable degree of contrast among the basic elements of a script. The distinction between the closed curve shape of ‹O› and the open curve shape of ‹C› is easy enough to see, but one would not expect a writing system to include several different C-like shapes that differ only by minuscule differences in the size of the opening. Nor would one expect to find a writing system in which a version of ‹C› with a narrow line and a version of ‹C› with a broader line constituted two different letters. To use a term traditional among historians of writing (Evans, 1894), the symbols of modern writing systems are *linear*. They require no shading, no fill, no color other than that needed to distinguish the writing from the background, and no distinctions between lighter and darker lines or wider and narrower lines.

Another visual property of writing that helps people to distinguish its elements is that most scripts contain few or no elements that differ only in their left–right orientation. English has the mirror-image pair ‹p› and ‹q›, and Korean has ㅏ and ㅓ, but most writing systems do not include any pairs of this sort. This is probably because people find it hard to learn to assign objects that differ only in orientation to different categories (Kolinsky et al., 2011).

Yet another visual property of writing that helps both beginning and skilled readers is the visual redundancy within the elements of a script. In many cases, this redundancy can allow an element to be successfully identified even if some of its visual characteristics are overlooked. For example, a reader of English who fails to detect the crossbar on ‹A› can still identify the letter because there is no other letter ‹Λ› in the Latin alphabet. According to one estimate, in fact, the identity of the elements of modern writing systems can be determined when, on average, half of the strokes are removed (Changizi & Shimojo, 2005).

Within a writing system, the elements show a certain *stylistic consistency*. For example, Chinese characters, such as 圆 'round' and 球 'ball', are squarish and angular; they do not include full circles or semicircles as do a number of the letters of the Latin alphabet. The similarities among the elements of a script reflect the importance of the aesthetic qualities of writing, above and beyond the message that it conveys. The set of forms in (1a) is more pleasing to the eye than the set of forms in (1b) for example, (a) showing a set of Hebrew letters, which are stylistically similar to each other, and (b) showing Hebrew letters interspersed with Arabic letters. As we discuss later in the chapter, stylistic consistency can benefit readers, allowing them to develop a familiarity with the visual patterns that are shared across the elements of their writing system.

(1)

a. לוכאל
b. כسאفל

The aesthetic nature of writing means that writers may choose certain fonts or visual styles because of their beauty or the qualities they express, and readers must deal with these variations. For example, 𝔄𝔟𝔢𝔯𝔡𝔢𝔢𝔫 ℭ𝔬𝔲𝔯𝔱 may appear on a street sign not because this font is easy for modern readers to read, but to convey adherence to tradition. The aesthetic benefits come with a cost: Readers must learn to deal with different forms of symbols, placing ‹𝔄› in the same category as ‹A›. For a child, it may be by no means obvious that these forms belong together.

Writing Represents Language

Communicating by visual marks on surfaces can work in any number of ways. For example, people can send a message by drawing or painting ad hoc images. This technique could, in principle, permit communication among people who do not share a language. However, communicating by drawing pictures is difficult and error-prone. Communication can be improved tremendously by using *symbols*: images that have conventionally assigned specific meanings for a specific group of users. Figure 2.2 shows some symbols that were developed in the 1970s to clearly and unambiguously direct travelers in the United States to essential services (AIGA, 1989). It is equally easy for speakers of any language to learn that these symbols direct travelers to departing flights, arrivals, and coffee shops, respectively. But the disadvantage of having a separate symbol for everything of interest is that life has an unbounded number of things of interest. Symbol sets such as those in Figure 2.2 almost never grow to more than a few hundred. This is because the effort in learning large numbers of symbols does not pay off in expressive power.

One improvement to a set of symbols is to add a *grammar*, a set of rules for expressing relations between symbols. These rules allow users to express concepts that vastly outnumber the individual symbols. The grammar of musical notation, for example, allows musicians to represent an indefinite number of musical pieces with great accuracy using a fairly small set of musical symbols. An even more familiar system is mathematical notation. The positional system, where every digit weights a different power of ten, lets people precisely indicate any natural number using no more than ten symbols, the digits. The two symbols ‹98›, for example, represent the number obtained by adding nine tens to eight because part of the grammar says that the second position from the right is the tens place; ‹89› represents a quite different number. Communication systems such as these, which include a grammar for combining symbols, are referred to by a special term in many languages, such as *writing* in English. People *draw* symbols of departing airplanes, but they *write* sheet music and mathematical formulas.

Yet even powerful systems such as those for music and math are limited to particular domains. Many philosophers have labored hard to develop grammars and symbol sets that could visually represent ideas regardless of their domain—numbers, dining preferences, political humor—but none have caught on. Such systems are difficult to learn, and they are still incapable of representing with any precision everything that people are accustomed to talk about. Writing in the sense that we discuss it in this chapter is an ingenious solution to the problem of generality. Instead of trying to directly represent the concepts that people can talk about, it represents the words of language itself. A writing system that can represent words accurately can represent anything people can talk about. It is truly a general-purpose system of communication.

Systems such as musical and mathematical notation have been characterized as *semasiographic*, literally, 'writing ideas' (Gelb, 1952, p. 11). General-purpose writing, in contrast, has been called *glottographic*, literally, 'language writing' (Pulgram, 1976, p. 4). The distinction between the

Fig. 2.2 Symbols developed by AIGA for the US Department of Transportation. Images courtesy of AIGA.

two types of systems can be appreciated by comparing the semasiographic ‹98› with glottographic forms like the English ‹ninety-eight› or the French ‹quatre-vingt-dix-huit›. The semasiographic form follows the grammar of mathematical notation and works the same way around the world. In particular, it can be read off in innumerable different ways, depending on the language one speaks. The English and French spellings, in contrast, represent words in those languages, with specific pronunciations that would be meaningless to people who have not learned those languages. Although writing systems have many differences in their outer form, as we saw in Figure 2.1, all full-scale systems in regular use are glottographic. Semasiography is sometimes mixed in, as when we write about "101 Dalmatians," but glottography is at the core.

If writing represented ideas directly, we might expect it to look like some sort of semantic network with nodes spread out all over the page, lines connecting the nodes, and so forth. But writing represents the words of a language, and the most natural way to do so is by arranging along a line a sequence of symbols representing units of language. This is what all modern writing systems do. We call the symbols that are laid out along a line the *characters*. Thus, ता is a single character in Hindi (representing the syllable [tɑ]) even though, as we discuss later, it includes one part (त) that stands for [t] and another part (the vertical bar on the right) that stands for [ɑ].

Stretches of speech can be indefinitely long, and writing media are decidedly finite. At some periods in the past, it was conventional when reaching the edge of a page to begin the next line directly underneath the end of that line. Writing would then proceed in the opposite direction, usually with all the letters reversed, so as to preserve to a reasonable extent the sequentiality between the two lines. In all modern writing systems the convention is now to write in straight lines until one runs out of space, then go back and begin another straight line in the same direction. This use of parallel straight lines is probably the best compromise between the desire to write words in their natural sequence and the desire to always present word symbols in the same orientation.

Although all general-purpose writing systems use symbols to represent language, language is complicated and open-ended. This leaves designers of writing systems many alternatives for mapping symbols onto language. As Martinet (1960) put it, language has a *double articulation*. Martinet's first articulation deals with words; to give it a more memorable name, we will call it the lexical level. At the lexical level, we can think of a sentence as being a sequence of words. Words, in turn, are sometimes composed of smaller meaningful units called morphemes. Thus *Kim's girlfriend uses a prepaid card* breaks down into the sequence shown in (2a); we use square brackets to enclose each word, some of which contain comma-separated morphemes. Martinet's second articulation, the phonetic level, deals with sounds. At this level, a sentence can be thought of as a sequence of syllables, which are in turn composed of phonemes; in (2b) we use brackets to enclose each syllable. Phonemes, in turn, can be described in terms of their *distinctive features*, the differences that distinguish them from each other in a given language. For example, /k/ is a voiceless velar stop, distinct from the voiced velar stop, /ɡ/. Each of Martinet's levels has its own grammar—the rules for combining words to make a clause have nothing in common with the rules for combining phonemes into syllables.

(2)

a. [Kim, 's] [girl, friend] [use, es] [a] [pre, pay, ed] [card]
b. [k ɪ m z] [ɡ ɚ l] [f r ɛ n d] [j u] [z ɪ] [z ə] [p ɹ i] [p e d] [k ɑ ɹ d]

The dual articulation of language pulls designers of writing systems in two directions. They could invent symbols at the lexical level, making a distinctive symbol for each word or morpheme. Such symbols are called *logograms*. Another choice is to focus on the phonetic level, making a distinctive symbol for each phoneme or each syllable. Such symbols are called *phonograms*. But even the most phonetic of modern writing systems do something special to stabilize word symbols and make them salient.

One way phonographic writing systems highlight the lexical level is through *lexical constancy*. In modern writing systems, there is typically only one conventional representation of a word: A horse is always a ‹horse› in English. A related property is *lexical distinctiveness*: ‹horse› represents the word *horse* and no other word. When lexical constancy and distinctiveness are present, it is easy for readers to directly map from the written form to the corresponding word. These properties attract particular attention in writing systems in which a given phoneme is not always spelled with the same letter. No basic phonetic spelling rule of English precludes spelling *horse* as ‹horce› (as in ‹force›) as well as ‹horse›, but lexical constancy has encouraged English spellers to stick with a single spelling. Inconsistent mappings also permit lexical distinctiveness among homophones: *horse* is always distinguished from *hoarse*, even though the two are pronounced alike. There is some question as to whether distinguishing homophones is a primary principle of writing systems—there are, after all, exceptions like *sow* (either /so/ 'plant seeds' or /saʊ/ 'female swine')—or whether it is a side effect of the conservatism of writing that we discuss later in this chapter. Spellings often stay the same even when pronunciations change, and the different vowel letters in ‹horse› and ‹hoarse› reflect the fact that the vowels used to be pronounced differently. But regardless of the origin of lexical distinctiveness, the reader is presented with an additional cue to identifying which homophone the writer intended.

Another way most modern phonographic writing systems highlight the lexical level is through *lexical demarcation*: explicit, visual means of showing where one word ends and the next one begins. Nowadays, almost all languages separate words physically from each other with spacing or other marks of punctuation, such as the colon-like symbol traditionally used in Amharic. Scripts sometimes demarcate words by other visual methods, as when Hindi writes most words with a solid bar at the top which connects with that of the adjacent elements in the same word, but, in modern times, not with the neighboring words (Figure 2.1, example 5). The bar reinforces the impression that the word is a unit. Similarly, Mongolian script (Figure 2.1, example 1) as well as Arabic joins letters within a word but not between words, a tactic used in Europe in many cursive handwriting styles but rarely in print. More subtly, Japanese does not space between words, but most content words start with a kanji character and end with a hiragana character, two types of character that look very different from each other. Thus, the transition between a simple, curvy hiragana character and a more complex, angular kanji character (as between を and 食 in Figure 2.1, example 2) usually is a visual marker of a lexical boundary. Not all writing systems have all three properties of lexical constancy, lexical distinctiveness, and lexical demarcation, but most have at least one.

Because writing represents language, its outer form is constrained both by the properties of speech in general and the properties of the specific language that it represents. One property of speech is that it is not highly repetitive. Repetition rarely goes beyond doubling. At any level, whether going by individual phonemes, syllables, or words, it is rare to see three instances of the same unit in a row. Because writing represents speech, this degree of repetition is uncommon in writing as well. We might accept that ‹Lorem ipsum dolor sit amet› means something in some language, but we would be very surprised to learn that there was a language in which ‹Looooremmmm iiiiipppsoooo› was a normal way of writing anything. In contrast, in mathematical notation ‹1,000,000› is a well-formed and unsurprising number.

Another general property of language that is reflected in the outer form of writing is its redundancy. Not all of the sounds in a word or all of the words in a message need to be processed in order to make sense of the message. Because writing represents language, it has some of the same redundancy. For example, standard French indicates negation by placing *ne* before the verb and *pas* after it. A reader who fails to perceive the word *ne* in *Je ne sais pas* 'I don't know' can still determine that the sentence is negative because of the *pas*. Or, because there is a word *mariposa* 'butterfly' in Spanish but not a word *maruposa*, a reader can identify the word without resolving whether the second vowel letter is ‹i› or ‹u›.

Writing also reflects many of the properties of the specific language that it represents. For example, it would be rare to see ‹the› at the end of an English sentence or を at the beginning of a Japanese sentence; this Japanese symbol in reality appears only immediately after the direct object of a verb. Likewise, readers of English would notice something peculiar about a word beginning with ‹ng›, ending with ‹pv›, or totally lacking a vowel letter. Such words are odd because the sounds they would represent violate the *phonotactics* of English, namely, the constraints on the ordering and position of phonemes in words. English words cannot begin with [ŋ], end with [pv], or lack a vowel sound. There should be no reason to come up with the spellings that represent these sounds, unless one is designing an experiment to tap people's processing of illegal sequences. Although the restriction against words like ‹ngim›, ‹lupv›, and ‹scvnkls› is motivated by the phonotactics of English, nothing would stop the beginning reader from learning about such restrictions as *graphotactic* irregularities—namely, those that violate the normal patterns by which letters are assembled.

At the same time, other graphotactic violations have little or nothing to do with the phonotactics of a language. For example, ‹hevvi mettl› could arguably be used as an improved spelling of *heavy metal* [ˈhɛvi ˈmɛtl̩] in English were it not for the fact that

English spelling generally avoids ‹vv›, word-final ‹i›, and word-final ‹l› after a consonant. As this example shows, writing takes on a life of its own. Although readers normally see through the visual forms to the language they represent, writing is a visual system with its own patterns.

Writing Does Not Represent All Aspects of Language

Writing represents a language, but it does not represent all aspects of the language. With rare exceptions, writing systems only represent distinctions that result in *lexical contrast*. For example, Spanish has the sound [d] (as in the English *den*) and the sound [ð] (as in the English *then*), but the choice between them is determined entirely by factors such as whether they come after a vowel. There can never be a situation where replacing one sound with the other could result in a new word. Thus, [d] and [ð] are *allophones* of the same phoneme in Spanish. Writing systems rarely have different symbols for different allophones of the same phoneme, and indeed Spanish [deðo] 'finger' is spelled ‹dedo›. The failure to distinguish allophones never seems to even be noticed by native speakers of a language, and it probably benefits them by reducing the number of symbols they must learn and distinguish when reading.

More surprisingly, it is not at all uncommon for writing systems to fail to represent distinctions that are lexically contrastive. Sometimes this *underrepresentation* is due to a change in a language's phonology that makes allophones become phonemes. For example, English spelling fails to differentiate the phoneme /θ/ (as in *thumb*) from /ð/ (as in *then*), in part because the two used to be allophones. In modern English they have become separate phonemes, but apparently there has not been sufficient confusion to prod people into inventing a new spelling for one or both of the sounds. In Hebrew, the letter פ once stood for the phoneme that had the allophones [p] and [f]. In modern times, these are now two separate phonemes, /p/ and /f/, but both are still written with the same letter. The fact that the same spelling can represent two different phonemes makes reading somewhat more complicated, but knowledge about words (e.g., /θɛn/ isn't an English word but /ðɛn/ is) and the use of discourse context help readers. Generally speaking, people are conservative and fairly tolerant of inconsistencies in a writing system, resisting spelling reforms that would break with tradition and require them to learn new spelling rules.

A related historical factor has to do with what writing system another writing system was borrowed from. People who devise a writing system by adapting that of another culture are often reluctant to change it greatly, even if ambiguities and inconsistencies arise. A great many modern writing systems use the Latin script, which, in its original and most widely disseminated form, has no basic way of representing more than five different vowel qualities. There is no direct way of representing vowel length, stress, tone, or virtually any sound that wasn't found in Latin. If the Latin script had had ways of representing these things, possibly many more languages today would as well. Instead, languages that use the Latin script, such as English, often let phonemic distinctions go unexpressed—‹wind› is either 'moving air' with a short vowel or 'turn coils' with a long vowel—or rely on often indirect ways of distinguishing them, such as consonant doubling—‹planing› /e/ versus ‹planning› /æ/—or silent ‹e›, as in ‹bite› /aɪ/, as opposed to ‹bit› /ɪ/. *Digraphs* (two-letter sequences that represent a single phoneme) are also widely used to supplement the alphabet, often with values that are hard for readers to deduce from their composition. For example, without considering its history, it is not obvious why ‹au› in English *taut* spells a vowel that has little to do with the common values of either ‹a› or ‹u›. Other languages that use the Latin script, such as Czech, have added *diacritics* to certain letters to express phonemic distinctions. Adding a small mark to a letter, such as the diacritic on ‹č›, comes across as a smaller change than creating a wholly new letter.

Another example of the effect of borrowing is afforded by the many other writing systems that use scripts derived from ancient Aramaic, which had no mechanism for writing short vowels. The most widespread of these scripts are Arabic and Hebrew. To this day, Hebrew and most of the languages that use the Arabic script, including Persian and Urdu, do not represent some of their vowels in the texts that are typically read by adults. Figure 2.3 illustrates how Hebrew writing omits representation of some but not all of its vowels. Hebrew and Arabic do have methods of representing all vowels, but these are used only in special situations, including texts for beginning readers. It is also quite likely that Hindi and related scripts of South Asia descended from Aramaic or one of its precursors. They mostly write vowels as diacritics and they leave out one vowel entirely. This suggests that vowel symbols may be late additions to the script. In Hindi, for

א∅ני י∅כול ל∅א∅כול זכוכית

t i χ u χ z   l o χ e ʔ e l   l o χ a j   i n a ʔ

Fig. 2.3 Omission of some vowels in Hebrew writing. Reading right to left, the Hebrew text from Fig. 2.1 (3) is aligned with its pronunciation [ʔani jaχol leʔeχol zχuχit]. ∅ is inserted in the Hebrew text where a vowel has no overt representation.

example, the vowel /ə/ is not represented: The word सकता /səktɑ/ 'can' in Figure 2.1, example 5, has overt representations for /s/ स, /k/ क, and /t/ त, and a diacritic for /ɑ/, the vertical bar to the right. However, it does not have anything overt for /ə/. While it is tempting to speculate on why some languages do not write many of their vowels, history has furnished us not with a natural experiment but with little more than a single anecdote: There is a single lineage of scripts that omit vowels. Languages that inherited scripts that descend from Aramaic may simply be retaining its ancestral properties because cultures tend to be fairly conservative about reforming their writing systems.

Certain features of pronunciation appear to be especially susceptible to underrepresentation in writing. All languages have *intonation*, using variations in pitch and timing over the range of a sentence to indicate attitude and emotion, to focus attention on part of a statement, and so forth. Intonation can be an important cue in understanding speech. But no writing system, to our knowledge, represents intonation in all but the most rudimentary and ad hoc way. The modern system of punctuation that has been adapted by most scripts is better than nothing—it can occasionally help the reader discern whether a panda who "eats shoots and leaves" is an herbivore or carries a gun ("eats, shoots, and leaves"; Truss, 2004)—but no language employs it to mark all aspects of intonation. Furthermore, much of the small set of symbols that is used for punctuation marks features other than intonation. For example, ‹?› marks a sentence as interrogative, regardless of its intonation. This widespread inattention to intonation may be due in part to the fact that writing systems focus on words. Representing words in different ways when they appear in different intonational patterns would undermine lexical constancy.

Writing systems are better at representing pitch differences when they help identify words, but they are not much better. Lexically contrastive pitch is called *tone*, and it is a feature of the majority of the world's languages. For example, in Hausa, one of the dominant languages of Africa, /baː˩ba˥/ 'father' and /baː˥ba˩/ 'mother' differ only in the tones. 'Father' has a low tone in the first syllable (indicated by the ˩ after the first vowel) and a high tone in the second (indicated by the ˥ after the second vowel). 'Mother' has the opposite pattern. Despite the fact that tone is very important in Hausa and there are established conventions for writing tone with diacritics, tone marks are almost never written in Hausa. The situation with Hausa is not unique: Kutsch Lojenga (2011) reported that speakers of tonal languages throughout Africa have greatly resisted writing tone.

Part of the resistance to writing tones is due to the perception that writing should look like the system on which it is modeled. In Africa this is mostly either English or French, which do not write tones. But tones have several other issues that appear to contribute to the reluctance to write them. Specifically, tones tend to be low in *segmentability*, *stability*, and *proprioception*. Segmentability refers to the inherent difficulty of explicitly factoring out tone as a phonetic feature that is different from the more concrete vowels and consonants over which it is superimposed. Even if a speaker understands that /a˥/ is a combination of the vowel /a/ and a high tone, the requirement to write separate symbols for the vowel and the tone impresses her as inconvenient clutter. Stability refers to the fact that tones tend to be pronounced differently in different phonetic and grammatical contexts. For example, in many African languages a high tone is pronounced with a very high pitch at the beginning of a sentence but with a much lower pitch toward the end of a sentence. Such things can make it difficult to associate a word with a specific tonal pattern. Proprioception refers to the fact that people can scarcely feel any differences in what their speech organs are doing when they make different tones, depriving them of a useful cue in reading and writing. It is not clear which of these issues is the most important, but underrepresentation of tone must be due to factors beyond adherence to models of writing that use the Latin alphabet. Many independently developed writing systems for tonal languages such as Cherokee, Vai, Japanese, and Chinese do not represent tones at all.

Similar factors appear to contribute to a tendency to underrepresent several other features in writing systems. Lexical stress, such as the distinction between /ɪnˈsɛns/ 'enrage' and /ˈɪnsɛns/ 'substance burned for its fragrance' in English, is not explicitly represented in the spelling: both of these examples are ‹incense›. Contrasting phoneme lengths often go underrepresented as well. For example, in Fijian, words may be distinguished by whether a vowel is pronounced longer or shorter: /mama/ means 'a ring', but /mamaː/,

with the second vowel lengthened, means 'chew it', and /maːmaː/, with both vowels lengthened, means 'lightweight'. Although the words differ phonemically, they all are written *mama*. Writers may find it difficult to identify length as a feature distinct from the vowel itself, especially because length is not stable: All vowels sound long when one tries to read or write a word by sounding it out slowly. Graphical considerations may also contribute to the underrepresentation of stress, length, and tone that we see across writing systems. In part for the historical reasons mentioned earlier, these features of speech are usually represented by small marks that clutter and complicate the written text (Kutsch Lojenga, 2011). Writers may omit the marks for this reason. That makes writing less visually complex than it would otherwise be, but it can also lead to ambiguity that readers must deal with.

Writing Changes but Lags Behind Language Change

We have already mentioned that writing may change over time, as when diacritics are added to indicate certain previously unexpressed aspects of pronunciation. In this section, we consider how writing changes and why such changes usually occur more slowly in writing than in spoken language.

Table 2.1 illustrates some changes to the shapes of letters that took place during the evolution of Arabic. The ultimate ancestors of these letters were the early Phoenician letters shown in the second column. As the script developed, it become more and more cursive—that is, simplified and adapted for quick writing—until the Aramaic letters begin to resemble today's Hebrew script, with many of its letters clearly pointing to the left, leading to the next character. By the time the earliest Arabic texts appeared, it had become customary to join many of the letters together, as in the traditional cursive form of such scripts as the Latin of English and French.

Table 2.1 Changes in Letter Shapes From Phoenician to Hebrew and Arabic.

Sound	Phoenician	Aramaic	Hebrew	Arabic	
				Early	Dotted
ʔ	𐤀	𐡀	א	ا	ا
b	𐤁	𐡁	ב	ـٮـ	ـبـ
t	𐤕	𐡕	ת	ـٮـ	ـتـ
p, f	𐤐	𐡐	פ	ـڡـ	ـفـ
q	𐤒	𐡒	ק	ـڡـ	ـقـ

Further, the letter forms had become so simplified that several of them had become indistinguishable. For example, as Table 2.1 shows, the forms for /b/ and /t/ in early Arabic were identical. This degree of similarity proved to be too high, so a system of dots to distinguish letters that stood for the same sound was standardized in the seventh century A.D.

As the Arabic example shows, changes to the shapes of symbols are often driven by a need for economy when writing by hand. The symbols become simpler, possible to be produced in fewer strokes and with fewer lifts of the writing implement. Symbols that have a pictorial quality (the Phoenician symbol in the first row in Table 2.1 was originally meant to depict the head of an ox) tend to lose that quality over time. However, a writing system must maintain a sufficient degree of contrast between symbols so that the human visual system, working quickly, can tell them apart. When contrast breaks down, methods may develop to allow for differentiation, as with the system of dots that was introduced in Arabic.

Although writing changes over time, it does not usually change quickly. This conservatism reflects, in part, the characteristics of its users. People become attached to the status quo, valuing the familiar in the case of writing as in the case of other things. For writing, conservatism is fostered by the fact that revered documents from the past, such as a sacred scripture, could become harder to read if the writing system changed. Spoken language is less conservative than written language, probably because speech fades quickly without special technology. One would have to compare the present against one's memory of the past to notice a change in how people speak. In contrast, the permanence of writing means that one need not rely on memory to notice changes in how people write. These considerations mean that the pace of change is generally faster for spoken language than for written language.

One manifestation of the conservatism of writing is a tendency to retain the spellings of words after they have been borrowed from another language. For example, in English, French, and German, the sound [f] is spelled *ph* in words and morphemes that were borrowed from Greek, such as *philosophy*, but as *f* in non-Greek morphemes. This etymological rule is very regular and would doubtless come naturally to children who learned to read and write in Greek before learning their native English, French, or German. But most children do not learn Greek first, of course, and so these etymological spellings come across to them as difficult.

Another manifestation of writing's conservative nature is that the grammar of the language that is encoded in print is somewhat different from the grammar that people normally use when they speak. It is more old-fashioned. Finnish is justly famous for having an almost perfectly consistent phonemic writing system, but it also has a wide gap between colloquial and formal styles of speech. The colloquial register is what people acquire as children and use for practically all spoken communication; the latter is what appears in practically all printed material (Karlsson, 2004). For example, schoolchildren in Finland are taught *Olemme talossa* for 'we are in the house'. That can be read off as /olemːe talosːa/ using simple relationships between letters and phonemes. In daily life, however, people would hear and say /me olːaːn talos/. The gap between written and spoken language is even larger for speakers of some other languages. For example, the standard literary form of Arabic is very close to the classical Arabic language spoken over 1,000 years ago and very different from the colloquial forms that are spoken today. Such is the conservatism of writing that historical and cultural factors are the main determinants of why a particular language has the writing system that it does.

Summary

The writing systems that are used by people with normal perceptual abilities employ visual symbols to represent language. In its outer form, writing appears as strings of characters that are arranged along lines. In its inner structure it concentrates on representing the words of a language. Although writing represents language, it does not represent all aspects of language. Some features tend to be left out. Too much under-representation can make reading difficult, but readers can get by with a surprising amount of it because they normally know the language they are reading. As we discuss later in the chapter, readers can often add missing information by using context: what they pick up elsewhere on the page or what they know about the language or about the world.

Differences Among Writing Systems

Writing systems also differ substantially among themselves. We begin by considering differences in their inner structure and conclude by looking at some important types of differences in their outer form.

Writing Represents Words in Different Ways

All general-purpose writing systems use symbols to represent language. It is common to categorize writing systems by how their most fundamental symbols do so (e.g., Sampson, 1985). Writing systems that use symbols that stand for morphemes are called *logographies*. Phonographic systems in which the basic elements represent syllables are *syllabaries*, and phonographic systems in which the basic elements represent phonemes are *alphabets*. However, placing writing systems into mutually exclusive categories such as these invites unproductive quibbling about which level is the most fundamental in a particular writing system and tends to neglect the fact that readers are likely to pick up on multiple levels of representation. In practice, most writing systems are mixed. Some contain logograms and phonograms side by side. Some have symbols for one level of representation that are composed of symbols of another level of representation. In what follows, we consider the choices that are made by several writing systems.

Chinese writing, like all modern writing systems, consists of a linear sequence of characters. Chinese gives priority to the lexical level of Martinet (1960) that was described earlier. Specifically, each Chinese character corresponds to a morpheme, which, by definition, has both a meaning and a pronunciation. In Chinese, those pronunciations are monosyllabic. The Chinese character 生 represents the specific morpheme that means 'born', for example. Because that morpheme has a pronunciation [ʂəŋ˥], then 生 itself represents the syllable [ʂəŋ˥]. Lexical constancy means that the morpheme [ʂəŋ˥] 'born' is always represented as 生, and a very strong drive toward lexical distinctiveness in Chinese means this morpheme is represented differently from the morpheme that is pronounced [ʂəŋ˥] and means 'raise' 升 and from the morpheme that is pronounced [ʂəŋ˥] and means 'nephew' 甥. Lexical demarcation is handled by packing together into one character all the visual elements that contribute to the representation of the morpheme. Characters in Chinese can be identified by the fact that they all occupy the same amount of square space in linear sequence. Thus, in the character 甥 [ʂəŋ˥] 'nephew', it is obvious that both the elements 生 [ʂəŋ˥] 'born' and 男 [nan˧˥] 'male', being squeezed into one square space, contribute to the representation of a single morpheme. Most modern Chinese words have more than one morpheme, and therefore they must be expressed by a sequence of characters. For example, the compound word [nan˧˥ʂəŋ˥] 'male student' is written 男生. Indeed, the structure of Chinese is such that all polysyllabic words must be spelled with multiple characters, even if their morphemic structure is not clear to present-day speakers. Chinese writing takes no special notice of words; it does not even put spaces between them. The fundamental level of representation in Chinese writing, at which all lexical demarcation is done, is at the level of the morpheme.

An interesting feature of Chinese writing is that a great number of characters, which by themselves represent a specific morpheme, can also occur as components of other characters, which stand for a different morpheme. In that usage they give hints as to what morpheme the character represents, typically by reference to other morphemes with related meaning or pronunciation. For example, the 男 in 甥 [ʂəŋ˥] 'nephew', representing a morpheme that means 'male', gives a hint as to the meaning of 甥. The 生, being pronounced [ʂəŋ˥], gives a very good hint as to the pronunciation of 甥. In this character, 生 is used phonetically. Some 81% of Chinese characters contain component symbols that give a hint to how the character is pronounced, although in only about half of those characters does the phonetic component match the pronunciation exactly, even if we ignore tone mismatches (Hsiao & Shillcock, 2006).

Chinese represents a situation where there are many symbols that can stand for syllables but none that stand for smaller units of sound. Symbols standing for syllables also exist in several other writing systems. In the Japanese sentence shown in Figure 2.4 (repeated here from Figure 2.1), all symbols except for the first are pronounced as one syllable. Many of the symbols were borrowed from Chinese. Others are members of Japanese syllabaries that can be used to represent syllables in a variety of different words. The katakana syllabary is used, among other purposes, to represent words borrowed from languages that do not use Chinese characters: in this case, [ɡarasu] from English *glass*. The hiragana syllabary is used more generally than katakana to phonetically spell out morphemes for which there is no ready Chinese character. These morphemes include function words and inflectional endings. For example, ます [masu] is an ending expressing politeness. The word [taberaremasu] 'can eat' illustrates an intermediate use of hiragana. The character 食 was borrowed from Chinese to represent the root morpheme [tabe-] 'eat.' Logically, one might expect that character to be followed here by hiragana that spell out only the inflectional endings, [raremasu]. Instead, it is followed by hiragana that spell out [beraremasu], which includes a spelling of the last part of the root. This apparent redundancy is useful because 食 can also represent other root morphemes in Japanese, such as [kuw-] 'feed on'. Adding a phonetic spelling for the last part of the root helps the reader determine which morpheme the writer intended to represent by the symbol 食. This strategy is used throughout Japanese, because the great majority of characters borrowed from Chinese can represent at least two different morphemes with different pronunciations.

In Japanese, the symbols that represent syllables do so atomically; they are not composed of smaller elements that represent the individual phonemes. For example, the symbol ら for /ra/ has nothing in common with the symbol れ for /re/, even though those syllables begin with the same consonant, nor with the symbol ま for /ma/, even though those syllables have the same vowel. This is true in other syllabaries as well. In Cherokee, Ꮃ represents [la], but no part of that symbol is reprised for other syllables that contain [l], such as Ꮄ [le], or for other syllables that contain [a], such as Ꮉ [ma].

In some other scripts where a character represents a syllable, the character is decomposable. In Korean, each syllable consists of elements that symbolize individual phonemes. For example, the symbol for [m] is ㅁ, which can be found at the top left of the character 명 [mjʌŋ] and at the bottom of the character 함 [ham]. When these two characters are sequenced along a line, one gets the word 명함 [mjʌŋham] 'business card'. Thus if we focus on characters, Korean can be said to represent syllables, like a syllabary. If we focus on the smallest component symbols, however, Korean represents phonemes. In this latter sense, it is an alphabet.

Many of the scripts of South Asia, including Hindi, are similar to Korean in that individual characters often represent syllables but may be composed of smaller symbols that represent phonemes.

Kanji	hira	kata			hira	Kanji	hira				
私	は	ガ	ラ	ス	を	食	べ	ら	れ	ま	す
watasi	**wa**	**ga**	**ra**	**su**	**o**	**ta(be-)**	**be**	**ra**	**re**	**ma**	**su**

Fig. 2.4 Characters in a Japanese sentence (from Figure 2.1, example 2) aligned with their pronunciation and labeled with character type. *hira*: hiragana; *kata*: katakana.

Such systems are sometimes called alphasyllabaries, syllabic alphabets, or *abugidas*. The Hindi example in Figure 2.5 reprises the 'I can eat glass' sentence from Figure 2.1, but inserts breaks between characters instead of between words. Even in this short extract, we can see that some of the characters are compositional. The characters for [kɑɲ] and [k] share the shape क, and syllables containing the vowel /ɑ/ have a vertical stroke at the right side of the character.

We talked earlier about the fact that underrepresentation is common across writing systems. It is particularly common in writing systems that represent syllables in a noncompositional way, like Japanese and Cherokee. Often, such systems are far from having different symbols for all of the syllables found in the language. In Cherokee, for example, distinctions of tone and vowel length are unrepresented. The phonemic distinction between plain and aspirated consonants is not noted for most consonants, and the glottal sounds [h] and [ʔ] are not noted at the end of syllables (Montgomery-Anderson, 2008).

Apart from neglecting tone, modern Japanese has eliminated such ambiguities in its syllabaries. But it has syllabic symbols—syllabograms—for only 44 syllables: one for each of five short vowels (e.g., あ for [a]), and 39 for combinations of a single short consonant and a short vowel (e.g., か for [ka]). Instead of using extra syllabograms for the hundreds of remaining syllables in the language, additional symbols are combined with the basic syllabograms to indicate such features as vowel length (かあ for [kaː]), consonant length (つか for [kːa]), syllable-final nasals (かん for [kan]), and consonant clusters in syllable onsets (きゃ for [kja]). From our perspective, the principle behind the design of the Japanese syllabaries is to minimize the number of syllabograms by providing them only for the simplest vowel and consonant–vowel syllables and using those syllabograms as the basis of more complex syllables. Many other writing systems agree with Japanese in handling things such as phoneme length by adding an element to that for a short phoneme. Another common perspective is to say that Japanese syllabaries are *moraic* by design. A *mora* is the length of a rime that consists of nothing but a short vowel. Rimes that have a long vowel, or include a consonant, have two moras. Moras are important to a phonological analysis of Japanese; for example, haiku have lines of five, seven, and five moras. Because of the importance of moras in Japanese, it is often considered significant that most one-mora syllables, such as [ka], are spelled with one symbol (か), whereas most two-mora syllables are spelled with two symbols, such as かあ for [kaː], かん for [kan], and かつ for [ka] followed by a stop. It is sometimes argued that the symbols spell moras, not syllables. However, syllables with onset consonant clusters break this rule, in that they have more symbols than moras, as in きゃ for the one-mora syllable [kja]. In our view, it is more useful to connect symbol counts with syllable complexity than with mora counts, because the former accounts for the use of multiple symbols in onsets as well as in rimes.

मैं	काँ	च	खा	स	क	ता	हूँ
mãĩ	kɑɲ	t͡ʃ	k^{h}ɑ	sə	k	tɑ	hũ

Fig. 2.5 Characters in a Hindi sentence (from Figure 2.1, example 5) aligned with their pronunciation.

In many writing systems, the individual characters that are arranged along the lines of print represent phonemes. To use a more familiar terminology, the characters are letters. Alphabets based on the Arabic, Cyrillic, Greek, Hebrew, and Latin scripts are familiar examples of systems that rely almost entirely on letters. The relationships between letters and phonemes are often not one-to-one, however. There are several historical sources of such mismatch. Sometimes when existing writing systems are adapted to new languages that have phonemes not provided for in the existing system, the lack is mitigated by assigning the sound to a sequence of two letters, a *digraph*. For example, the English sound [θ] did not exist in the Latin script, so the digraph ‹th› was used. Another reason for letter–phoneme mismatches is the conservatism of writing that was mentioned earlier. When the pronunciations of words change, spellings often do not change to keep up with them. In much of the English-speaking world the pronunciation of [ɹ] in a syllable coda has changed or been lost, resulting in many homophones such as *lord* and *laud*, but an ‹r› remains in the spelling. This conservatism in the spelling means there are more ways to spell several sounds such as [ɔ], which makes spelling harder for people who have this sound change. At the same time, the conservative spelling means that people who have already learned to spell ‹lord› do not have to change their ways, and existing books with such spellings still look fresh. Perhaps most importantly, reluctance to drop the ‹r› from the spellings avoids conflict with dialects in which the [ɹ] still exists in words like *lord*. Conservatism in writing means that spelling has remained more consistent throughout the world than pronunciation has.

Because sounds sometimes change in particular positions of words or when near other specific sounds, the pronunciation of a letter may depend on the position in which it occurs in a word or on the surrounding letters. For example, the vowel of English *fall* was originally the same as that of *fat*, but it moved further back in the mouth under the influence of the following /l/, which is pronounced here with the back of the tongue raised. The spellings did not change. As a result, *fall* now has the same vowel phoneme as *fault* in most dialects of English, but a different vowel spelling. A reader who considered the following letter when deciding on the pronunciation of the ‹a› would have a better chance of being correct than a reader who used a context-free rule that linked the letter ‹a› to the phoneme /æ/. This type of situation is very common in English (Kessler & Treiman, 2001). Purely graphotactic patterns can also make links between letters and phonemes more complex. For example, English spelling generally avoids word-final ‹c› or ‹k› after a single vowel letter, especially in one-syllable words. The spelling ‹ck› is typically used instead, as in *back*. This means that the reader must learn an additional digraph for no phonetic reason.

To some extent, alphabetic writing systems represent other levels of language besides the phoneme. Almost all such writing systems clearly demarcate individual words and spell words in a consistent manner. This lexical constancy helps readers to recognize lexical elements as well as individual phonemes. However, syllables tend not to be separately demarcated in alphabets, nor, in general, are sublexical morphemes such as prefixes and suffixes. But some systems do make an attempt at morphemic constancy, such that a morpheme has the same spelling in different words even if the pronunciation varies somewhat. For example, the first part of the English *cleanliness* [ˈklɛnlinɪs] is spelled the same way as *clean* [ˈklin], even though ‹ea› is a rare spelling for [ɛ]. As another example, German spells *Tag* [ˈtɑːk] 'day' with the same ‹g› that appears more phonetically in the plural *Tage* [ˈtɑːɡə]. Although many cases of morphemic constancy may be due to the conservative retention of older spellings, they probably help to make morphemes easily identifiable.

Some alphabetic writing systems, often called *shallow*, have primarily one-to-one links between letters and phonemes. Finnish fits this description, at least if readers use the pronunciations of its formal register, as mentioned earlier. The term *deep* is often used to refer to writing systems such as English that have "orthographic inconsistencies and complexities, including multi-letter graphemes, context-dependent rules, irregularities, and morphological effects" (Seymour, Aro, & Erskine, 2003, p. 146). However, these different sorts of patterns probably have different effects on readers. Readers may find it easier to use a rare pronunciation of a letter when that allows for morphemic constancy, for example, than when it does not. Indeed, when Chomsky and Halle (1968) introduced the term *deep* in reference to writing, they used it for cases such as *cleanliness* in which spellings reflect morphological considerations. These are cases in which people must go deeper than the surface phonology in order to make sense of the spelling. Chomsky and Halle would not consider multiletter graphemes such as ‹ck› or letter-to-sound rules that depend on context to fall into the same category as the deep spellings to which they refer.

The individual characters that are arrayed along a line of print may represent morphemes, syllables, or phonemes, but there is no writing system in which they represent distinctive features such as velar place of articulation or aspiration. In a few writing systems, including Korean, the shapes of some letters bear a relationship to the features of the sounds that they represent. For example, ㅌ for the aspirated stop [t^h] and ㅋ for the aspirated stop [k^h] add a horizontal line to the letter that represents the corresponding unaspirated stop, ㄷ for [t] and ㄱ for [k]. However, these patterns are not always geometrically or phonetically consistent.

Differences in Outer Form

In discussing differences among writing systems, we have focused so far on differences in their inner structure: how they represent language. In this section we briefly consider differences among writing systems in their outer form.

As we discussed earlier, all scripts lay out their characters sequentially. They differ, though, in the *direction* in which the characters are laid out. Nowadays the great majority of languages are written in horizontal rows from left to right. The popularity of that direction invites explanation. One might speculate that left-to-right writing has won out because most people are right-handed, and ink is less likely to be smudged if the right hand proceeds to the right after writing a word. Vertical writing may be less common because the effective field of vision extends further horizontally than vertically, and because the eye may be more used to tracking horizontal movement than vertical movement. Upward writing may be least common of all because objects fall more often than they rise. The prevalence of left-to-right writing systems may be misleading, however, because so many writing systems descend from just a couple of ancestors—Classical Greek and Latin (among them English and most other European systems) and Brahmi (most scripts of south Asia)—or were inspired by one of those descendant systems. For example, Chinese and Korean were traditionally written vertically but are predominantly written nowadays from left to right, in order to better integrate with European books and other writing technologies.

To a small extent, the outer forms of scripts vary in how *pictorial* the symbols are. Certain ancient scripts, such as those of Egyptian and Mayan inscriptions, contained a high proportion of symbols that pictured objects. In modern scripts, pictoriality is harder to come by. Korean consonant letters are supposed to be based on images of the vocal tract—ㅁ represents the closed lips, for example—but it is doubtful that people would recognize such pictures without being taught their meaning. Chinese characters are often characterized as pictorial, and indeed a few hundred of them started out as pictures three thousand years ago. Nowadays, however, people who have not been taught what those characters mean cannot usually guess what they represent (Xiao & Treiman, 2012). In general, even scripts that start out with largely pictorial symbols make them highly simplified from the very beginning, and the symbols typically become unidentifiable as they are changed to make writing faster and more compact.

A third way in which scripts vary in their outer form is in the *complexity* of their characters. Among the factors that most commonly contribute to character complexity are the number of contrasts the characters must represent and whether they are composed of multiple elements. In alphabetic writing systems that treat letters as characters, characters are fairly simple—a little less than three strokes on average (Changizi & Shimojo, 2005). Characters are more complex in writing systems such as Hindi and Korean, in which symbols that stand for phonemes combine into characters that stand for syllables. Chinese characters are quite complex. This is in part because Chinese needs a way to make several thousand visually distinct characters and in part because it achieves that by combining multiple components in each character.

As mentioned earlier, each script has a certain stylistic consistency: A character in a particular script tends to look more like other characters in that script than like other signs or images. The specific stylistic features that are shared differ from one script to another. For example, circles are absent from modern Chinese character forms but are present in some letters of the Latin alphabet. Such patterns let people make generalizations and simplifying assumptions when reading and writing. For example, a Chinese writer who imperfectly remembers that a character contains a certain closed shape does not need to consider whether that shape was a circle. A learner of the Hebrew script will apprehend that letters tend to open to the left but not to the right, so if the correct left–right orientation of כ is not immediately recalled per se, it can quickly be deduced. Of course, learning a rule may be injurious when encountering exceptions. When a letter of the Latin alphabet contains a vertical line to which additional segments are attached, the appendage is usually on the right side of the vertical line, as with ‹b› and ‹p›. Children who learn that regularity may have trouble writing the few exceptions, such as ‹d› and ‹q› (Treiman & Kessler, 2011).

Summary

Although all writing systems represent language, they do so in different ways. Systems differ in the emphasis that they place on the lexical level and the phonological level, although many systems include elements of both. Writing systems also differ in their outer form, including the visual properties of their characters and the kinds of lines along which the characters are arranged.

Implications for Reading and for Learning to Read

Having discussed the nature of writing systems, including their similarities and their differences, we focus in this section on the implications for reading. We briefly discuss some general implications for how the reading process must take place, both in skilled readers and in those who are learning to read. We do not have the space to review the empirical evidence for the predictions, but we point to a few studies that have addressed some of the issues.

Reading Involves the Recovery of Linguistic Form

Because writing represents language, successful reading involves recovering the linguistic form that is represented in print. Reading begins with the eyes, because writing is visual, but "successful skilled reading enables the language system to take over from the visual system with astonishing speed" (Perfetti, 2012, p. 299). Indeed, much research has documented the involvement of phonology in silent reading across a variety of writing systems (see Pollatsek, this volume). And it is not just phonology that is recovered from print, but other aspects of language as well. Models of the cognitive processes involved in skilled silent reading must specify how linguistic processing shapes and then takes over from visual processing, and our consideration of the nature of writing systems suggests that models that postulate a close relationship between visual and linguistic processing have more psychological plausibility than models that do not.

Given the nature of writing systems, children cannot get very far in learning to read until they have learned that elements of writing stand for specific elements of language. Thus, ‹cat› stands for *cat* and not *kitty*. Children must learn that reading a word is different from naming an object or a picture, where it could be appropriate to use either *cat* or *kitty* for a young feline. Learning to read is easiest, moreover, if children already know the language that is represented. By six years of age or so, when children in many societies begin learning to read, language development is quite advanced. Children know many words and many aspects of grammar. However, the language that is represented in writing may include certain lexical items and grammatical structures that are not common in colloquial speech and that may be unfamiliar to children. As mentioned earlier, this is especially true in certain cultures, including those that use Arabic. In any language, though, people learn some aspects of language primarily from reading.

Readers Need Not Always Take in All Aspects of Writing's Visual Form

As we have discussed, there is some redundancy in writing by virtue of the redundancy in language itself and by virtue of the redundancy in symbols of writing. This means that readers can sometimes get by with only partial uptake of visual information. Thus, a reader of English could recognize ‹Alligator› even if he missed the crossbar on the ‹A› and even if he did not take in the last few letters. This is because there is no other letter in the Latin alphabet that is the same as ‹A› except for the crossbar and no other word that differs from ‹Alligator› in just the last few letters. As another example, a reader of English could recognize ‹campus› even if she did not resolve the relative order of ‹m› and ‹p›. Readers learn which aspects of writing's visual form are critical to take in and which are less critical, and this can depend on which language they are reading (see Frost, this volume). In any written language, though, it is rarely necessary to process each and every visual element. In order to write a word correctly, on the other hand, complete information about the identity and the order of the elements is required. This is a major reason why, across writing systems, spelling is more difficult than reading (see Treiman & Kessler, 2014, for a discussion of spelling).

Reading Often Requires Information That Is Not on the Page

All writing systems represent language, but no writing system represents all aspects of the language. For example, we have seen that intonation and stress are often not represented. Underrepresentation means that in order to understand a text, readers sometimes must supplement what is on the page with other things that they know about the language or the world. For example, readers of English might decode the spelling ‹give› as [ɡaɪv], but because they do not recognize that as a word of their language, they will quickly rule it out. They can use their knowledge about the world to suggest that ‹sewer› in *The dirty water drained into the sewer* stands for [ˈsuɚ] 'a conduit for carrying off waste' as opposed to [ˈsoɚ] 'a person who sews'. Sometimes, the information that is needed to disambiguate a word comes only after the word has been read. Consider the sentence *Since the old sewer was awful, the tailor's shop got a bad reputation* (Folk & Morris, 1995). When the reader gets to the part about the tailor's shop, it becomes clear that *sewer* is a person who sews. But before this point it is possible that *sewer* could be a conduit for carrying off waste. Readers have difficulty with sentences such as this, sometimes moving their eyes back to *sewer* when they reach the point where the meaning is disambiguated. Readers' knowledge about things that are not on the page can often reduce or even eliminate the effects of any remaining ambiguities.

Learning to Read Is Challenging

Writing is not a part of the human genetic endowment, as spoken language is; it is an "optional accessory that must be painstakingly bolted on" (Pinker, 1997, p. ix). The ease with which educated modern people read can be misleading: That ease is achieved only through years of practice. Even when a child already knows the language that is represented, learning to read accurately and fluently is hard. Some of the challenges, such as the need to learn that a written word represents its meaning differently than a picture does, are the same for learners of different writing systems. Others are different, such as learning the direction the characters are written in or what specific shapes correspond to what specific sounds or morphemes.

Learning to read and write typically requires a degree of explicit instruction that is not required in order to learn to speak and listen. That instruction should be based on a solid understanding of how the writing system works, but that is not always true. For example, phonics instruction for English does not typically give adequate consideration to the fact that the pronunciation of a letter or digraph can be influenced by the context in which it occurs. It does not typically consider how morphology and graphotactics can help in choosing among alternative pronunciations. Children could benefit from learning about such patterns as that ‹oo› is typically pronounced as [ʊ] when it occurs before ‹k›, as in *book* and *look*, but as [u] in other contexts. They could benefit from learning that the final ‹e› of words like *give* and *have* is motivated by graphotactic considerations (English words do not normally end with ‹v›), and the final ‹e› does not influence the pronunciation of the vowel as it does for *guide* and *hate*.

Knowledge of Written Language Is Used for More Than Reading and Writing

Writing developed as a tool to allow people to freeze language, which fades quickly without special technology. By placing language in a relatively permanent form, writing permits communication among people who are distant in time and space. Once learned, however, writing becomes a tool for other purposes. For example, people use writing to help remember things. Because writing represents language, people use their knowledge about the written forms of words in thinking about the words' spoken forms. Indeed, a number of studies support the idea that literate people's ideas about speech can be influenced, sometimes accurately, sometimes inaccurately, by their knowledge of writing (see Kolinsky, this volume). For example, people often judge that the word *lagoon* [ləˈgun] contains the syllable [læg] because it contains the letters ‹lag› (Taft & Hambly, 1985). Sometimes, awareness of the spelling even leads people to change their pronunciation to more closely reflect the written form of a word. An example of a *spelling pronunciation* that has become widespread in English is *often*, which is now often pronounced with a [t] that had been silent for 500 years.

Conclusions

Theories about how people read and about how they learn to read should be based on a good understanding of writing systems and how they work. If our knowledge about writing is too narrow, perhaps limited to our own writing system and others like it, our theories may be too narrow. The teaching of reading may also suffer. In this chapter, we have presented some basic information about writing systems. We have seen that although writing's outer form can vary a good deal from one writing system to another, there are a surprising number of similarities across writing systems in both outer form and inner structure. The properties of writing mean that skilled reading involves the recovery of linguistic forms. Readers do not always have to take in all aspects of the print in order to recover the linguistic form; however, they often need to use information outside the print itself.

References

AIGA. (1989). *Tools and resources*. http://www.aiga.org/symbol-signs

Changizi, M. A., & Shimojo, S. (2005). Character complexity and redundancy in writing systems over human history. *Proceedings of the Royal Society B, 272,* 267–275.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York, NY: Harper and Row.

Evans, A. J. (1894). Primitive pictographs and a prae-Phoenician script from Crete and the Peloponnese. *Journal of Hellenic Studies, 14,* 270–372.

Folk, J. R., & Morris, R. K. (1995). Multiple lexical codes in reading: Evidence from eye movements, naming time, and oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1412–1429.

Gelb, I. J. (1952). *A study of writing: The foundations of grammatology*. Chicago, IL: University of Chicago Press.

Hsiao, J. H., & Shillcock, R. (2006). Analysis of a Chinese phonetic compound database: Implications for orthographic processing. *Journal of Psycholinguistic Research, 35,* 405–426.

Karlsson, F. (2004). *Finnische Grammatik* [Finnish grammar]. 4th ed. Transl. of *Suomen peruskielioppi*. Hamburg, Germany: Buske.

Kessler, B., & Treiman, R. (2001). Relations between sounds and letters in English monosyllables. *Journal of Memory and Language, 44,* 592–617.

Kolinsky, R., Verhaeghe, A., Fernandes, T., Mengarda, E. J., Grimm-Cabral, L., & Morais, J. (2011). Enantiomorphy through the looking glass: Literacy effects on mirror-image discrimination. *Journal of Experimental Psychology: General, 140,* 210–238.

Kutsch Lojenga, C. (2011, January). *Orthography and tone: Tone system typology and its implications for orthography development*. Paper presented at the meeting of the Linguistic Society of America, Pittsburgh, PA.

Martinet, A. (1960). *Éléments de linguistique générale*. Paris, France: Colin.

Mollick, E. (1996). The "I Can Eat Glass" project. http://web.archive.org/web/20040201212958/http://hcs.harvard.edu/~igp/glass.html

Montgomery-Anderson, B. (2008). *A reference grammar of Oklahoma Cherokee* (Doctoral dissertation, University of Kansas). Retrieved from http://kuscholarworks.ku.edu

Perfetti, C. (2012). Thru but not wisht: Language, writing, and universal reading theory. *Behavioral and Brain Sciences, 35,* 299–300.

Pinker, S. (1997). Foreword. In D. McGuinness, *Why our children can't read and what we can do about it* (pp. ix–x). New York, NY: Free Press.

Pulgram, E. (1976). Typologies of writing systems. In E. Pulgram & W. Haas (Eds.), *Writing without letters* (pp. 1–28). Manchester, England: Manchester University Press.

Sampson, G. (1985). *Writing systems: A linguistic introduction.* Stanford, CA: Stanford University Press.

Seymour, P. H. K., Aro, M., & Erskine J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94,* 143–174.

Taft, M., & Hambly, G. (1985). The influence of orthography on phonological representations in the lexicon. *Journal of Memory and Language, 24,* 320–335.

Treiman, R., & Kessler, B. (2011). Similarities among the shapes of writing and their effects on learning. *Written Language and Literacy, 14,* 39–57.

Treiman, R., & Kessler, B. (2014). *How children learn to write words*. New York, NY: Oxford University Press.

Truss, L. (2004). *Eats, shoots & leaves: The zero tolerance approach to punctuation*. New York, NY: Penguin.

Xiao, W., & Treiman, R. (2012). Iconicity of simple Chinese characters. *Behavior Research Methods, 44,* 954–960.

CHAPTER

3 Visual Word Recognition

Melvin J. Yap *and* David A. Balota

Abstract

Visual word recognition is an integral aspect of reading. Although readers are able to recognize visually presented words with apparent ease, the processes that map orthography onto phonology and semantics are far from straightforward. The present chapter discusses the cognitive processes that skilled readers use in order to recognize and pronounce individual words. After a historical overview of the broad theoretical developments in this rich field, the chapter provides a description of methods and a selective review of the empirical literature, with an emphasis on how the recognition of an isolated word is modulated by its lexical- and semantic-level properties and by its context. The final section of the chapter briefly considers some recent approaches and analytic tools in visual word recognition research, including megastudy, analysis of response time distributions, and the important role of individual differences.

Key Words: visual word recognition, lexical decision, speeded pronunciation, masked priming, semantic priming, orthographic priming, phonological priming, megastudies, individual differences, response time distributional analysis

Skilled reading is a remarkably complex and multifaceted behavior, which relies on the recognition of individual words. The squiggly marks on the page need to somehow map onto a word representation so that the meaning of the word can be accessed. At first blush, this appears to be a relatively straightforward process of pattern recognition. However, words code and convey multiple domains of information, including orthography, phonology, morphology, and ultimately meaning. Indeed, because of the multidimensional nature of word recognition, this literature has made seminal contributions to (1) the distinctions between automatic and attentional mechanisms (e.g., Neely, 1977), (2) the development of computational models (e.g., McClelland & Rumelhart, 1981), and (3) cognitive neuroscience (e.g., Petersen, Fox, Posner, Mintun, & Raichle, 1989). Given the extensive influence of word recognition research on cognitive science, attempting to provide a concise overview of this area is a daunting task. We have chosen to first provide a brief historical overview of the area, with an emphasis on the wide-ranging theoretical contributions. We then turn to some basic findings in the literature and conclude with more recent developments in studying word recognition. Our goal is to expose the reader to the major issues, as opposed to providing detailed expositions of each of the research topics.

Historical and Theoretical Overview

Although a number of writing systems exist, reading research has been dominated by the study of alphabetic writing systems, where the unit of language symbolized by writing is the phoneme (Treiman & Kessler, 2007). In alphabetic writing systems, the building blocks of words are letters, and so the recognition of letters was central to early models of visual word processing. If printed words are recognized via their constituent letters, then it is natural to wonder whether letters are also recognized via their constituent features (see Grainger, Rey, & Dufau, 2008, for a review). An important approach in this area is the feature analytic approach. According to this view, there is a set of visual features (e.g., vertical lines, horizontal lines, diagonal lines, curved closed forms, closed open forms, intersections) that are critical for discriminating among the letters. So, the letter ‹H› would be defined by the convergence of two vertical lines and one horizontal line. Indeed, component features such as these laid the foundation for the first computational model of letter perception (pandemonium model; Selfridge & Neisser, 1960). About the same time, Hubel and Wiesel (1962) were able to identify receptive fields of cortical neurons in alert cats; these receptive fields appeared to be sensitive to vertical lines, horizontal lines, oblique lines, and intersections. Although it is likely that such features play an important initial role in letter perception, many questions remain. These include (1) how the features are bound together to form a letter (see Treisman, 1999, for a review of the binding problem); (2) how the system flexibly codes different sets of features that are necessary for recognizing letters across fonts, visual angles, and levels of degradation; and (3) how the system adjusts to handwritten text wherein the features appear to be very different from standard text (see Plamondon & Srihari, 2000, for a detailed review).

Moving on to the letter level, letters vary in the extent of feature overlap, and, as expected, this influences the ease of searching for a letter in a background of letters (e.g., it is more difficult to locate ‹Z› when it is embedded within the letters ‹F›, ‹N›, ‹K›, and ‹X›, than when it is embedded within ‹O›, ‹J›, ‹U›, ‹D›; see Neisser, 1967). Appelman and Mayzner (1981), in a comprehensive review of isolated letter recognition, considered studies that measured (1) participants' accuracy for identifying single letters under varying levels of degradation or (2) their response times for letter naming, letter matching, and letter classification (i.e., letter vs. nonletter forms). The results, based on over 800,000 observations from 58 studies, revealed that the frequency of a letter in the language (e.g., ‹T› is approximately three times more frequent than ‹C›) had no effect on accuracy-based studies where participants simply report letters. Interestingly, however, there was a clear effect of frequency on response latencies. Appelman and Mayzner (1981) suggested that the consistent absence of letter frequency effects in accuracy was incompatible with the idea that early letter encoding is modulated by letter frequency. We primarily note this pattern because it is surprising that the simple effect of frequency of exposure would produce varying influences across tasks, and hence it is important to remind the reader that there are always important between-task differences when considering the influence of a variable on performance.

Recognizing Letters Within Words

Letters are rarely presented in isolation, but are typically embedded in words. Interestingly, Cattell (1886) argued that letters (e.g., ‹n›) were more easily reported when presented in the context of letters that form words (*born*) than in the context of letters that form nonwords (*gorn*). There are many interpretations of this simple effect. For example, partial information from words (*bor_*) might be more useful for helping participants guess the identity of the critical letter ‹n›. This led to the development of an experimental paradigm that involved a forced-choice test for letters embedded in words, nonwords, and in isolation (Reicher, 1969; Wheeler, 1970). By providing the participant with two plausible response alternatives (e.g., *bore* vs. *born*), guessing is ruled out as an explanation, along with other interpretations of Cattell's original observation. Remarkably, the superior reporting of letters embedded in words, compared with when they were embedded in nonwords or presented in isolation, was upheld. This became known as the *word superiority effect* or the *Reicher-Wheeler effect*.

The theoretical significance of the word superiority effect is profound because one is confronted with the following conundrum: If letters are a necessary first step for recognizing a word, how can word-level information influence the perception of the letters making up the word? This effect stimulated the highly influential interactive activation model of letter perception developed by McClelland and Rumelhart (1981) and Rumelhart and McClelland (1982) (see Figure 3.1). This powerful computational model involves three levels (features, letters, and words) and two types of connections across representations—facilitatory (represented by arrows) and inhibitory (represented by filled circles). Presenting a word activates the feature-, letter-, and word-level representations consistent with that word. Importantly, as word-level nodes receive activation, they begin to provide feedback to position-specific letters. This additional top-down influence of word-level on letter-level representations drives the word superiority effect.

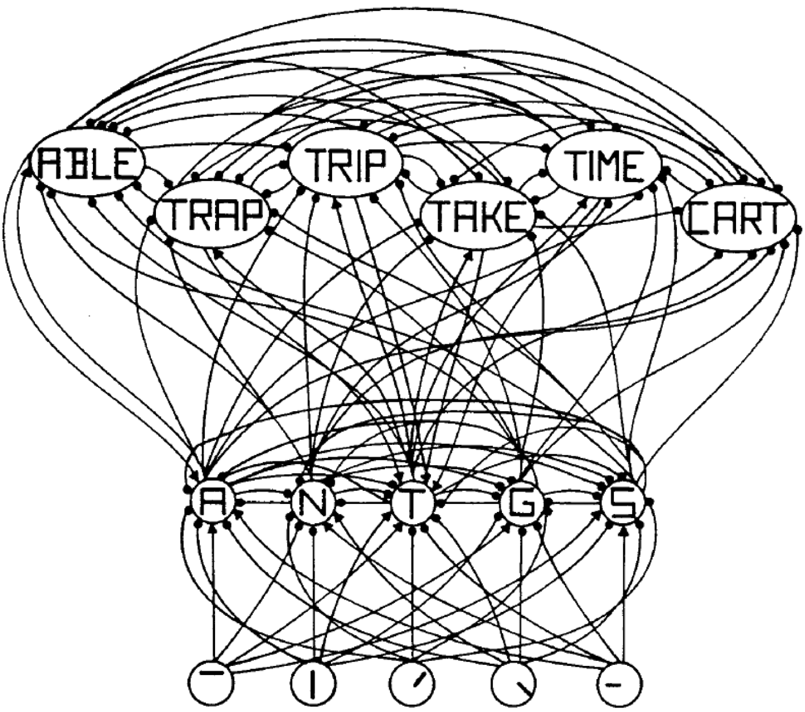


Fig. 3.1 McClelland and Rumelhart's (1981) interactive activation model of letter recognition.

The interactive activation model is historically important for many reasons. First, the model emphasized cascaded, rather than staged, processing (see McClelland, 1979), wherein all nodes accumulate activation across time via the spread of activation and inhibition across the connection paths. Second, the activation dynamics of all units are constrained by the activation and inhibition of other similarly spelled words (i.e., neighbors). This is an important difference from the classic Logogen model developed by Morton (1970), wherein lexical representations (logogens) accumulate activation across time independently of each other. Third, the interactive activation framework is a critical component of a number of computational models of visual word recognition, and predates the principles of the parallel distributed processing (PDP) approaches described in the next section.

Models and Tasks of Lexical Processing

Although the interactive activation model (McClelland & Rumelhart, 1981) contains word-level representations, it was primarily developed to explain letter-rather than word-recognition performance. However, forced-choice letter recognition is rather removed from word-level processing, and one should consider tasks that reflect processes at the word level. Many tasks have been developed to investigate lexical-level processing, including *category verification* and *semantic classification* (e.g., classifying a word as living or nonliving), *perceptual identification* (identifying a perceptually degraded stimulus), and *reading* (with eye-fixation durations on a target word measured). Although all of these tasks have important advantages and some disadvantages, here we focus on two tasks that have been dominant in work on isolated word recognition, *speeded pronunciation* (reading a word or nonword, e.g., *flirp*, aloud) and *lexical decision* (classifying letter strings as words and nonwords via a button press). In these two tasks, researchers respectively measure the amount of time needed by participants to initiate the pronunciation of a word or to press a button. Both tasks a priori appear to map onto processes involved in a word-level representation, reaching threshold to produce the appropriate response, either the correct pronunciation or the correct word/nonword response.

MODELS OF SPEEDED PRONUNCIATION

We will first consider computational models of word-pronunciation performance, since this task has been particularly influential in model development. Our focus is on models of English pronunciation, although it should be noted that models have been implemented in other languages (e.g., French; Ans, Carbonnel, & Valdois, 1998). Historically, there have been two major classes of models of speeded pronunciation: dual-route models and single-route models. The dual-route cascaded (DRC) model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) has two distinct pathways for pronouncing a word aloud: a direct *lexical* route that maps the full visual letter string onto a lexical representation and an assembled *sublexical* route that maps the letter string onto its pronunciation based on abstract grapheme-phoneme correspondence rules (see Figure 3.2). These rules (e.g., ‹k› → /k/) were selected on purely statistical grounds; that is, /k/ is the phoneme most commonly associated with ‹k› in English monosyllables. The DRC model accounts for many findings in the visual word recognition literature. One particularly important finding is the frequency by regularity interaction. That is, *regular* words that adhere to abstract grapheme-phoneme correspondence rules (e.g., ‹k› → /k/) are pronounced faster than *irregular* words (those that violate the rules, e.g., *pint*), and this effect is exaggerated for words that are rarely encountered in printed language. This result follows the assumption that the lexical route (based on whole-word representations) is frequency modulated, but the assembled route (based on smaller sublexical units) is insensitive to whole-word frequency. Hence, irregular low-frequency words (e.g., *pint*) are recognized more slowly than regular low-frequency words (e.g., *hint*), because the two routes produce conflicting pronunciations for *pint*, and extra time is needed to resolve the competition before the correct pronunciation can be produced. In contrast, for high-frequency words, the difference in recognition times for regular (e.g., *save*) and irregular (e.g., *have*) words is attenuated or absent, because the lexical route produces an output before there is competition from the slower sublexical route.

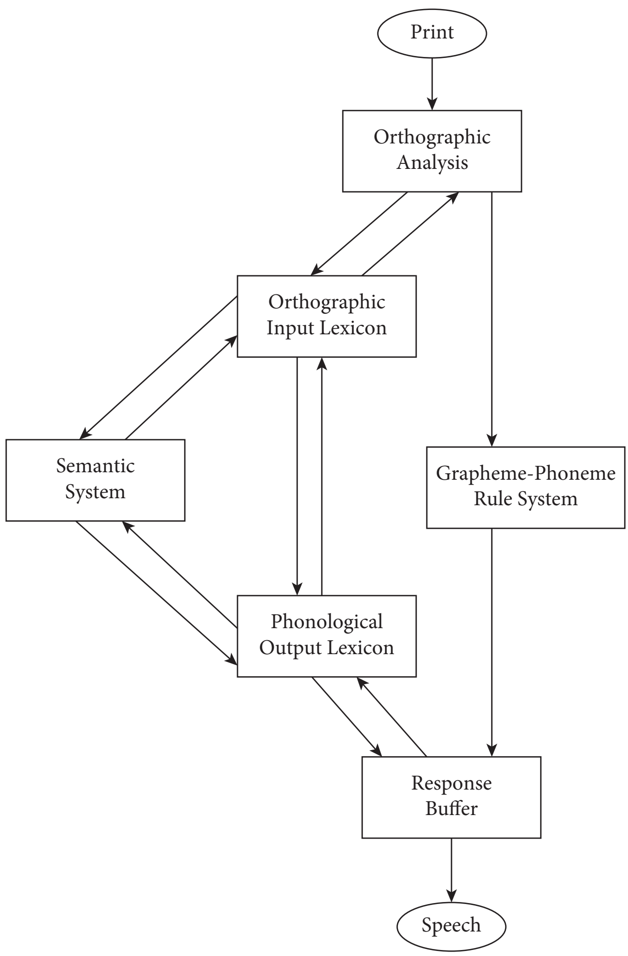


Fig. 3.2 Coltheart et al.'s (2001) DRC model of visual word recognition and reading aloud.

Coltheart et al. (2001) noted that a dual-route model also easily accommodates an important neuropsychological dissociation between acquired surface and phonological dyslexia. Individuals with surface dyslexia appear to have a breakdown in the lexical route, since they are relatively good at pronouncing nonwords and regularize words that do not conform to English spelling-to-sound rules (i.e., they pronounce *pint* such that it rhymes with *hint*). In contrast, individuals with phonological dyslexia appear to have a breakdown in the sublexical route such that they have particular difficulty with nonwords but are relatively good at pronouncing both regular and irregular words, which have lexical representations.

The second major class of models of speeded pronunciation is nicely reflected in the parallel distributed connectionist model developed by Seidenberg and McClelland (1989). The general structure of this model is displayed in Figure 3.3, in which a set of input units codes the orthography of the stimulus and these units map onto a set of hidden units, which in turn map onto a set of phonological units that code the pronunciation of the stimulus. Initially, the pathway weights are set to random levels. Gradually, through the learning mechanism of *backpropagation* (a common method for training computational neural networks), the connections across levels are adjusted to capture the correct pronunciation when a given orthographic string is presented. This model was trained on over 2,400 single-syllable words; the number of times a word is presented to the model is related to its frequency of occurrence in the language. Remarkably, after training, Seidenberg and McClelland found that the network produced many of the effects observed in speeded pronunciation performance. A particular noteworthy finding is that this connectionist network was able to account for the frequency by regularity interaction noted above. Importantly, the connectionist perspective is appealing because (1) it includes a learning mechanism; (2) it does not contain any formal spelling-to-sound "rules," but instead mimics rule-like behavior based on the statistical properties of spelling-to-sound mappings (see discussion of *consistency effects* later); and (3) it involves one, as opposed to two, pathways for pronunciation.

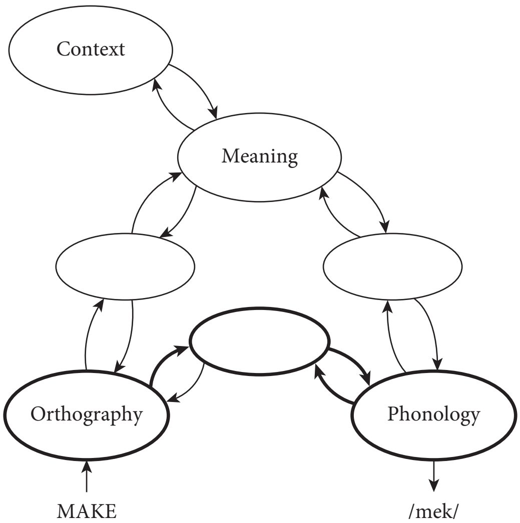


Fig. 3.3 Seidenberg and McClelland's (1989) parallel distributed processing model.

A hybrid model of speeded pronunciation called developed by Perry, Ziegler, and Zorzi (2007) was the CDP+ (connectionist dual process) model. The CDP+ model is very much like Coltheart et al.'s (2001) model, except that the DRC model's rule-based sublexical route is replaced by a two-layer connectionist network that learns the most reliable spelling-sound relationships in the language. This model is important because it not only accommodates the major empirical benchmarks in the literature but also accounts for considerably more item-level word recognition variance in large-scale databases (see discussion of *megastudies* later). A disyllabic version of this model, the CDP++ model, is also available (Perry, Ziegler, & Zorzi, 2010). The extension to disyllabic words is important because most major word recognition models have focused on single-syllable words (for an exception, see Ans et al., 1998). However, the majority of English words are multisyllabic, which involve additional processing demands such as syllabification and stress assignment. In this light, the CDP++ model is an important advance that extrapolates dual-route and connectionist principles to a much larger set of words.

Models of Lexical Decision Performance

The modeling of lexical decision performance has taken a somewhat different path than the modeling of speeded word pronunciation. This is not surprising, since the demands of producing the correct pronunciation for a visual letter string are quite different from the demands of discriminating familiar words from unfamiliar nonwords. For example, within the DRC model, a deadline mechanism has been implemented to simulate lexical decision (Coltheart et al., 2001). That is, a word response is produced when lexical activity in the orthographic lexicon exceeds some threshold, while a nonword response is made if lexical activity does not exceed that threshold after some deadline has elapsed (see also Grainger & Jacobs, 1996). The connectionist network can also be embellished to distinguish between words and nonwords by monitoring a measure of familiarity based on semantic activity (Plaut, 1997). Both approaches are useful for making contact with the lexical processing literature.

In contrast to these models, there are more general approaches that focus on the binary decision processes involved in the lexical decision task. One early model in this area was proposed by Balota and Chumbley (1984; also see Balota & Spieler, 1999). According to this model, lexical decisions can be based on two processes: a relatively fast-acting familiarity-based process and a slower, more attention-demanding process that checks the specific spelling or meaning of a given stimulus. This model was useful for emphasizing the decision-related processes in this task, further underscoring the distinction between task-general and task-specific processes in lexical decision. More recently, computational models of lexical decision have been developed that also emphasize the decision process. For example, Ratcliff, Gomez, and McKoon's (2004) diffusion model assumes that decisions are produced by a process that accumulates noisy information over time from a starting point toward a word or nonword boundary. This model is noteworthy because it captures not only mean response time and accuracy but also response time distributions for both correct and incorrect responses. Hence, this model captures the full range of behavior within the lexical decision task, a problem for previous models. An alternative approach is the Bayesian Reader model developed by Norris (2006). This model assumes that readers in the lexical decision task behave like optimal decision-makers who compute the probability that the presented letter string is a word rather than a nonword, given the input (see Kinoshita, this volume, for further discussion).

It should be evident from the foregoing discussion that models of lexical decision performance are quite different from their speeded-pronunciation counterparts. The latter emphasize processes mediating spelling-to-sound translation, whereas the former emphasize processes mediating word/nonword discrimination. Indeed, the effect sizes of major variables differ remarkably across lexical decision and speeded pronunciation (e.g., Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). Hence, a flexible and adaptive lexical-processing system is more consistent with the extant literature than one that is relatively static and modular. One such framework is presented in Figure 3.4, wherein one can see how task demands may emphasize different pathways within a more general lexical architecture (Balota & Yap, 2006). Of course, this is simply a general perspective, but the potentially crucial point

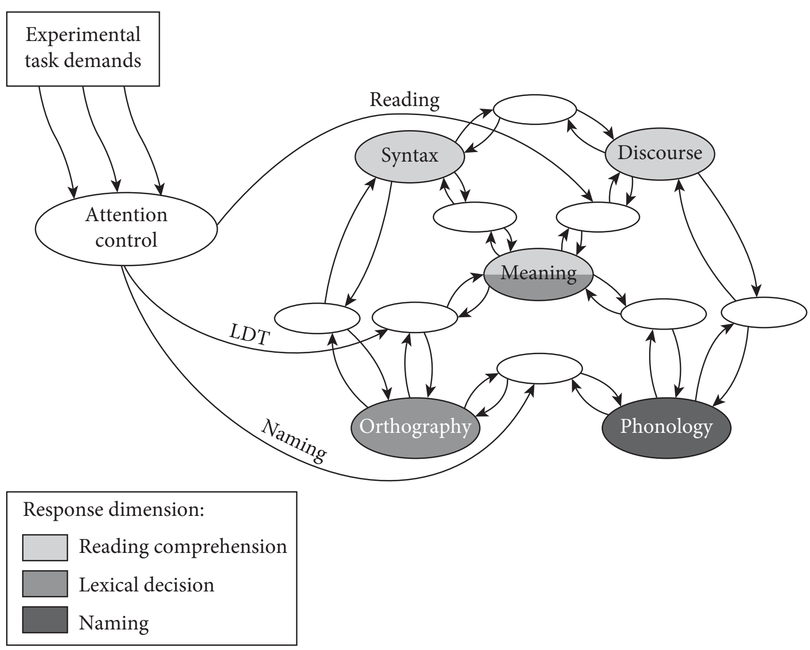


Fig. 3.4 The flexible lexical processor.

is that the lexical processing system adaptively considers different sources of information to maximize performance in response to the demands of a task.

In sum, the visual word recognition domain has provided a powerful test bed for the development of both metaphorical and computational models of mapping visual patterns onto phonology and meaning. This section provides only a snippet of some of the historical developments. Armed with these theoretical perspectives, we now turn to an analysis of how aspects of the empirical literature are interpreted within these models.

Lexical- and Semantic-Level Influences on Word Recognition

In order to better understand the processes underlying visual word recognition, researchers have identified how the many statistical properties associated with words (e.g., frequency of occurrence, number of letters, imageability) influence performance on different word recognition tasks. In this next section, we selectively review the impact of the most important *lexical* variables, which are quantified at the level of the whole word. There is also a rich literature examining the functional *sublexical* units (i.e., representations smaller than a word, such as letters, morphemes, and syllables) mediating word recognition (Carreiras & Grainger, 2004), but this is beyond the scope of the present chapter and is covered in other chapters (see Taft, this volume, and Perea, this volume).

Word Frequency

The frequency with which a word appears in print is the most robust predictor of word recognition performance (Whaley, 1978). Across virtually all lexical processing tasks, participants respond more quickly and accurately to high-frequency than low-frequency words. The word-frequency effect yields important insights into the nature of the human information-retrieval mechanism (Murray & Forster, 2004) and represents a fundamental constraint for all word recognition models. Despite its apparent simplicity, the theoretical interpretation of the word-frequency effect is far from straightforward (see also Kinoshita, this volume).

For example, one general class of lexical access models involves a type of serial *search* or *verification* process (Becker, 1980; Forster, 1976; Paap, McDonald, Schvaneveldt, & Noel, 1987), in which candidates compatible with the initial analysis of the stimulus are compared (or verified) against the visually presented letter string in descending order of frequency. The influential interactive activation model (e.g., Coltheart et al., 2001; McClelland & Rumelhart, 1981; Perry et al., 2007) described earlier assumes that the resting-level activations or activation thresholds of words (logogens in Morton's, 1970, nomenclature) vary with frequency of exposure. High-frequency words are responded to faster because they have higher resting-activation levels (or lower thresholds), thereby requiring less stimulus information to be recognized. Of course, within the connectionist frameworks (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989) that rely on distributed, rather than local, representations, frequency is coded by the strength of the weights between input and output representations. The Bayesian Reader model (Norris, 2006), which is predicated on the assumption that people recognize words in an optimal manner, takes a more functional approach to word-frequency effects. Specifically, word-frequency effects are a consequence of ideal observers taking the prior probabilities of words (indexed by their word frequencies) into account when resolving an ambiguous input as the stimulus unfolds during perception.

Researchers have also recently examined how different theoretical frameworks are able to account for the form of the relationship between word-frequency and word recognition measures. For example, a frequency-ordered serial search model predicts a linear relationship between the rank position of a word in a frequency-ordered list and access times, whereas the Bayesian Reader model predicts a logarithmic relationship between frequency and response times (Adelman & Brown, 2008). The work by Murray and Forster (2004) indicated that rank frequency was a better predictor of response times than log-transformed frequency, although this is qualified by more recent analyses by Adelman and Brown (2008) which suggest that word-frequency effects are most consistent with instance models (e.g., Logan, 1988) where each encounter with a word leaves an instance or trace in memory. The functional form of the word-frequency effect has been particularly well studied because researchers have developed large databases of lexical-decision and speeded-pronunciation performance while concurrently generating much better estimates of word frequency within the language (e.g., Brysbaert & New, 2009).

Although printed word frequency plays a central role in lexical access, there is also ample evidence that word-frequency effects partly implicate task-specific processes occurring after lexical access.

For example, in lexical decision, participants may particularly attend to the familiarity and meaningfulness of the letter string to help them discriminate between words and nonwords. This emphasis on familiarity-based information consequently exaggerates the frequency effect in lexical decision, compared with pronunciation (Balota & Chumbley, 1984). Specifically, low-frequency words are more similar to nonwords on the dimension of familiarity/meaningfulness than are high-frequency words. It is therefore more difficult to discriminate low-frequency words from nonwords, thereby slowing response times to low-frequency words and making the frequency effect larger. Indeed, researchers who have manipulated the overlap between words and nonwords by varying nonword wordlikeness (e.g., *brnta, brant, brane*; see Stone & Van Orden, 1993) report that such manipulations modulate the size of the word-frequency effect. The important point here is that frequency effects (and probably most other psycholinguistic effects) do not unequivocally reflect word recognition processes.

Length

Length here refers to the number of letters in a word. In perceptual identification, lexical decision, pronunciation, and reading, one generally observes longer latencies for longer words (see New, Ferrand, Pallier, & Brysbaert, 2006, for a review). Although the length effect is partly attributable to processes (e.g., early visual or late articulatory) that are beyond the scope of word recognition models, simulations indicate that the inhibitory influence of length on pronunciation onset latencies is especially difficult to reconcile with models that fully rely on parallel processing (e.g., Plaut et al., 1996). Instead, length effects are more compatible with models that incorporate serial processing, such as the DRC model (Coltheart et al., 2001), which contains a sublexical pathway that assembles phonology in a serial, letter-by-letter manner (Rastle & Coltheart, 2006). In fact, Weekes (1997) found that length effects are particularly large for nonwords compared with words, consistent with the DRC model perspective that length effects primarily reflect the influence of the sublexical pathway.

Orthographic and Phonological Similarity

In their classic study, Coltheart, Davelaar, Jonasson, and Besner (1977) explored the effects of an orthographic similarity metric they termed *orthographic neighborhood size* on lexical decision. Orthographic neighborhood size is defined by the number of orthographic neighbors associated with a letter string, where an orthographic neighbor is any word that can be obtained by substituting a single letter of a target word (e.g., *sand*'s neighbors include *band, send, said*, and *sank*). Assuming that lexical retrieval involves a competitive process, one might expect words with many neighbors to elicit more competition and hence produce slower response latencies. However, a review by Andrews (1997) suggested that across a number of languages, both lexical decision and pronunciation latencies are generally faster for words with many neighbors, and this effect is larger for low-frequency than for high-frequency words. The facilitatory effects of neighborhood size appear to be difficult to accommodate within any model (e.g., DRC model) that includes an interactive activation mechanism (McClelland & Rumelhart, 1981), because there should be more within-level inhibition to words with more orthographic neighbors. In addition to number of neighbors, researchers (e.g., Sears, Hino, & Lupker, 1995) have also considered the influence of *neighborhood frequency* (i.e., whether the target word possesses a higher-frequency neighbor, see Perea, this volume, for a discussion of such effects).

Like orthographic similarity, phonological similarity is defined by counting the number of *phonological neighbors*, that is, words created by changing a single phoneme of a target word (e.g., *gate*'s neighbors include *hate, get*, and *bait*). Yates (2005) and Yates, Friend, and Ploetz (2008a) have shown that in lexical decision, speeded pronunciation, semantic classification, and reading, words with many phonological neighbors are responded to faster than words with few phonological neighbors. There is also evidence that as the number of phonological neighbors overlapping with the *least supported phoneme* (i.e., the phoneme position within a word with which the fewest phonological neighbors coincide) increases, pronunciation latencies become faster (Yates, Friend, & Ploetz, 2008b). Generally, these results are consistent with the idea that words with many phonological neighbors receive additional activation within the phonological system, and help provide useful constraints for how phonology plays a role in word recognition.

The original definition of neighborhood size is somewhat restrictive. For example, a neighbor had to be matched in length to the target and differing only by the substitution of a single letter or phoneme. More expansive and flexible metrics of neighborhood size have been proposed (see Perea, this volume), including one based on the mean

Levenshtein distance (i.e., the number of single letter insertions, deletions, and substitutions needed to convert one string of elements to another) between a target word and its closest 20 neighbors in the lexicon. This measure (OLD20) has been shown to be a particularly powerful predictor for longer words (Yarkoni, Balota, & Yap, 2008).

Regularity and Consistency

As described earlier, the *regularity* of a word is defined by whether it conforms to the most statistically reliable spelling-to-sound correspondence rules in the language. *Hint* is regular because it follows these rules, whereas *pint* is irregular because it does not. Another theoretically important variable that quantifies the relationship between spelling and sound is *consistency*, which reflects the extent to which a word is pronounced like similarly spelled words. For example, *kind* is considered consistent because most similarly spelled words (e.g., *bind, find, hind, mind*) are pronounced the same way. In contrast, *have* is inconsistent because its pronunciation is different from most similarly spelled words (e.g., *cave, gave, save*). Generally, consistent words are recognized faster than inconsistent words, and the consistency effect is stronger in speeded pronunciation than in lexical decision, because the pronunciation task emphasizes the generation of the correct phonology (Jared, 2002). Such graded consistency effects fall naturally out of the connectionist perspective, where there is no sharp dichotomy between items that obey the "rules" and items that do not. Instead, lexical processing reflects the statistical properties of spelling-sound mappings at multiple grain sizes (Plaut et al., 1996). Consistency effects appear to pose a special challenge for the DRC model (Coltheart et al., 2001), which has some difficulty simulating them (Zevin & Seidenberg, 2006).

Although regularity and consistency correlate highly, these dimensions are separable. Distinguishing between these two variables is particularly valuable for adjudicating between the rule-based DRC approach (which predicts regularity effects) and the connectionist approach (which predicts consistency effects). Indeed, Cortese and Simpson (2000) crossed these two variables factorially in a speeded pronunciation experiment, and compared their results with simulated data from three computational models of word recognition. They observed stronger effects of consistency than regularity, a pattern that was captured best by Plaut et al.'s (1996) PDP model.

The above-mentioned studies have all emphasized the consistency of the *rime* unit (i.e., the vowel and consonant cluster after the onset of a syllable); *bind, find, hind*, and *mind* are all rime neighbors of *kind*. However, Treiman, Kessler, and Bick (2003) showed that the pronunciation of a vowel can also be influenced both by the consistency of its onset and coda. Thus, consistency in pronunciation appears to be sensitive to multiple grain sizes.

Semantic Richness

A growing number of reports in the literature indicate that word recognition is facilitated for semantically *richer* words (i.e., words that are associated with relatively more semantic information; for reviews, see Balota, Ferraro, & Connor, 1991; Pexman, 2012). This is theoretically intriguing because in virtually all models of word recognition, it would appear that a word has to be recognized before its meaning is obtained (Balota, 1990). This is at odds with available empirical evidence which suggests that the system has access to meaning before a word is fully identified, possibly via feedback activation from semantic to orthographic and phonological units (Balota et al., 1991; Pexman, 2012). Although the ultimate goal of reading is to extract meaning from visually printed words, the influence of meaning-level influences on word recognition remains poorly understood.

A number of dimensions have been identified that appear to tap the richness of a word's semantic representation, including the *number of semantic features* associated with its referent (McRae, Cree, Seidenberg, & McNorgan, 2005); its *number of semantic neighbors* (Shaoul & Westbury, 2010); the *number of distinct first associates* elicited by the word in a free-association task (Nelson, McEvoy, & Schreiber, 1998); *imageability*, the extent to which a word evokes mental imagery (Cortese & Fugett, 2004); *number of senses*, the number of meanings associated with a word (Miller, 1990); *body-object interaction*, the extent to which a human body can interact with a word's referent (Siakaluk, Pexman, Aguilera, Owen, & Sears, 2008); and *sensory experience ratings*, the extent to which a word evokes a sensory or perceptual experience (Juhasz & Yap, 2013). Across tasks, words from denser semantic neighborhoods, which possess more meanings and evoke more imagery, and whose referents are associated with more features or are easier for the human body to interact with are recognized faster (e.g., Yap, Pexman, Wellsby, Hargreaves, & Huff, 2012).

Importantly, the different richness variables account for unique (i.e., nonoverlapping) variance in word recognition performance (Yap, Pexman, et al., 2012), implying that no single richness dimension (and its associated theoretical framework) can adequately explain how meaning is derived from print. Instead, semantic memory is best conceptualized as multidimensional (Pexman, Siakaluk, & Yap, 2013).

In addition to the richness dimensions described above, the emotional *valence* (positive, neutral, negative) and *arousal* of a word influence lexical decision and speeded pronunciation performance. For example, *snake* is a negative, high-arousal word, while *sleep* is a positive, low-arousal word. A number of early studies suggested that negative, compared with neutral and positive, stimuli are responded to more slowly. This slowing is consistent with the idea that negative stimuli attract attention in early processing, and more time is needed to disengage attention from these stimuli before a lexical decision or pronunciation response can be made (see Kuperman, Estes, Brysbaert, & Warriner, 2014, for a review). However, this conclusion is qualified by a meta-analysis revealing that the negative and neutral words used in the studies were not always well matched on lexical characteristics (Larsen, Mercer, & Balota, 2006). Although the results of better-controlled studies are somewhat mixed, a recent large-scale analysis of valence and arousal effects for over 12,000 words, which controlled for many lexical and semantic factors, suggests that valence and arousal exert independent and monotonic effects, such that negative (compared with positive) and arousing (compared with calming) words are recognized more slowly (Kuperman et al., 2014).

Finally, an intriguing aspect of the semantic richness literature involves the extent to which is that the strength of these effects is modulated by the specific demands of a lexical processing task (Balota & Yap, 2006). For example, semantic richness accounts for much more item-level variance in the category verification task than in tasks where semantic processing is not the primary basis for responding. Yap, Tan, Pexman, and Hargreaves (2011) also found that words with more senses were associated with faster lexical decision times but less accurate category verification performance. This result is consistent with the notion that multiple meanings can hurt performance in a task that requires participants to resolve the specific meaning of a word.

Context/Priming Effects

Thus far we have described variables that influence isolated word recognition. There is also a rich literature directed at how different contexts or primes influence word recognition processes. In a typical priming paradigm, two letter strings are presented successively that have some dimension of similarity. Specifically, the two strings might be morphologically (*touching-TOUCH*), orthographically (*couch-TOUCH*), phonologically (*much-TOUCH*), or semantically/associatively related (*feel-TOUCH*). Primes can either be *unmasked* (i.e., consciously available) or *masked* (i.e., presented briefly to minimize conscious processing). The key advantage of the masked priming paradigm is that participants are usually unaware of the relationship between the prime and the target, thereby minimizing strategic effects (Forster, 1998; see also Kinoshita & Lupker, 2003). In this section, we limit our coverage to phonological, morphological, and semantic priming effects. Kinoshita (this volume) and Perea (this volume) provide excellent reviews of orthographic priming effects and discuss how this important work constrains models that address how readers code letter position in words (see also Frost, this volume).

Phonological Priming Effects

What is the role of phonological codes in visual word recognition (Frost, 1998)? Do these codes automatically precede and constrain the identification of words, or is phonology generated after lexical access? These controversial questions have been extensively investigated with the masked priming paradigm and other paradigms (see Halderman, Ashby, & Perfetti, 2012, for a review). For example, Lukatela and Turvey (2000) reported that compared with a control prime (e.g., *clep*), phonologically related primes (e.g., *klip*) facilitated lexical decision responses to targets (i.e., *CLIP*), even when primes were presented for only 14 ms. Indeed, in an important meta-analysis of masked phonological priming studies in English, Rastle and Brysbaert (2006) concluded that there were small but reliable effects of masked phonological priming in perceptual identification, pronunciation, and lexical decision. To confirm this, Rastle and Brysbaert (2006) conducted two masked priming experiments that demonstrated that words (e.g., *GROW*) were recognized 13 ms faster on average when they were preceded by phonologically similar primes (*groe*) than by orthographic controls (*groy*). Collectively, these results provide compelling evidence for an early and pervasive influence of phonological processes in word recognition. These phonological processes potentially help in stabilizing the identity of words so that they can be perceived accurately (Halderman et al., 2012; see also Pollatsek, this volume).

Morphological Priming Effects

Morphemes are the smallest units of meaning in words, and many English words are multimorphemic. An important debate in the literature concerns the extent to which the morphemic constituents in a word serve as access units during word recognition (see Taft, this volume). For example, are morphologically complex words such as *painter* automatically decomposed into their morphemic subunits (i.e., *paint* + *er*) prior to lexical access (Taft & Forster, 1975) or does each complex word have its own representation? Relatedly, does the morphological decomposition procedure distinguish between inflected words that are more semantically *transparent* (i.e., the meaning of the word can be predicted from its constituents, e.g., *sadness*) and words that are more semantically *opaque* (e.g., *department*)? The answers to such questions help shed light on the representations and processes underlying morphological processing.

To better delineate the time course of morphological processes, researchers rely heavily on the masked morphological priming paradigm. Using this tool, they have established that recognition of a target word (e.g., *SAD*) is facilitated by the masked presentation of morphologically related words (i.e., *sadness*) (Rastle, Davis, Marslen-Wilson, & Tyler, 2000). By using appropriate controls, Rastle et al. (2000) have shown that such morphological priming effects cannot be simply attributed to semantic or orthographic overlap between primes and targets, and hence provide compelling evidence for early and obligatory decomposition of morphologically complex words into morphemes prior to lexical access.

Interestingly, Rastle, Davis, and New (2004) have also reported that masked morphological priming effects are equivalent in magnitude for transparent (e.g., *cleaner—CLEAN*) and opaque (e.g., *corner—CORN*) prime-target pairs,[1] suggesting that the initial morphological decomposition process is blind to semantics and based entirely on the analysis of orthography. That being said, the role of semantics in morphological processing is still not entirely clear. A meta-analysis of the literature revealed a small but reliable effect of *semantic transparency*. That is, transparent primes facilitate target recognition to a greater extent than opaque primes (Feldman, O'Connor, & del Prado Martin, 2009), consistent with an early semantic influence on morphological processing (but see Davis & Rastle, 2010).

These patterns are theoretically important because they challenge the connectionist frameworks which posit that morphemic effects emerge via interactions among orthography, phonology, and semantics (e.g., Gonnerman, Seidenberg, & Andersen, 2007); such frameworks predict less priming for opaque than for transparent prime-target pairs (Plaut & Gonnerman, 2000). For a more extensive discussion of the morphological processing literature, readers are encouraged to consult Diependaele, Grainger, and Sandra (2012).

"Semantic" Priming Effects

The *semantic priming effect* refers to the robust finding that words are recognized faster when preceded by a semantically related prime (e.g., *cat-DOG*) than when preceded by a semantically unrelated prime (e.g., *mat-DOG*) (Meyer & Schvaneveldt, 1971).[2] The semantic priming literature provides important insights into the architecture of the mental lexicon and the processes used to retrieve information from that network. The "semantic" in semantic priming effect is largely an expository convenience (McNamara, 2005), since the effect may reflect an associative relationship between the two words rather than an overlap in their semantic features. For example, *dog* and *cat* share both a semantic and associative relationship, whereas *mouse* and *cheese* primarily share an associative relationship. While a review by Lucas (2000) suggests there are instances where semantic priming effects truly reflect shared semantic information, a follow-up review by Hutchison (2003) yields the more guarded conclusion that a simple associative account can accommodate most of the priming literature. What else do we know about semantic priming?

Related primes facilitate target recognition even when primes are heavily masked and cannot be consciously identified (Balota, 1983; Fischler & Goodman, 1978), suggesting that the meaning of a prime word can be processed, even if it is not consciously identifiable. This claim is consistent with an intriguing phenomenon known as the *mediated priming effect*. In mediated priming, *lion* is able to prime *STRIPES* (Balota & Lorch, 1986). Although there is no obvious direct relationship between the two words, priming is able to occur through the mediating concept *tiger*. These results are consistent with the classic study by Neely (1977), who demonstrated that semantic priming effects can occur at short stimulus onset asynchronies even when attention is directed to a different area of semantic memory.

A number of theoretical mechanisms have been proposed to explain semantic priming; these mechanisms are not mutually exclusive and may well operate together (see McNamara, 2005, for a review). *Automatic spreading activation* (Posner & Snyder, 1975) is the canonical explanation for semantic priming. That is, a prime (e.g., *cat*) automatically preactivates related nodes (e.g., *DOG*) via associative/semantic pathways, facilitating recognition of these related words when they are subsequently presented (see Collins & Loftus, 1975). Priming may also partly reflect *expectancy*, or the strategic generation of potential candidates for the upcoming target (Becker, 1980); facilitation is observed when the expectancy is correct. Finally, there is evidence that priming effects in the lexical decision task implicate postlexical decision processes. Specifically, participants may engage in backward semantic checking from the target to the prime (Neely, Keefe, & Ross, 1989), since the absence or presence of a prime-target relationship is diagnostic of the target's lexicality (nonwords are never related to the primes). Space constraints preclude a comprehensive survey of this interesting and important area of research, but readers are directed to Neely (1991) and McNamara (2005) for excellent reviews of the semantic/associative priming literature.

Joint Effects of Variables

Heretofore we have emphasized the main effects of variables. However, researchers are typically more interested in the extent to which multiple variables interact to influence word recognition performance. Indeed, such interactions are particularly useful for constraining theoretical models. For example, stimulus length interacts with orthographic neighborhood size, such that there is an increasing facilitatory effect of orthographic neighborhood size for long, compared to short, words (see Balota et al., 2004). In addition, low-frequency words produce larger effects of both orthographic neighborhood size and length than high-frequency words (Balota et al., 2004) in the speeded pronunciation task, but not in the lexical decision task. It is possible that the reduced effects of variables for high-frequency words may reflect better established lexical representations for these items.

There is also considerable evidence for interactions within the priming literature. For example, semantic priming typically interacts with word frequency and stimulus quality, such that priming effects are larger for low-frequency (e.g., Becker, 1979) and degraded (Becker & Killion, 1977) word targets. However, stimulus quality and word frequency produce robust additive effects (Stanners, Jastrzembski, & Westbrook, 1975) in the lexical decision task but not in either the word pronunciation or semantic classification task (Yap & Balota, 2007). There is also recent evidence that priming produces additive effects with the difficulty of the nonword distracters in the lexical decision task (Lupker & Pexman, 2010). Traditional priming accounts (e.g., spreading activation, expectancy) are too simple to capture this complex constellation of additive and interactive effects (McNamara, 2005), and it may be necessary to turn to models that possess multiple stages or levels of lexical-semantic representation (for an example, see Yap, Balota, & Tan, 2013). An important next step within computational modeling will be the development of models that can account for both the additive and interactive effects of targeted variables (see Plaut & Booth, 2000, 2006, for a potential framework, and also Borowsky & Besner, 2006, for a discussion of limitations of this approach).

Newer Approaches and Analytic Tools in Visual Word Recognition Research

Megastudies Versus Factorial Studies of Word Recognition

The most common experimental design in word recognition research is the factorial design, where independent variables of interest are manipulated and extraneous variables are controlled for. Although this approach is useful, like all approaches, it has some limitations (see Balota, Yap, Hutchison, & Cortese, 2012, for a review). The *megastudy* approach allows the language to define the stimuli, rather than the experimenter selecting stimuli based on a limited set of criteria. In megastudies, researchers examine word recognition for very large sets of words, such as virtually all English monosyllabic words (Balota et al., 2004; Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995) or multisyllabic monomorphemic words (Yap & Balota, 2009). In addition to identifying the unique predictive power of a large set of targeted variables, along with their interactive effects (Balota et al., 2004), megastudies have proven valuable for adjudicating between computational models of word recognition

(Perry et al., 2007), comparing competing metrics of word frequency (Brysbaert & New, 2009), evaluating the impact of novel psycholinguistic variables (Juhasz & Yap, 2013; Yarkoni et al., 2008), exploring potential nonlinear functional relationships between factors and word recognition performance (New et al., 2006), and investigating the role of individual differences in word recognition (Yap, Balota, Sibley, & Ratcliff, 2012).

The megastudy approach is aided by the availability of freely accessible online databases containing lexical characteristics and behavioral data for large sets of words. For example, the English Lexicon Project (ELP; Balota et al., 2007; http://elexicon.wustl.edu) provides lexical decision and speeded pronunciation measures for over 40,000 English words, along with a search engine that indexes a wide variety of lexical variables (see also the British Lexicon Project; Keuleers, Lacey, Rastle, & Brysbaert, 2011). The ELP has stimulated a flurry of related megastudies in other languages, including the French Lexicon Project (Ferrand et al., 2010), the Dutch Lexicon Project (Keuleers, Diependaele, & Brysbaert, 2010), the Malay Lexicon Project (Yap, Rickard Liow, Jalil, & Faizal, 2010), and the Chinese Lexicon Project (Sze, Rickard Liow, & Yap, 2014). Researchers have been turning to crowd-sourcing tools such as Amazon's Mechanical Turk (Mason & Suri, 2012) or smartphone apps to rapidly collect norms (e.g., concreteness ratings; Brysbaert, Warriner, & Kuperman, 2014) and behavioral data (Dufau et al, 2011). Researchers have also recently started developing databases that explore the influence of context on word recognition. For example, the Semantic Priming Project (Hutchison et al., 2013; http://spp.montana.edu) and the Form Priming Project (Adelman et al., 2014), respectively, serve as behavioral databases of semantic priming and masked form priming performance.

While one might be concerned that large-scale data may not be sensitive to more subtle manipulations (e.g., the interaction between frequency and consistency; Sibley, Kello, & Seidenberg, 2009), recent analyses indicate that databases such as the English Lexicon Project reproduce the standard effects in the literature (Balota et al., 2012). Thus megastudies provide a useful complement to the factorial studies in the literature.

Analyses of Response Time Distributions

In the overwhelming majority of studies in word recognition, researchers compare the *mean* response time across different conditions to determine whether their data are consistent with the predicted hypotheses. To the extent that empirical response time distributions are symmetrical and experimental manipulations primarily shift distributions, this approach works quite well. However, empirical distributions are virtually always positively skewed, and experimental effects can both shift and modulate the shape of a distribution (Heathcote, Popiel, & Mewhort, 1991). Thus, relying solely on analyses comparing means is potentially both inadequate and misleading (Heathcote et al., 1991). Fortunately, a number of approaches have been developed for understanding the influence of variables on the underlying response time distribution. The first and ultimately optimal method is to fit the data to a computational model (e.g., diffusion model; Ratcliff, 1978) that is able to generate specific predictions about experimental effects on the characteristics of the response time distribution. In the absence of such a model, researchers can (1) evaluate the influence of manipulations on the parameters of a mathematical function (e.g., the ex-Gaussian function, the sum of the normal and exponential distribution) fitted to an empirically obtained response time distribution or (2) generate descriptive plots (e.g., quantile plots) of how a manipulation differentially affects different regions of the distribution.

By augmenting conventional means-based analyses with distributional methods, researchers have gained finer-grained insights into the processes underlying isolated word recognition and semantic priming (see Balota & Yap, 2011, for a selective review). Consider the classic semantic priming effect, in which participants recognize *CAT* faster when it is preceded by *dog* than by an unrelated word like *dig*. Across a series of studies, there is evidence that the semantic priming effect in highly skilled readers is purely mediated by distributional shifting (Balota, Yap, Cortese, & Watson, 2008). That is, the benefit afforded by a related prime is constant, regardless of target difficulty (for a replication in masked semantic priming, see Gomez, Perea, & Ratcliff, 2013). Distributional shifting is most consistent with the idea that for such readers priming reflects relatively modular processes, whereby primes preactivate related words through automatic spreading activation and provide readers with a processing head-start when the words are subsequently presented. When word identification is compromised in some way, priming is no longer entirely mediated by a shift; instead, priming effects increase monotonically as target difficulty increases. One sees this pattern when targets are visually degraded (Balota et al., 2008; Yap et al., 2013) or when less skilled readers are processing unfamiliar low-frequency words (Yap, Tse, & Balota, 2009). That is, when target identification is effortful, readers can strategically retrieve prime information to aid in resolving the target (Thomas, Neely, & O'Connor, 2012).

Although it is tempting to map distributional parameters or aspects of the response time distribution onto specific cognitive processes, it is important not to do this in the absence of converging evidence (Matzke & Wagenmakers, 2009). The key point here is that there is a growing literature which suggests that one can gain important insights into lexical processes by moving beyond simple measures of central tendency and considering response time distributional analyses.

Individual Differences

Empirical work and models of word recognition have traditionally focused on group-level performance (but see Zevin & Seidenberg, 2006, for an exception). However, there is compelling evidence that individual differences in reading skill can modulate word recognition performance (see Andrews, this volume; see also Yap, Balota, et al., 2012, for a review). For example, vocabulary knowledge appears to moderate the joint effects of priming and word frequency (Yap et al., 2009). For readers with smaller vocabularies, priming and word frequency interact; priming effects are larger for low-frequency words. In contrast, highly skilled readers with a larger vocabulary produce robust main effects of priming and word frequency but no interaction.

The advent of large datasets containing individual participant data makes it possible to explore individual differences with very large samples. For example, in their analysis of the trial-level lexical decision and speeded pronunciation data contributed by over 1,200 participants in the English Lexicon Project, Yap, Balota, et al. (2012) made a number of noteworthy observations. Importantly, Yap, Balota, et al. reported considerable within- and between-session reliability across distinct sets of items with respect to overall mean response time, response time distributional characteristics, diffusion model parameters, and effects of theoretically important variables such as word frequency and length. Readers with more vocabulary knowledge showed faster, more accurate word recognition performance and attenuated sensitivity to stimulus characteristics. Collectively, results such as these suggest that participants are associated with relatively stable distributional and processing profiles that extend beyond average processing speed. Moving forward, it will be increasingly important to develop models that can capture both group-level performance and the variability across individual readers.

Concluding Remarks

The research in visual word recognition provides exciting insights into the early stages of reading and has been the focus of important principles in cognitive modeling, including interactive activation, rule-based coding, connectionist modeling, and more recently, notions of optimal perceivers from a Bayesian perspective. Although there has been considerable progress, different tasks bring with them task-specific operations that can influence the results. Hence one must be cognizant of the interplay between task-general lexical processes and task-specific processes when considering this literature. Finally, because of space constraints, the reader should be reminded that this is at best a brief snapshot of the visual word recognition literature, and we have focused primarily on behavioral studies in adult readers. For example, research in cognitive neuroscience continues to provide important constraints for word recognition models (Taylor, Rastle, & Davis, 2013; see Woollams, this volume). We anticipate that visual word recognition will continue to be at the heart of fundamental breakthroughs in understanding how people read.

Notes

1 Rastle and colleagues do not distinguish between semantically opaque prime-target pairs that share both an etymological and surface morphological relationship (e.g., *department-DEPART*) and pairs that share only the surface relationship (e.g., *corner-CORN*), because such a distinction is difficult to reconcile with a plausible theory of language acquisition (Rastle & Davis, 2003). However, there are researchers (e.g., Longtin, Segui, & Hallé, 2003) who make this distinction and would consider *corner* a pseudoaffixed word.

2 The extent to which two words (e.g., *cat* and *dog*) are related is typically captured by free association norms (e.g., Nelson et al., 1998), which are derived from participants' responses to cue words. An alternative approach, which assumes that a word's meaning is tied to the contexts it appears in, examines the co-occurrence of words in a large text corpus (Landauer & Dumais, 1997). Word pairs which that co-occur more frequently are considered to be more strongly related (Jones, Kintsch, & Mewhort, 2006).

References

Adelman, J. S., & Brown, G. D. (2008). Modeling lexical decision: The form of frequency and diversity effects. *Psychological Review, 115,* 214–229.

Adelman, J. S., Johnson, R. L., McCormick, S. F., McKague, M., Kinoshita, S., Bowers, J. S., . . . Davis, C. J. (2014). A behavioral database for masked form priming. *Behavior Research Methods, 46,* 1052–1067.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4,* 439–461.

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multi-trace memory model of polysyllabic word reading. *Psychological Review, 105,* 678–723.

Appelman, I. B., & Mayzner, M. S. (1981). The letter-frequency effect and the generality of familiarity effects on perception. *Perception & Psychophysics, 30,* 436–446.

Balota, D. A. (1983). Automatic semantic activation and episodic memory encoding. *Journal of Verbal Learning and Verbal Behavior, 22*, 88–104.

Balota, D. A. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 9–32). Hillsdale, NJ: Erlbaum.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 340–357.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283–316.

Balota, D. A., Ferraro, F. R., & Connor, L. T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187–218). Hillsdale, NJ: Erlbaum.

Balota, D. A., & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 336–345.

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General, 128,* 32–55.

Balota, D. A., & Yap, M. J. (2006). Attentional control and flexible lexical processing: Explorations of the magic moment of word recognition. In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 229–258). New York, NY: Psychology Press.

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science, 20,* 160–166.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459.

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language, 59,* 495–523.

Balota, D. A., Yap, M. J., Hutchison, K. A., & Cortese, M. J. (2012). Megastudies: What do millions (or so) of trials tell us about lexical processing? In J. S. Adelman (Ed.), *Visual word recognition: Vol. 1. Models and methods, orthography and phonology* (pp. 90–115). Hove, England: Psychology Press.

Becker, C. A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 252–259.

Becker, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & Cognition, 8*, 493–512.

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 389–401.

Borowsky, R., & Besner, D. (2006). Parallel distributed processing and lexical-semantic effects in visual word recognition: Are a few stages necessary? *Psychological Review, 113,* 181–193.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41,* 977–990.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46,* 904–911.

Carreiras, M., & Grainger, J. (2004). Sublexical units and the 'front end' of visual word recognition. *Language and Cognitive Processes, 19,* 321–331.

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind, 11,* 63–65.

Collins, A., & Loftus, E. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82,* 407–428.

Coltheart, M., Davelaar, E., Jonasson, J., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments andComputers, 36,* 384–387.

Cortese, M. J., & Simpson, G. B. (2000). Regularity effects in word naming: What are they? *Memory & Cognition, 28,* 1269–1276.

Davis, M. H., & Rastle, K. (2010). Form and meaning in early morphological processing: Comment on Feldman, O'Connor, and Moscoso del Prado Martin (2009). *Psychonomic Bulletin & Review, 17,* 749–755.

Diependaele, K., Grainger, J., & Sandra, D. (2012). Derivational morphology and skilled reading: An empirical overview. In M. J. Spivey, K. McRae, & M. Joanisse (Eds.), *The Cambridge handbook of psycholinguistics* (pp. 311–333). Cambridge, England: Cambridge University Press.

Dufau, S., Duñabeitia, J. A., Moret-Tatay, C., McGonigal, A., Peeters, D., Alario, F.-X., . . . Grainger, J. (2011). Smart phone, smart science: How the use of smartphones can revolutionize research in cognitive science. *PLoS ONE, 6,* e24974.

Feldman, L. B., O'Connor, P. A., & del Prado Martin, F. M. (2009). Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review, 16,* 684–691.

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., . . . Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods, 42,* 488–496.

Fischler, I., & Goodman, G. O. (1978). Latency of associative activation in memory. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1731–1740.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. C. T. Walker (Eds.), *New approaches to language mechanisms* (pp. 257–287). Amsterdam, the Netherlands: North-Holland.

Forster, K. I. (1998). The pros and cons of masked priming. *Journal of Psycholinguistic Research, 27,* 203–233.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123,* 71–99.

Gomez, P., Perea, M., & Ratcliff, R. (2013). A diffusion model account of masked vs. unmasked priming: Are they qualitatively different? *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1731–1740.

Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General, 136,* 323–345.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103,* 518–565.

Grainger, J., Rey, A., & Dufau, S. (2008). Letter perception: From pixels to pandemonium. *Trends in Cognitive Sciences, 12,* 381–387.

Halderman, L. K., Ashby, J., & Perfetti, C. A. (2012). Phonology: An early and integral role in identifying words. In J. S. Adelman (Ed.), *Visual word recognition: Vol. 1. Models and methods, orthography and phonology* (pp. 207–228). Hove, England: Psychology Press.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. K. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin, 109,* 340–347.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology, 160,* 106–154.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A micro-analytic review. *Psychonomic Bulletin & Review, 10,* 785–813.

Hutchison, K. A., Balota, D. A., Neely, J. H., Cortese, M. J., Cohen-Shikora, E. R., Tse, C-S., . . . Buchanan, E. (2013). The semantic priming project. *Behavior Research Methods, 45*, 1099–1114.

Jared, D. (2002). Spelling-sound consistency and regularity effects in word naming. *Journal of Memory and Language, 46,* 723–750.

Jones, M. N., Kintsch, W., & Mewhort, D. J. K. (2006). High-dimensional semantic space accounts of priming. *Journal of Memory and Language, 55,* 534–552.

Juhasz, B. J., & Yap, M. J. (2013). Sensory experience ratings (SERs) for over 5,000 mono- and disyllabic words. *Behavior Research Methods, 45,* 160–168.

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology, 1,* 174.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2011). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods, 44,* 287–304.

Kinoshita, S., & Lupker, S. J. (Eds.). (2003). *Masked priming: State of the art*. New York, NY: Psychology Press.

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology: General, 143,* 1065–1081.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211–240.

Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion, 6,* 62–72.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Longtin, C.-M., Segui, J., & Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes, 18,* 313–334.

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7,* 618–630.

Lukatela, G., & Turvey, M. T. (2000). An evaluation of the two-cycles model of phonology assembly. *Journal of Memory and Language, 42,* 183–207.

Lupker, S. J., & Pexman, P. M. (2010). Making things difficult in lexical decision: The impact of pseudohomophones and transposed-letter nonwords on frequency and semantic priming effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1267–1289.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44,* 1–23.

Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review, 16,* 798–817.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 86,* 287–330.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. Hove, England: Psychology Press.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods, 37,* 547–559.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227–234.

Miller, G. A. (1990). Word Net: An on-line lexical database. *International Journal of Lexicography, 3,* 235–312.

Morton, J. (1970). A functional model for memory. In D. A. Norman (Ed.), *Models of human memory* (pp. 203–254). New York, NY: Academic Press.

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review, 111,* 721–756.
Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106,* 226–254.
Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 236–264). Hillsdale, NJ: Erlbaum.
Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1003–1019.
Neisser, U. (1967). *Cognitive psychology*. New York, NY: Appleton-Century-Crofts.
Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. w3.usf.edu/FreeAssociation/.
New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review, 13,* 45–52.
Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review, 113,* 327–357.
Paap, K. R., McDonald, J. E., Schvaneveldt, R. W., & Noel, R. W. (1987). Frequency and pronounceability in visually presented naming and lexical-decision tasks. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 221–243). Hillsdale, NJ: Erlbaum.
Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114,* 273–315.
Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the connectionist dual process (CDP++) model. *Cognitive Psychology, 61,* 106–151.
Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1989). Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience, 1,* 153–170.
Pexman, P. M. (2012). Meaning-based influences on visual word recognition. In J. S. Adelman (Ed.), *Visual word recognition: Vol. 2. Meaning and context, individuals and development* (pp. 24–43). Hove, England: Psychology Press.
Pexman, P. M., Siakaluk, P. D., & Yap, M. J. (2013). Introduction to the research topic meaning in mind: Semantic richness effects in language processing. *Frontiers in Human Neuroscience, 7,* 723.
Plamondon, R., & Srihari, S. N. (2000). On-line and off-line handwriting recognition: A comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22,* 63–84.
Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes, 12,* 765–805.
Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review, 107,* 786–823.
Plaut, D. C., & Booth, J. R. (2006). More modeling but still no stages: Reply to Borowsky and Besner. *Psychological Review, 113,* 196–200.
Plaut, D. C., & Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes, 15,* 445–485.
Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.
Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.
Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology, 53,* 97–145.
Rastle, K., & Coltheart, M. (2006). Is there serial processing in the reading system; and are there lexical representations? In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 3–24). New York, NY: Psychology Press.
Rastle, K., & Davis, M. H. (2003). Reading morphologically complex words: Some thoughts from masked priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art* (pp. 279–305). Hove, England: Psychology Press.
Rastle, K., Davis, M. H., Marslen-Wilson, W., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes, 15,* 507–538.
Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review, 11,* 1090–1098.
Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.
Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111,* 159–182.
Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81,* 274–280.
Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60–94.
Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 876–900.
Seidenberg, M. S., & McClelland, J. L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review, 96*, 523–568.
Selfridge, O. G., & Neisser, U. (1960). Pattern recognition by machine. *Scientific American, 203*, 60–68.
Shaoul, C., & Westbury, C. (2010). Exploring lexical co-occurrence space using HiDEx. *Behavior Research Methods, 42,* 393–413.
Siakaluk, P. D., Pexman, P. M., Aguilera, L., Owen, W. J., & Sears, C. R. (2008). Evidence for the activation of sensorimotor

information during visual word recognition: The body-object interaction effect. *Cognition, 106,* 433–443.

Sibley, D. E., Kello, C. T., & Seidenberg, M. S. (2009). *Error, error everywhere: A look at megastudies of word reading*. Proceedings of the Annual Meeting of the Cognitive Science Society. Amsterdam, The Netherlands.

Stanners, R. F., Jastrzembski, J. E., & Westbrook, A. (1975). Frequency and visual quality in a word-nonword classification task. *Journal of Verbal Learning and Verbal Behavior, 14,* 259–264.

Stone, G. O., & Van Orden, G. C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 744–774.

Sze, W. P., Rickard Liow, S. J., & Yap, M. J. (2014). The Chinese Lexicon Project: A repository of lexical decision behavioral responses for 2,500 Chinese characters. *Behavior Research Methods, 46*, 263–273.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14,* 638–647.

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin, 139,* 766–791.

Thomas, M. A., Neely, J. H., & O'Connor, P. (2012). When word identification gets tough, retrospective semantic processing comes to the rescue. *Journal of Memory and Language, 66,* 623–643.

Treiman, R., & Kessler, B. (2007). Learning to read. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 657–666). Oxford, England: Oxford University Press.

Treiman, R., Kessler, B., & Bick, S. (2003). Influence of consonantal context on the pronunciation of vowels: A comparison of human readers and computational models. *Cognition, 88,* 49–78.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General, 124,* 107–136.

Treisman, A. (1999). Solutions to the binding problem: Progress through controversy and convergence. *Neuron, 24,* 105–110.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50A,* 439–456.

Whaley, C. (1978). Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior, 17,* 143–154.

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology, 1,* 59–85.

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 274–296.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language, 60,* 502–529.

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognitions: Insights from the English lexicon project. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 53–79.

Yap, M. J., Balota, D. A., & Tan, S. E. (2013). Additive and interactive effects in semantic priming: Isolating lexical and decision processes in the lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 140–158.

Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. J. (2012). An abundance of riches: Cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience, 6,* 72.

Yap, M. J., Rickard Liow, S. J., Jalil, S. B., & Faizal, S. S. B. (2010). The Malay Lexicon Project: A database of lexical statistics for 9,592 words. *Behavior Research Methods, 42,* 992–1003.

Yap, M. J., Tan, S. E., Pexman, P. M., & Hargreaves, I. S. (2011). Is more always better? Effects of semantic richness on lexical decision, speeded pronunciation, and semantic classification. *Psychonomic Bulletin & Review, 18,* 742–750.

Yap, M. J., Tse, C.-S., & Balota, D. A. (2009). Individual differences in the joint effects of semantic priming and word frequency: The role of lexical integrity. *Journal of Memory and Language, 61,* 303–325.

Yarkoni, T., Balota, D. A., & Yap, M. J. (2008). Beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15,* 971–979.

Yates, M. (2005). Phonological neighbors speed visual word processing: Evidence from multiple tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1385–1397.

Yates, M., Friend, J., & Ploetz, D. M. (2008a). The effect of phonological neighborhood density on eye movements during reading. *Cognition, 107,* 685–692.

Yates, M., Friend, J., & Ploetz, D. M. (2008b). Phonological neighbors influence word naming through the least supported phoneme. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1599–1608.

Zevin, J. D., & Seidenberg, M. (2006). Simulating consistency effects and individual differences in nonword naming. *Journal of Memory and Language, 54,* 145–160.

CHAPTER

4 The Work of the Eyes During Reading

Elizabeth R. Schotter *and* Keith Rayner

Abstract

This chapter discusses the movements of the eyes during reading: how they support and reveal the reading process. The chapter starts by describing basic facts about eye movements. It then describes some factors that account for variability in eye movement behaviors and then discusses some important methodological paradigms (namely gaze-contingent display change paradigms) that have contributed to our knowledge about eye movements and cognitive processing in reading. In particular, these paradigms have been used to study what types of information readers obtain from a word before they look directly at it. The chapter reviews what has been learned from experiments using these paradigms. It also discusses the issue of eye movement control in reading (i.e., what determines where the eyes move and how long they linger on a particular word) and describes several studies demonstrating that eye movement control in reading is determined directly by ongoing cognitive process.

Key Words: eye movement control, fixation duration, reading, perceptual span, preview effect

Eye movements are a critical part of the reading process. However, as Rayner and Pollatsek (1989) noted, we have the impression when we read that our eyes (and mind) sweep continuously across the text, and it is only when we encounter difficulty and pause to consider what we have just read or go back to reread earlier material that we are even aware of the movements of our eyes. That impression is an illusion, as the progress of the eyes across the page is not continuous. The eyes remain relatively still for periods called *fixations*, which usually last 150 to 500 ms (with the average being 200–250 ms). Between fixations the eyes move rapidly in what are called *saccades*, after the French word for *jump*. Saccades are *ballistic movements* (i.e., once they start, they cannot be altered). The eyes generally move forward about 7 to 9 letter spaces with each saccade for readers of English. The duration of a saccade varies with the distance moved, with a typical saccade in reading taking about 20 to 35 ms. As little visual information is extracted from the printed page during saccades (Matin, 1974), all useful visual information enters the reading system during fixations. The pattern of information extraction during reading is thus a bit like seeing a slide show (Rayner & Pollatsek, 1989). Readers see a slide for about a quarter of a second; there is then a brief off-time, and then a new slide of a different view of the page appears for about a quarter of a second. This pattern of fixations and saccades is not unique to reading. The perception of any static display (i.e., a picture or a scene) proceeds in the same way, although the pattern and timing of fixations differs from that in reading (Rayner, Li, Williams, Cave, & Well, 2007).

The second way in which our subjective impression is an illusion is that the eyes do not move relentlessly forward through the text. While most saccades do move forward (i.e., in English, left-to-right movements), about 10 to 15% move backward (i.e., right-to-left) and are termed *regressive saccades* (or *regressions* for short). Thus, since we make about 4 to 5 saccades per second, readers make a regression about once every 2 seconds. Readers are generally unaware of most regressions except those that reflect major confusion, requiring them to go back a considerable distance in the text to straighten things out. However, most regressions are quite short, only going back a few characters (to the preceding word or two). Often in these cases the regressions do not reflect comprehension difficulty, but instead are corrections of oculomotor error (e.g., overshooting the intended saccade target).

Another type of eye movement in reading is the *return sweep*, when the eyes move from near the end of one line to near the beginning of the next. While return sweeps are right-to-left movements they are not counted as regressions, since they move the reader forward through the linguistic progression of the text. Return sweeps are actually fairly complicated as they often start 5 to 7 letter spaces from the end of a line and they generally end on about the third to seventh letter space of the next line. Return sweeps generally fall short of their goal, and there is often an additional short right-to-left saccade after the large return sweep. However, the leftmost fixation is still sometimes on the second word of the line. Thus most of the time, about 80% of the line falls between the extreme fixations on it. As mentioned above, the small regressions following return sweeps are probably corrections for errors in aiming the eyes; it is difficult to execute a long saccade perfectly, with the eyes usually undershooting the target position.

Another important point about the general properties of eye movements during reading is that the two eyes are not always perfectly aligned on the same position in a word. For a long time it was assumed that the two eyes typically landed on the same letter in a word, or that they were perfectly aligned. However, while on over 50% of the fixations the two eyes are aligned on the same letter, they are on different letters quite often, and sometimes the two eyes are even crossed (i.e., the left eye is fixating further to the right than the right eye). While this is a fact about the characteristics of eye movements during reading, it is also true that how long the eyes remain in place is not dramatically affected by whether or not the two eyes are on the same letter (see Kirkby, Webster, Blythe, & Liversedge, 2008, for a review).

Silent Versus Oral Reading

In this chapter our focus will be on eye movements during silent reading. However, there are both similarities and differences between the eye movements readers make in silent and oral reading. Much of what we know about eye movements during oral reading stems from Buswell (1922), but there have been some recent investigations of eye movements during oral reading (see Ashby, Yang, Evans, & Rayner, 2012; Inhoff, Solomon, Radach, & Seymour, 2011) using much better and more accurate eye tracking systems than Buswell had available. Nevertheless, most of Buswell's findings have held up rather well.

What are the differences between silent and oral reading? The average fixation duration in oral reading is about 50 ms longer than in silent reading, the average forward saccade length is shorter, and there are more regressions. These differences are undoubtedly related to the fact that the eyes can move faster through the text in silent reading than the words can be pronounced in oral reading and readers don't want their eyes to get too far ahead of their voice. Thus there are places in the eye-movement record where readers are keeping their eyes in a holding pattern in oral reading so that this doesn't happen. The *eye-voice span*, the distance the eyes are ahead of the voice, is often the focus of research on oral reading. The main finding is that they eyes are typically about two words ahead of the voice.

Variations in Fixations and Saccades

The eyes move forward (about 7–9 letter spaces on average) in reading, but not relentlessly so. They pause for periods of approximately 150 to 500 ms (the large majority of the fixations are between 200 and 250 ms), and move backward about 10% to 15% of the time. However, there is considerable variation between and even within individual readers. Thus some readers have average fixation durations closer to 200 ms while others are closer to 300 ms, and some may only move about six letter spaces on average (with similar variations in the rate of regressions). But, it is also the case that, for any reader, individual fixation durations can range from under 200 ms to over 300 ms, and some saccades will be as short as 1 to 2 letter spaces and some will be longer than 10 letter spaces.

Viewing Distance Effects

When reading English, the average saccade of 7 to 9 letter spaces (about 2–3 degrees of visual angle) appears to be fundamental in that this is how far the eyes move regardless of the retinal size of the text (as long as the letters are not too big or too small). For example, regardless of whether a given text is 36 cm or 72 cm from the eyes, the average saccade length is still about 8 letters, even though 8 letters subtends twice the visual angle at 36 cm as it does at 72 cm (Morrison & Rayner, 1981; O'Regan, 1983). This fact suggests that the visibility of the text is relatively invariant to absolute size over an extended range of distances. As a result, data on saccade length are typically expressed in letter spaces; this appears to be the natural metric in reading rather than degrees of visual angle. A recent clever manipulation (the *parafoveal magnification paradigm*) has shown that when letters outside the center of vision are magnified in relation to their distance from fixation on each fixation (thereby compensating in some sense for the poorer acuity with letters outside the center of vision), how far the eyes move is still driven by number of characters (Miellet, O'Donnell, & Sereno, 2009). The fact that the distance of the text (and hence the absolute size of the letters) makes little difference for saccade size is probably due to a tradeoff between two factors: (1) when the text is nearer, the letters are bigger and easier to see; however, (2) when the text is nearer, a given letter will be further from the center of fixation, hence harder to see. Of course, there are limits; the text will be impossible to read if it is a mile away or against your face.

Orthographic Differences

Thus far we have focused mostly on results from studies of reading in English; however, English uses only one of many writing systems. Do the characteristics of eye movements change when people read text in other writing systems? The answer to this question is clearly "yes," as demonstrated by experiments that have examined the patterns of eye movements of Chinese and Japanese readers. However, a major problem with comparing saccade sizes in English with other languages is what to use as the unit of measurement. The previous section implied that the letter (or letter space) is the fundamental unit of measurement for English. However, there are no letters per se in either of these languages: the characters stand for morphemes and syllables. If one measures by characters (assigning letters the role of characters), then eye movements of Chinese and Japanese readers tend to be much smaller than eye movements of readers of English. Chinese readers move their eyes about 2 characters on average (Shen, 1927; a character stands for a morpheme rather than a word, so that this is less than two words). Readers of Japanese text, which is made up of characters that stand for morphemes (Kanji) and syllables (Kana), move their eyes about 3.5 characters (Ikeda & Saida, 1978). This is less than 3.5 words, since it often takes several characters to make a word. Since the average saccade in English is about 7 to 9 characters (about a word and a half), it appears that the average saccade length is, if anything, a bit less in English than in Chinese and Japanese if one equates for number of words or morphemes.

Readers of Hebrew also have smaller saccades (about 5.5 characters) than readers of English (Pollatsek, Bolozky, Well, & Rayner, 1981). Hebrew differs structurally and orthographically from English in some important ways. First, not all vowels are represented orthographically in Hebrew. In addition, many function words in Hebrew are attached to content words. The net effect of these differences is that Hebrew sentences normally contain fewer words and fewer letters than their English counterparts. In short, although Hebrew is basically an alphabetic system, the information is also more densely packed than in English.

The average saccade lengths of Chinese, Japanese, and Hebrew readers suggest that the informational density of the text determines how far the eyes move in each saccade. This finding seems consistent with the fact that, for readers of English, as the text becomes more difficult (and hence, the informational density is greater), saccade length decreases. However, it is an open question whether the differences in informational density across languages are best thought of in terms of the density of the meaning or the amount of visual information per character (measured perhaps by the number of strokes or lines in the character). For Hebrew, the characters seem of approximately equal complexity to English, so the differences between Hebrew and English are more likely to be explained by differences in amount of meaning per character. However, the Chinese and Japanese writing systems are so different from English that it is hard to say which type of informational density is operating to produce the differences in reading. We suspect that both the visual and semantic factors are contributing.

Fixation durations for readers of Japanese, Chinese, and Hebrew are fairly similar to those of readers of English. Despite the fact that reading in these languages is slower when measured superficially, the reading rates seem to be equivalent when measured in terms of amount of meaning extracted per unit time. In fact, when the reading rate in Hebrew was based on the number of words in the English translations of the Hebrew sentences, the average reading rate for the Hebrew-and

English-speaking subjects was nearly identical (Pollatsek et al., 1981).

The remainder of the current chapter deals with how visual information is extracted from text (see also Rayner, 1998, 2009; Schotter, Angele, & Rayner, 2012, for reviews). Thus we will focus on what useful information readers extract during fixations and on how the eyes are guided through text. It is important to note that readers' eye movements are very much influenced by the lexical and contextual properties of the fixated words. So, for example, how long readers look at a word is strongly influenced by factors like word frequency and word predictability. We will document these findings in more detail later, but for now it is important to keep in mind that how easy a word is to process has a large impact on how long the eyes remain on that word.

We turn now to three important issues with respect to eye movements in reading: (1) What is the size of the perceptual span (or region of effective vision) during reading, (2) what kind of information is integrated across eye movements in reading, and (3) what factors control where we fixate next and how long? However, before doing so, we will discuss the gaze-contingent display change paradigms that have contributed a great deal to our knowledge of the answers to the above questions.

Gaze-Contingent Display Change Paradigms

Moving Window/Moving Mask Experiments

Gaze-contingent techniques can be used to examine the use of foveal and parafoveal information during reading. The distinction between the fovea and parafovea relates to acuity in the visual field. The fovea is the center of vision (one degree from fixation in any direction) and has the highest acuity. Outside the fovea, acuity decreases with increasing distance from fixation. The parafovea (from 1–5 degrees away from the center of fixation) has moderate acuity, and the periphery (beyond 5 degrees from fixation) has the lowest acuity. One indication of how much readers rely on foveal and parafoveal information is how well they can read when only one of the two sources of information is available. To do this, McConkie and Rayner (1975; see also Rayner & Bertera, 1979) introduced a paradigm in which readers could only see the word in the fovea and a specific area around it. In this *moving window paradigm*, the eyes are monitored and valid information is provided within the window area, with the text outside the window replaced by

```
(A)  This is an example sentence in the moving window paradigm.
      *
     Thisxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.
          *
     xxxs is axxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.
                  *
     xxxxxxxxxxxexampxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.

(B)  This is an example sentence in the moving mask paradigm.
                 *
     This is anxxxxxple sentence in the moving mask paradigm.
                 *
     This ixxxxxxxxxxxxxsentence in the moving mask paradigm.

(C)  This is an example sentence in the boundary paradigm.
                                     *
     This is an example sentence in the xxxxxxxx paradigm.
                                           *
     This is an example sentence in the boundary paradigm.
```

Fig. 4.1 Examples of displays in gaze-contingent display change paradigms. Within each panel, the top row represents the sentence without any manipulation (all words are always visible). The following rows represent the display during a sequence of fixations, and the asterisk above each line represents the location of fixation. Panel A represents the display in the moving window paradigm (McConkie & Rayner, 1975) with a 5-character window condition (the fixated character and two on either side are visible and all other characters, including spaces, are masked with an x). Panel B represents the display in the moving mask paradigm (Rayner & Bertera, 1979) with a 5-character mask (middle line) and in a 13-character window (bottom line). Panel C represents the display in the boundary paradigm (Rayner, 1975) with an x-mask preview (middle line) that changes to the target word (boundary) once the reader makes a saccade to it.

other letters (typically with x's or random letters; see Figure 4.1, panel A). The extent to which reading is disrupted when only valid foveal information is available can be compared with the extent of disruption when only parafoveal information is available. To assess this, the *moving mask paradigm* masks foveal letters while retaining the letters in the parafovea and periphery (Rayner & Bertera, 1979; see also Fine & Rubin, 1999; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981; see Figure 4.1, panel B).

Gaze-Contingent Boundary Paradigm

While the moving window and moving mask paradigms are used to test globally how efficiently people can read with only foveal or parafoveal information, the *gaze-contingent boundary paradigm* (generally called the *boundary paradigm*; Rayner, 1975) is used to assess what type of information readers access from a word before it is fixated. In this paradigm there is an invisible boundary to the left of a particular *target* word in the sentence. This word is replaced with a different, *preview*, word (or nonword) while the reader fixates prior parts of the text. When the reader's eyes cross the boundary location, the preview changes to the target (see Figure 4.1, panel C). The display change occurs during a saccade when vision is effectively suppressed and therefore is generally not seen by the reader. The preview may share all (e.g., identical control condition) or very few properties with the target word (e.g., an unrelated word or random letters). If the target is processed faster (as evidenced by shorter fixation times) when the preview was related compared with when it was unrelated to the target word, this is considered *preview benefit*. While experiments designed to study preview benefit may seem unnatural (when we read, words do not normally change in front of our eyes), they must reflect some aspect of the underlying process of reading; there would be no preview benefit if the preview had not been processed parafoveally. As noted earlier, most subjects are not aware that any of the words are changing in such paradigms so it is unlikely that the experimental manipulation would alter their reading behavior. Furthermore, because the preview does not change to the target until the subject makes a saccade toward it, it is reasonable to assume that the reader would not process it differently than any other parafoveal word in the experimental sentence or, indeed, in a nonexperimental situation (Schotter et al., 2012).

What Is the Size of the Perceptual Span?

Research using the moving window paradigm has revealed that readers of English typically have a perceptual span (i.e., the area from which they obtain useful information) extending from 3 to 4 letter spaces to the left of fixation (McConkie & Rayner, 1976; Rayner, Well, & Pollatsek, 1980) to 14 to 15 letter spaces to the right of fixation (McConkie & Rayner, 1975; Rayner & Bertera, 1979; see Rayner, 2014, for further review). Another way to describe the size of the perceptual span is that it extends from the beginning of the currently fixated word to 2 words to the right of fixation for readers of English (Rayner, Well, Pollatsek, & Bertera, 1982). As long as the letters of the fixated word and the following word are available and the rest of the letters are replaced with visually similar letters (and the spaces between words are left intact), readers are typically unaware that anything is strange about the text and reading is only about 10% slower than without a window (Rayner et al., 1982). If 3 words (the currently fixated and the next 2 words to the right) are available, reading is generally equivalent to normal (Rayner et al., 1982). As mentioned earlier, the reading process is influenced by properties of the writing system; the asymmetry of the perceptual span is reversed for readers of Hebrew, which is printed right-to-left (Pollatsek et al., 1981), and is smaller for readers of Hebrew than readers of English. Similarly, the perceptual span is smaller in Chinese (1 character to the left and 3 characters to the right of fixation; Inhoff & Liu, 1998) and Japanese (Ikeda & Saida, 1978). However, as with saccade length described, the spans are equivalent when the number of words is considered instead of the number of characters.

The size of the perceptual span can vary from fixation to fixation. As we will discuss, the perceptual span becomes smaller as the difficulty of the fixated word increases (Henderson & Ferreira, 1990; Kennison & Clifton, 1995). Other interesting findings are that (1) beginning (Häikiö, Bertram, Hyönä, & Niemi, 2009; Rayner, 1986) and dyslexic readers (Rayner, Murphy, Henderson, & Pollatsek, 1989) have smaller spans then skilled readers; (2) fast readers (~ 330 wpm) have larger perceptual spans than slower readers (~ 200 wpm; Rayner, Slattery, & Bélanger, 2010); (3) older readers (mean age over 70 years) have a smaller and less asymmetric perceptual spans than college-aged readers (Rayner, Castelhano, & Yang, 2009); and (4) skilled deaf readers have larger perceptual spans than hearing controls (Bélanger, Slattery, Mayberry, & Rayner, 2012).

Now we consider the results from moving mask experiments (Rayner & Bertera, 1979; see also Rayner, Yang, Schuett, & Slattery, 2013; Figure 4.1, panel B). When the mask was small enough to allow some information to reach the fovea (i.e., if it was only 1 to 5 letters wide, with 3 letters equaling one degree of visual angle), subjects read at a reduced rate but were still able to obtain information beyond the mask. As the size of the mask increased, reading efficiency dropped precipitously. When the mask was extremely wide (13–17 letters) subjects were able to report very little information about the sentence; in the intermediate conditions (7–11 letters), where the mask covered the entirety of the fovea but only some of the parafovea, readers made a large number of errors when reporting the sentences. The nature of these errors (e.g., reading *pretty* as *priest* or *profits* as *politics*) indicate that readers were struggling to guess the correct word based on mostly low-level features such as word-initial letters and word length. Even though the manipulation is quite distinct from natural reading, this study clearly demonstrates the limits of parafoveal processing in reading and, consequently, the importance of foveal processing (see also the later descriptions of disappearing text studies).

What Type of Information Is Obtained From Parafoveal Words?

Although it is clear that parafoveal information is limited compared with foveal information, preview benefit effects demonstrate the importance of parafoveal information to reading. Obviously, these effects require that some parafoveal information obtained on a previous fixation persists during the saccade and is available during the subsequent fixation. We will review the evidence for preview benefit separately at each level of representation of words (for more thorough reviews, see Cutter, Drieghe, & Liversedge, this volume; Reingold, Sheridan, & Reichle, this volume; Schotter et al., 2012).

ORTHOGRAPHIC PROCESSING

It has been robustly demonstrated that orthographic information is obtained from words parafoveally and yields preview benefit when it is shared between preview and target (Balota, Pollatsek, & Rayner, 1985; Rayner, 1975). Importantly, this information is based on abstract letter codes and does not depend on the overall shape of the word; the results do not change when presenting words in alternating case (e.g., AlTeRnAtInG cAsE; McConkie & Zola, 1979; Rayner et al., 1980; Slattery, Angele, & Rayner, 2011). Furthermore, having the first 2 to 3 letters preserved in the preview facilitates processing of the target word (Inhoff, Pollatsek, Posner, & Rayner, 1989). Having the letters at the end of a word preserved does not as consistently yield preview benefit except for words shorter than 6 letters (Johnson, Perea, & Rayner, 2007). Lastly, Johnson et al. (2007) found that readers obtained more benefit from transposed-letter previews (*jugde* as a preview for *judge*) than from previews with replacement letters (*jupbe*). The same was true for longer (7-letter) targets except when the first or the last letters of the preview were transposed. The transposed letters did not have to be adjacent; readers obtained more preview benefit from *flewor* than *flawur* for the target word *flower*. Furthermore, transposed letter preview benefits were obtained from either transpositions that produce words (*clam-calm*) or nonwords (*clam-caml*; Johnson & Dunne, 2012), suggesting that these effects operate at the orthographic rather than at the lexical level.

PHONOLOGICAL PROCESSING

Phonological preview benefit effects show that readers can use phonological information about a parafoveal word to help guide processing when the word is subsequently fixated (Ashby & Rayner, 2004; Rayner, Sereno, Lesch, & Pollatsek, 1995; Sparrow & Miellet, 2002). Specifically, a parafoveal preview of a phonologically related (homophone) word facilitates processing of the target for readers of English (Pollatsek, Lesch, Morris, & Rayner, 1992; cf. Chace, Rayner, & Well, 2005, for a lack of facilitation for less skilled readers). Additionally, there is preview benefit from homophone and pseudohomophone previews, demonstrated in French (Miellet & Sparrow, 2004), English (Ashby, Treiman, Kessler, & Rayner, 2006) and Chinese (Liu, Inhoff, Ye, & Wu, 2002; Pollatsek, Tan, & Rayner, 2000; Tsai, Lee, Tzeng, Hung, & Yen, 2004), which is not an alphabetic language and therefore does not always code phonology as transparently through orthography (see Pollatsek, this volume).

In some situations, typographical information obtained parafoveally can be used to guide phonological processing upon fixation. Slattery, Schotter, Berry, and Rayner (2011) conducted a boundary experiment with abbreviations as target words that were presented in normal (mostly lowercase) sentences or all capital sentences. They manipulated whether the target abbreviation (which was always printed in all capitals) was an *acronym* (i.e., pronounced as a word such as *NASA*) or an *initialism* (i.e., pronounced as a series of letter names such as *NCAA*). They found that, readers were biased to process these abbreviations as initialisms when they were presented in mostly lower case sentences, so that they were typographically distinct. On the other hand, when the abbreviations were presented in all capital sentences and not typographically distinct, readers defaulted to processing these strings as words, indicating that, in some ways, typographical information obtained parafoveally can bias phonological processing once a word is fixated.

MORPHOLOGICAL PROCESSING

There is mixed evidence for preview benefit of morphological information in reading in alphabetic languages. While there is no evidence for parafoveal morphological processing in English (e.g., Drieghe, Pollatsek, Juhasz, & Rayner, 2010; Lima, 1987) or Finnish (e.g., Hyönä, Bertram, & Pollatsek, 2004), there is evidence that morphological information is processed parafoveally in Hebrew (e.g., Deutsch, Frost, Pollatsek, & Rayner, 2000). The discrepancy between the Hebrew and English/Finnish studies may reflect differences in the morphological structure of the languages. In Hebrew, all verbs and many nouns and adjectives are marked semantically (by a *word root*, generally consisting of three consonants) and morphologically (by a *word pattern*, consisting of vowels or a mixture of vowels and consonants). The word root and the word pattern are interleaved instead of concatenated, so it is not the case that the word root is the beginning of the word and the word pattern is the end, or vice versa. Furthermore, in this interleaved structure the positions of the constituent letters of the word root or the word pattern are not fixed, so the orthographic or phonological structure of the word does not transparently indicate morphology. For this reason, any preview benefit in Hebrew provided in the morphologically related preview condition above and beyond the orthographically related preview condition is due to morphology being processed parafoveally and not due to a stronger orthographic relationship between the morphological preview and the target.

Chinese may be similar to Hebrew in that morphological structure plays a more important role in word identification than in English or Finnish. In Chinese, the morpheme that a character represents can differ depending on the word in which it is embedded. Yen et al. (2008) found that preview benefits were larger when the preview and target word shared a character representing the same morpheme than when they shared a character that represented different morphemes in the context of the different words. Recently, Yang (2013) reported a preview benefit for 2-character (i.e., bimorphemic) Chinese compound words. Readers obtained the same amount of preview benefit from a reverse character order preview as from the identical/correct character order preview, as long as the transposition fit into the sentence context. Similarly, Angele and Rayner (2013) found that in English, readers obtained a small preview benefit from a word in which the order of the morphemes was transposed (*boycow* as a preview for *cowboy*). Furthermore, Rayner, Angele, Schotter, and Bicknell (2013) found that there is no preview benefit for transpositions of two words separated by a space (e.g., "My neighbor painted the *white walls/walls white* yesterday."), suggesting that while morphemes within a word might be processed in parallel to some extent, this is not the case for two separate words.

Control of Eye Movements During Reading

There are two important questions regarding eye movement control in reading. First, what determines where to look next? The second, and more frequently investigated, issue is what determines when to move the eyes. Rayner and Pollatsek (1981) provided the first unambiguous evidence that the length of the next saccade and the duration of the current fixation are directly affected by ongoing cognitive processing. In one experiment they used the moving window paradigm, and the size of the window of normal text was randomly varied from fixation to fixation. They found that saccade length of the next saccade varied according to the size of the window: if the window was small the saccade was short, whereas if the window was large the saccade was much larger. These data suggest that if readers have access to more information parafoveally, they may not need to directly fixate those locations on the next fixation and can therefore make a longer saccade. In another experiment, the onset of access to information from the foveal text was delayed by the presence of a mask (with time varying randomly from fixation to fixation), and fixation durations were lengthened by approximately the duration of the mask onset (see also Morrison, 1984). These data suggest that the purpose of fixations is to allow visual access to the text. If this information is denied (e.g., by the presence of the mask), the eyes wait in that location until the useful information is provided, and then continue normally. Furthermore, these masking manipulations affected saccade length and fixation duration independently; thus there is reason to believe that the decisions of where to move the eyes and when to move the eyes are made independently (Rayner & McConkie, 1976). While some researchers have suggested that the two decisions overlap in time and influence each other, for the sake of exposition we will discuss these two topics separately.

Where to Move the Eyes Next

The strongest influence on where the eyes move is low-level cues provided by word length and space information; this is less true for unspaced languages like Chinese and Thai (how readers of unspaced texts segment words is a major challenge for understanding reading in these languages; see Li, Rayner, & Cave, 2009; see Li, Zang, Liversedge, & Pollatsek, this volume). Saccade length is strongly related to the length of the fixated word and the subsequent word (e.g., O'Regan, 1980; Rayner, 1979); readers make longer saccades into long words than into short words. For languages that are typically written with spaces between words, reading slows down by as much as 30% to 50% or more when the spaces are removed (Morris, Rayner, & Pollatsek, 1990; Rayner, Fischer, & Pollatsek, 1998; Rayner, Yang, et al., 2013). Interestingly, Kohsom and Gobet (1997) demonstrated that when space information was provided for readers of Thai (who are not used to reading with spaces between words), they read more effectively than normal; however, this was not true for readers of Chinese (Bai, Yan, Liversedge, Zang, & Rayner, 2008) or Japanese (Sainio, Hyönä, Bingushi, & Bertram, 2007). Whereas inserting spaces between characters interfered with Chinese reading, inserting spaces between words did not. Actually, it is quite surprising that the insertion of spaces between words did not interfere, given that the Chinese readers have a lifetime of experience reading without spaces. All of these pieces of evidence suggest that interword spaces benefit reading as long as they are not orthographically illegal.

Landing Position Effects

The spaces between words provide information about an upcoming word's length, which leads to systematic tendencies with respect to where the eyes typically land. Rayner (1979) demonstrated that readers' eyes tend to land halfway between the middle of a word and the beginning of that word, the *preferred viewing location*. It is generally argued that readers attempt to target the center of words, but their saccades tend to fall short (McConkie, Kerr, Reddix, & Zola, 1988; Rayner, 1979). When readers' eyes land at a nonoptimal position in a word, they are more likely to refixate that word (O'Regan, 1990; Rayner, Sereno, & Raney, 1996). Using the boundary paradigm, Inhoff et al. (2003; see also Juhasz, White, Liversedge, & Rayner 2008; White, Rayner, & Liversedge, 2005) provided readers with an incorrect length preview of an upcoming word in the parafovea, causing readers to send their eyes to what will turn out to be a nonoptimal position in the word. They found that this increases reading time on the word once it is fixated. The location of a fixation in a word can be viewed not only as a landing site for that word but also as the launch site for the next saccade. Although the preferred viewing location in a word lies between the beginning and the middle of a word, this position varies as a function of the prior launch site (McConkie et al., 1988; Rayner et al., 1996). If the launch site for a saccade landing on a target word is far from that word (e.g., 8–10 letter spaces), the landing position will be shifted to the left. Likewise, if the distance is small (2–3 letter spaces), the landing position is shifted to the right.

In contrast to the preferred viewing location, which represents where readers tend to fixate words, the optimal viewing position represents the location in a word at which recognition time is minimized (i.e., efficiency is maximized). The optimal viewing location effect was originally studied in the context of isolated word recognition studies, in which eye movements were monitored (O'Regan & Jacobs, 1992; O'Regan, Lévy-Schoen, Pynte, & Brugaillère, 1984), and two general effects have been reported. First, there is a refixation effect such that the further the eyes are from the optimal viewing position, the more likely it is that a refixation will be made on the word. Second, there is a processing cost effect such that for every letter that the eyes deviate from the optimal viewing position, there is a cost of roughly 20 ms (O'Regan et al., 1984). Interestingly, however, although the refixation effect remains in reading (as opposed to isolated word recognition), the processing cost is either greatly attenuated or absent (Rayner et al., 1996; Vitu, O'Regan, & Mittau, 1990). There are two reasons for this: Contextual information in reading may override low-level visual processing, and the parafoveal preview of the word before it is directly fixated facilitates subsequent processing.

Skipping Effects

As noted earlier, words are sometimes skipped during reading. Obviously, skipped words must usually be processed in parafoveal vision (where stimuli are degraded by acuity limitations), which also reduces the speed of processing of these words (Rayner & Morrison, 1981). It is a mistake to think that if a word is skipped it is not processed. Fisher and Shebilske (1985) demonstrated this by examining the eye movements of readers on a passage of text. They then deleted all words from the passage that these readers had skipped and asked a second group of readers to read the passage. This second group had a difficult time understanding the text. So skipped words do get processed on the fixation prior to or after the skip (though some words may be guessed by readers). Thus the fixation prior to skipping is inflated compared with fixations prior to a word that is not skipped (Kliegl & Engbert, 2005; Pollatsek, Rayner, & Balota, 1986), and the same is true for fixations after skipping (Reichle, Rayner, & Pollatsek, 2003).

Two factors have a big impact on skipping: word length (short words are much more likely to be skipped than long words; see Drieghe, Brysbaert, Desmet, & De Baecke, 2004) and contextual constraint (predictable words are much more likely to be skipped than unpredictable words; e.g., Ehrlich & Rayner, 1981; Rayner & Well, 1996). When two or three short words occur in succession, there is a good chance that two of them will be skipped. And short words (like *the*) preceding a content word are often skipped (Drieghe, Pollatsek, Staub, & Rayner, 2008; Gautier, O'Regan, & LaGargasson, 2000). Angele and Rayner (2013; see also Angele, Laishly, Rayner, & Liversedge, 2014) demonstrated that *the* is skipped frequently even when it is in a position in the sentence where it would not be licensed (e.g., grammatically) by the prior context. Word frequency also has an effect on word skipping, but the effect is smaller than that of predictability (Rayner et al., 1996). While predictability influences whether or not a word is skipped, it does not influence where in the word the fixation lands (Rayner, Binder, Ashby, & Pollatsek, 2001; Vainio, Hyönä, & Pajunen, 2009), though it does influence fixation durations.

When to Move the Eyes

It is clear that the difficulty associated with processing the fixated word is strongly related to the duration of fixations on it (i.e., the decision of when to move the eyes off of the word) and this difficulty is mostly determined by a host of linguistic variables (see Liversedge & Findlay, 2000; Rayner, 1998, 2009, for a more complete discussion). These linguistic variables include word frequency (Inhoff & Rayner, 1986; Kliegl, Grabner, Rolfs, & Engbert, 2004; Rayner & Duffy, 1986; Schilling, Rayner, & Chumbley, 1998), word predictability (Balota et al., 1985; Ehrlich & Rayner, 1981; Rayner & Well, 1996; Zola, 1984), number of meanings (Duffy, Morris, & Rayner, 1988; Folk & Morris, 2003; Leinenger & Rayner, 2013; Rayner, Cook, Juhasz, & Frazier, 2006; Sereno, O'Donnell, & Rayner, 2006), age of acquisition (Juhasz & Rayner, 2003, 2006), phonological properties of words (Ashby & Clifton, 2005; Folk, 1999; Jared, Levy, & Rayner, 1999; Rayner, Pollatsek, & Binder, 1998; Sereno & Rayner, 2000), semantic relations between the fixated word and prior words (Carroll & Slowiaczek, 1986; Morris, 1994), and word familiarity (Chaffin, Morris, & Seely, 2001). In addition, the variable of neighborhood frequency (i.e., the number of words that can be created by replacing a single letter in its position from the target word) generally exerts effects not during the initial encounter with the word (Perea & Pollatsek, 1998), but rather in later reading-time measures (like the probability of rereading the word after moving past it or time spent on words following the target). Moreover, reading time is influenced by the letters in the word in that even if the words contain all the correct letters but in the wrong order, reading time is impaired compared with reading words with correctly ordered letters. Furthermore, these findings are not confined to English (the language most often examined); there are similar effects of word frequency (Yan et al., 2006) and word predictability (Rayner, Li, Juhasz, & Yan, 2005) when reading Chinese.

Interestingly, while these effects are quite robust during reading for comprehension, the magnitude of these effects varies depending on the task in which the person engages (e.g., Kuperman, Drieghe, Keuleers, & Brysbaert, 2013; Schilling et al., 1998). For example, the word frequency effect disappears or is attenuated (compared with normal reading) when searching for a particular word in the text (Rayner & Fischer, 1996; Rayner & Raney, 1996) and when readers "zone out" during reading (Reichle, Rennenberg, & Schooler, 2010; Schad & Engbert, 2012). In contrast, in tasks that require that readers engage more deeply with the words (e.g., when proofreading to detect spelling errors) frequency effects always become exaggerated (Kaakinen & Hyönä, 2010; Schotter, Bicknell, Howard, Levy, &

Rayner, 2014), and predictability effects become exaggerated only when a comparison between the word and the sentence context is necessary to detect errors (e.g., when the spelling errors produce real words that are incompatible with the sentence context, as in "The marathon runners trained on the trial behind the high school" as opposed to "… trained on a trcak"; Schotter et al., 2014).

It is thus quite clear that lexical variables have strong and immediate effects on how long readers look at a word. While other linguistic variables can have an influence on how soon readers move on in the text, it is generally the case that higher-level linguistic variables have somewhat later effects, unless the variable more or less smacks you in the eye. So for example, when readers fixate on the disambiguating word in a syntactic garden path sentence (e.g., "While Mary bathed the baby spat up on the bed") there are increased fixation times on it (Frazier & Rayner, 1982; Rayner & Frazier, 1987) or regressions from it back to earlier parts of the sentence (Frazier & Rayner, 1982; Meseguer, Carreiras, & Clifton, 2002). Readers also have longer fixations at the end of clauses and sentences (Hirotani, Frazier, & Rayner, 2006; Just & Carpenter, 1980; Rayner, Kambe, & Duffy, 2000; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). In addition, when readers encounter an anomalous word, they fixate on it longer, and the effect is quite immediate (Rayner, Warren, Juhasz, & Liversedge, 2004; Staub, Rayner, Pollatsek, Hyönä, & Majewski, 2007; Warren & McConnell, 2007). However, when a word indicates a merely implausible but not truly anomalous event, the effect registered in the eye movement record is typically delayed a bit, showing up in later processing measures (Joseph et al., 2009; Rayner et al., 2004).

Interestingly, when sentences with an anomalous word (such as *carrot* in "Jane used a pump to inflate the large carrot") are embedded in cartoon or fantasy-like contexts and are compared with real-world contexts where inflating a carrot with a pump is anomalous (Warren, McConnell, & Rayner, 2008), the earliest effects on fixation time measures (first-fixation duration and gaze duration) still were on the anomalous word rather than on the control word, *carrot*, in a sentence like "Jane used a knife to chop the large carrot." However, the reading time measure that includes time reading *carrot* and rereading any previous words before moving on (i.e., go-past time) revealed disruption only in the real-world context. These results suggest that contextual information did not eliminate the initial disruption, but moderated it quickly thereafter. In short, some variables have strong influences immediately when a word is fixated (such as frequency, age of acquisition, predictability, discussed previously), while other variables mostly seem to have later occurring effects. However, there is no doubt that cognitive processing activities have a strong influence on when the eyes move.

Disappearing Text Experiments

Perhaps the most compelling evidence that cognitive processing of the fixated word drives the eyes through the text comes from disappearing text experiments, in which the fixated word either disappears or is masked 50 to 60 ms after it is fixated (Ishida & Ikeda, 1989; Liversedge et al., 2004; Rayner et al., 1981; Rayner, Liversedge, & White, 2006; Rayner, Liversedge, White, & Vergilino-Perez, 2003). These studies show that if readers are allowed to see the fixated word for 50 to 60 ms before it disappears, they read quite normally. Furthermore, there was still a word frequency effect (i.e., longer fixations on low- than high-frequency words) when the fixated word disappeared. This result provides very good evidence that lexical processing is the engine that drives the eyes through text (see also Staub, White, Drieghe, Holloway, & Rayner, 2010). These findings do not mean that readers are able to fully process and identify the fixated word in 50 to 60 ms, but rather that 50 to 60 ms is the time that is needed to get the visual information into the processing system (i.e., the eye-brain lag; Reichle & Reingold, 2013). It further suggests that readers fixate words much longer than necessary to obtain sufficient visual information to identify them. But under normal circumstances, the extra time is not wasted, since readers can begin to process the upcoming word parafoveally.

Despite the fact that moving mask and disappearing text experiments (described previously) suggest that foveal information is most important for reading, parafoveal information is very important as well. Rayner et al. (2006; see also Inhoff, Eiter, & Radach, 2005) found that readers' performance declined drastically when, on each fixation, the word to the right of fixation disappeared or was masked 60 ms after fixation onset. This is in marked contrast to findings that readers are able to read sentences fairly normally when the fixated word disappeared or was masked 60 ms after fixation onset (Liversedge et al., 2004; Rayner et al., 1981; Rayner et al., 2003; Rayner, Yang, Castelhano, & Liversedge, 2011).

In short, these studies contribute strong evidence for *direct control* of eye movements by cognitive processing; they suggest that the decisions of when to move the eyes reflect the difficulty of ongoing cognitive processing. But these are not the only studies to support this idea (see Dambacher, Slattery, Yang, Kliegl, & Rayner, 2013; Reingold & Rayner, 2006; Reingold, Yang, & Rayner, 2010; Schad, Risse, Slattery, & Rayner, 2014). Other strong evidence comes from sophisticated statistical analyses of the durations of fixations on words (see Reingold, et al., this volume).

Survival Analyses

It has been suggested, despite the robust evidence for the influence of lexical factors on eye-movement decisions, that eye movements are primarily determined by low-level factors (e.g., physical size of letters, clarity of letters) and that lexical factors only have an influence on extremely long fixations (e.g., Feng, 2006; McConkie & Yang, 2003; Yang, 2006; Yang & McConkie, 2001). To test this, Reingold, Reichle, Glaholt, and Sheridan (2012) developed a technique that is based on survival analyses whereby they manipulated the frequency of target words (high vs. low) as well as their availability for parafoveal processing during fixations on the pretarget word such that there was either a valid or invalid preview. They investigated the distributions of fixation durations to assess whether these manipulations affected the entire distribution (suggesting direct control) or whether the influence was only observed for the extremely long fixations. They used a survival analysis technique that provided precise estimates of the timing of the first discernible influence of word frequency on the fixation on the target word. Using this technique, they found a significant influence of word frequency on fixation duration in normal reading (valid preview) as early as 145 ms from the start of fixation. The time course of frequency effects was strongly influenced by preview validity, with the frequency effect being apparent much later in processing when there was not a valid preview. These results demonstrate the crucial role of parafoveal processing in enabling direct lexical control of fixation times. These survival analyses have also been used to study the time course of word predictability (Sheridan & Reingold, 2012a), lexical ambiguity resolution (Sheridan & Reingold, 2012b), and text segmentation (Sheridan, Rayner, & Reingold, 2013; see Reingold et al., this volume).

Final Comments

In this chapter we have reviewed the basic information concerning the work of the eyes during reading. Much of the research in the field over the past few years has been dominated by the development and appearance of a number of computational models of reading with the most prominent of these models being the E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998), SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005), and Glenmore (Reilly & Radach, 2006). Given that we are intimately involved with the original development and later instantiations of E-Z Reader (Schotter, Reichle, & Rayner, 2014), we obviously see the advantages of computational models in terms of accounting for data and making interesting and informative predictions. But we also feel that the pendulum has swung a bit too far and that cleverly designed experimental studies to further investigate the relationship between eye movements and reading should be heavily valued and at least on par with research designed to test differences between the models.

References

Angele, B., Laishley, A. E., Rayner, K., & Liversedge, S. P. (2014). The effect of high-and low-frequency previews and sentential fit on word skipping during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1181–1203.

Angele, B., & Rayner, K. (2013). Eye movements and parafoveal preview of compound words: Does morpheme order matter? *Quarterly Journal of Experimental Psychology, 66,* 505–526.

Ashby, J., & Clifton, C., Jr. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition, 96,* B89–B100.

Ashby, J., & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements. *Language and Cognitive Processes, 19,* 391–426.

Ashby, J., Treiman, R., Kessler, B., & Rayner, K. (2006). Vowel processing during silent reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 416–424.

Ashby, J., Yang, J., Evans, K. H., & Rayner, K. (2012). Eye movements and the perceptual span in silent and oral reading. *Attention, Perception,* and *Psychophysics, 74,* 634–640.

Bai, X., Yan, G., Zang, C., Liversedge, S. P., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1277–1287.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17,* 364–390.

Bélanger, N. N., Slattery, T. J., Mayberry, R. I., & Rayner, K. (2012). Skilled deaf readers have an enhanced perceptual span in reading. *Psychological Science, 23,* 816–823.

Buswell, G. T. (1922). *Fundamental reading habits: A study of their development*. Chicago, IL: Chicago University Press.

Carroll, P., & Slowiaczek, M. L. (1986). Constraints on semantic priming in reading: A fixation time analysis. *Memory & Cognition, 14,* 509–522.

Chace, K. H., Rayner, K., & Well, A. D. (2005). Eye movements and phonological parafoveal preview: Effects of reading skill. *Canadian Journal of Experimental Psychology, 59,* 209–217.

Chaffin, R., Morris, R. K., & Seely, R. E. (2001). Learning new word meanings from context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 225.

Dambacher, M., Slattery, T. J., Yang, J., Kliegl, R., & Rayner, K. (2013). Evidence for direct control of eye movements during reading. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1468.

Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2000). Early morphological effects in word recognition in Hebrew: Evidence from parafoveal preview benefit. *Language and Cognitive Processes, 15,* 487–506.

Drieghe, D., Brysbaert, M., Desmet, T., & De Baecke, C. (2004). Word skipping in reading: On the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology, 16,* 79–103.

Drieghe, D., Pollatsek, A., Juhasz, B. J., & Rayner, K. (2010). Parafoveal processing during reading is reduced across a morphological boundary. *Cognition, 116,* 136–142.

Drieghe, D., Pollatsek, A., Staub, A., & Rayner, K. (2008). The word grouping hypothesis and eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1552–1560.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27,* 429–446.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20*, 641–655.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112,* 777–813.

Feng, G. (2006). Eye movements as time-series random variables: A stochastic model of eye movement control in reading. *Cognitive Systems Research, 7*, 70–95.

Fine, E. M., & Rubin, G. S. (1999). The effects of simulated cataract on reading with normal vision and simulated central scotoma. *Vision Research, 39,* 4274–4285.

Fisher, D. F., & Shebilske, W. L. (1985). There is more that meets the eye than the eyemind assumption. In R. Groner, G. W. McConkie, & C. Menz (Eds.), *Eye movements and human information processing* (pp. 149–158). Amsterdam, the Netherlands: North Holland.

Folk, J. R. (1999). Phonological codes are used to access the lexicon during silent reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 892–906.

Folk, J. R., & Morris, R. K. (2003). Effects of syntactic category assignment on lexical ambiguity resolution in reading: An eye movement analysis. *Memory & Cognition, 31,* 87–99.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14,* 178–210.

Gautier, V., O'Regan, J. K., & Le Gargasson, J. F. (2000). "The-skipping" revisited in French: Programming saccades to skip the article 'les'. *Vision Research, 40,* 2517–2531.

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology, 102,* 167–181.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 417–429.

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language, 54,* 425–443.

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement–contingent display change study. *Memory & Cognition, 32*, 523–532.

Ikeda, M., & Saida, S. (1978). Span of recognition in reading. *Vision Research, 18,* 83–88.

Inhoff, A. W., Eiter, B. M., & Radach, R. (2005). Time course of linguistic information extraction from consecutive words during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 979–995.

Inhoff, A. W., & Liu, W. (1998). The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 20–34.

Inhoff, A. W., Pollatsek, A., Posner, M. I., & Rayner, K. (1989). Covert attention and eye movements during reading. *Quarterly Journal of Experimental Psychology, 41A,* 63–89.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception* and *Psychophysics, 40,* 431–439.

Inhoff, A. W., Solomon, M., Radach, R., & Seymour B. A. (2011). Temporal dynamics of the eye–voice span and eye movement control during oral reading. *Journal of Cognitive Psychology, 23,* 543–558.

Ishida, T., & Ikeda, M. (1989). Temporal properties of information extraction in reading studied by a text-mask replacement technique. *Journal of the Optical Society of America A, 6*, 1624–1632.

Jared, D., Levy, B. A., & Rayner, K. (1999). The role of phonology in the activation of word meanings during reading: evidence from proofreading and eye movements. *Journal of Experimental Psychology: General, 128,* 219–264.

Johnson, R. L., & Dunne, M. D. (2012). Parafoveal processing of transposed-letter words and nonwords: Evidence against parafoveal lexical activation. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 191–212.

Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 209–229.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research, 49,* 2078–2086.

Juhasz, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 1312–1318.

Juhasz, B. J., & Rayner, K. (2006). The role of age of acquisition and word frequency in reading: Evidence from eye fixation durations. *Visual Cognition, 13*, 846–863.

Juhasz, B. J., White, S. J., Liversedge, S. P., & Rayner, K. (2008). Eye movements and the use of parafoveal word length

information in reading. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1560–1579.
Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329–354.
Kaakinen, J., & Hyönä, J. (2010). Task effects on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1561–1566.
Kennison, S. M., & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 68–81.
Kirkby, J. A., Webster, L. A. D., Blythe, H. I., & Liversedge, S. P. (2008). Binocular coordination during reading and non-reading tasks. *Psychological Bulletin, 134,* 742–763.
Kliegl, R., & Engbert, R. (2005). Fixation durations before word skipping in reading. *Psychonomic Bulletin* and *Review, 12,* 132–138.
Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16,* 262–284.
Kohsom, C., & Gobet, F. (1997). Adding spaces to Thai and English: Effects on reading. In *Proceedings of the nineteenth annual conference of the Cognitive Science Society* (Vol. 19, pp. 388–393). Ablex Press.
Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora megastudies. *Quarterly Journal of Experimental Psychology, 66,* 563–580.
Leinenger, M., & Rayner, K. (2013). Eye movements while reading biased homographs: Effects of prior encounter and biasing context on reducing the subordinate bias effect. *Journal of Cognitive Psychology, 25,* 665–681.
Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology, 58,* 525–552.
Lima, S. D. (1987). Morphological analysis in sentence reading. *Journal of Memory and Language, 26,* 84–99.
Liu, W., Inhoff, A. W., Ye, Y., & Wu, C. (2002). Use of parafoveally visible characters during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1213–1227.
Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences,* 4, 6–14.
Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., & Kentridge, R. W. (2004). Eye movements when reading disappearing text: Is there a gap effect in reading? *Vision Research, 44,* 1013–1024.
Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin, 81,* 899–917.
McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations on words. *Vision Research, 28,* 1107–1118.
McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception* and *Psychophysics, 17,* 578–587.
McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society, 8*, 365–368.
McConkie, G. W., & Yang, S. N. (2003). How cognition affects eye movements during reading. In R. Radach, J. Hyönä, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 413–427). Amsterdam, the Netherlands: Elsevier.
McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception* and *Psychophysics, 25,* 221–224.
Meseguer, E., Carreiras, M., & Clifton, C. (2002). Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Memory & Cognition, 30,* 551–561.
Miellet, S., O'Donnell, P. J., & Sereno, S. C. (2009). Parafoveal magnification: Visual acuity does not modulate the perceptual span in reading. *Psychological Science, 20*, 721–728.
Miellet, S., & Sparrow, L. (2004). Phonological codes are assembled before word fixation: Evidence from boundary paradigm in sentence reading. *Brain and Language, 90,* 299–310.
Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 92–103.
Morris, R. K., Rayner, K., & Pollatsek, A. (1990). Eye movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 268–281.
Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 667–682.
Morrison, R. E., & Rayner, K. (1981). Saccade size in reading depends upon character spaces and not visual angle. *Attention, Perception,* and *Psychophysics, 30,* 395–396.
O'Regan, J. K. (1980). The control of saccade size and fixation duration in reading: The limits of linguistic control. *Attention, Perception,* and *Psychophysics, 28,* 112–117.
O'Regan, J. K. (1983). Elementary perceptual and eye movement control processes in reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 121–140). San Diego, CA: Academic Press.
O'Regan, J. K. (1990). Eye movements and reading. In E. Kowler (Ed.), *Eye Movements and their role in visual and cognitive processes* (pp. 395–453). Amsterdam, the Netherlands: Elsevier.
O'Regan, J. K., & Jacobs, A. M. (1992). Optimal viewing position effect in word recognition: A challenge to current theory. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 185–197.
O'Regan, J. K., Lévy-Schoen, A., Pynte, J., & Brugaillère, B. E. (1984). Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 250–257.
Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 767–779.
Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language, 14,* 174–180.
Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of*

Experimental Psychology: Human Perception and Performance, 18, 148–162.
Pollatsek, A., Rayner, K., & Balota, D. A. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception & Psychophysics, 40,* 123–130.
Pollatsek, A., Tan, L. H., & Rayner, K. (2000). The role of phonological codes in integrating information across saccadic eye movements in Chinese character identification. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 607–633.
Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology,* 7, 65–81.
Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8,* 21–30.
Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41,* 211–236.
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422.
Rayner, K. (2009). The thirty fifth Sir Frederick Bartlett lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62,* 1457–1506.
Rayner, K. (2014). The gaze-contingent moving window in reading: Development and review. *Visual Cognition, 22,* 242–258.
Rayner, K., Angele, B, Schotter, E. R., & Bicknell, K. (2013). On the processing of canonical word order during eye fixations in reading: Do readers process transposed word previews? *Visual Cognition, 21,* 353–381.
Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science, 206,* 468–469.
Rayner, K., Binder, K. S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research, 41,* 943–954.
Rayner, K., Castelhano, M. S., & Yang, J. (2009). Eye movements and the perceptual span in older and younger readers. *Psychology and Aging, 24,* 755–760.
Rayner, K., Cook, A. E., Juhasz, B. J., & Frazier, L. (2006). Immediate disambiguation of lexically ambiguous words during reading: Evidence from eye movements. *British Journal of Psychology, 97,* 467–482.
Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14,* 191–201.
Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics, 5,* 734–747.
Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38,* 1129–1144.
Rayner, K., & Frazier, L. (1987). Parsing temporarily ambiguous complements. *Quarterly Journal of Experimental Psychology, 39,* 657–673.
Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 167–179.
Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology, 53,* 1061–1080.
Rayner, K., Li, X., Juhasz, B. J., & Yan, G. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review, 12,* 1089–1093.
Rayner, K., Li, X., Williams, C. C., Cave, K. R., & Well, A. D. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Research, 47,* 2714–2726.
Rayner, K., Liversedge, S. P., & White, S. J. (2006). Eye movements when reading disappearing text: The importance of the word to the right of fixation. *Vision Research, 46,* 310–323.
Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading disappearing text. *Psychological Science, 14,* 385–388.
Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research, 16,* 829–837.
Rayner, K., & Morrison, R. E. (1981). Eye movements and identifying words in parafoveal vision. *Bulletin of the Psychonomic Society, 17,* 135–138.
Rayner, K., Murphy, L., Henderson, J. M., & Pollatsek, A. (1989). Selective attentional dyslexia. *Cognitive Neuropsychology, 6,* 357–378.
Rayner, K., & Pollatsek, A. (1981). Eye movement control during reading: Evidence for direct control. *Quarterly Journal of Experimental Psychology, 33,* 351–373.
Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.
Rayner, K., Pollatsek, A., & Binder, K. S. (1998). Phonological codes and eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 476–497.
Rayner, K., & Raney, G. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review, 3,* 245–248.
Rayner, K., Sereno, S. C., Lesch, M. F., & Pollatsek, A. (1995). Phonological codes are automatically activated during reading: Evidence from an eye movement priming paradigm. *Psychological Science, 6,* 26–32.
Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C., Jr (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes, 4,* SI21–SI49.
Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1188–1200.
Rayner, K., Slattery, T. J., & Bélanger, N. N. (2010). Eye movements, the perceptual span, and reading speed. *Psychonomic Bulletin & Review, 17,* 834–839.
Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 1290–1301.
Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review,* 3, 504–509.
Rayner, K., Well, A. D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics, 27,* 537–544.
Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception & Psychophysics, 31,* 537–550.

Rayner, K., Yang, J., Castelhano, M. S., & Liversedge, S. P. (2011). Eye movements of older and younger readers when reading disappearing text. *Psychology and Aging, 26,* 214–223.

Rayner, K., Yang, J., Schuett, S., & Slattery, T. J. (2013). Eye movements of older and younger readers when reading unspaced text. *Experimental Psychology, 60,* 354–361.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26,* 445–476.

Reichle, E. D., & Reingold, E. M. (2013). Neurophysiological constraints on the eye-mind link. *Frontiers in Human Neuroscience, 7,* 361.

Reichle, E. D., Rennenberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science, 21,* 1300–1310.

Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research, 7,* 34–55.

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z reader model. *Psychological Science, 17,* 742–746.

Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive Psychology, 65,* 177–206.

Reingold, E. M., Yang, J., & Rayner, K. (2010). The time course of word frequency and case alternation effects on fixation times in reading: Evidence for lexical control of eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1677–1683.

Sainio, M., Hyönä, J., Bingushi, K., & Bertram, R. (2007). The role of interword spacing in reading Japanese: An eye movement study. *Vision Research, 47,* 2575–2584.

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulated shuffled versus normal text reading using the SWIFT model. *Visual Cognition, 20,* 391–421.

Schad, D. J., Risse, S., Slattery, T., & Rayner, K. (2014). Word frequency in fast priming: Evidence for immediate cognitive control of eye movements during reading. *Visual Cognition, 22,* 390–414.

Schilling, H., Rayner, K., & Chumbley, J. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26,* 1270–1281.

Schotter, E. R, Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception & Psychophysics, 74,* 5–35.

Schotter, E. R., Bicknell, K., Howard, I., Levy, R., & Rayner, K. (2014). Task effects reveal cognitive flexibility responding to frequency and predictability: Evidence from eye movements in reading and proofreading. *Cognition, 131,* 1–27.

Schotter, E. R., Reichle, E. D., & Rayner, K. (2014). Rethinking parafoveal processing in reading: Serial attention models can account for semantic preview benefit and n+2 preview effects. *Visual Cognition, 22,* 309–333.

Sereno, S. C., O'Donnell, P. J., & Rayner, K. (2006). Eye movements and lexical ambiguity resolution: Investigating the subordinate-bias effect. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 335–350.

Sereno, S. C., & Rayner, K. (2000). Spelling-sound regularity effects on eye fixations in reading. *Perception & Psychophysics, 62,* 402–409.

Shen, E. (1927). An analysis of eye movements in the reading of Chinese. *Journal of Experimental Psychology, 10,* 158–183.

Sheridan, H., Rayner, K., & Reingold, E. M. (2013). Unsegmented text delays word identification: Evidence from a survival analysis of fixation durations. *Visual Cognition, 21,* 38–60.

Sheridan, H., & Reingold, E. M. (2012a). The time course of predictability effects in reading: Evidence from a survival analysis of fixation durations. *Visual Cognition, 20,* 733–745.

Sheridan, H., & Reingold, E. M. (2012b). The time course of contextual influences during lexical ambiguity resolution: Evidence from distributional analyses of fixation durations. *Memory & Cognition, 40,* 1122–1131.

Slattery, T. J., Angele, B., & Rayner, K. (2011). Eye movements and display change detection during reading. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1924–1938.

Slattery, T. J., Schotter, E. R., Berry, R. W., & Rayner, K. (2011). Parafoveal and foveal processing of abbreviations during eye fixations in reading: Making a case for case. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 1022–1031.

Sparrow, L., & Miellet, S. (2002). Activation of phonological codes during reading: Evidence from errors detection and eye movements. *Brain and Language, 81,* 509–516.

Staub, A., Rayner, K., Pollatsek, A., Hyönä, J., & Majewski, H. (2007). The time course of plausibility effects on eye movements in reading: Evidence from noun-noun compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 1162–1169.

Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye movements in reading. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1280–1293.

Tsai, J. L., Lee, C. Y., Tzeng, O. J. L., Hung, D. L., & Yen, N. S. (2004). Use of phonological codes for Chinese characters: Evidence from processing of parafoveal preview when reading sentences. *Brain and Language, 91,* 235–244.

Vainio, S., Hyönä, J., & Pajunen, A. (2009). Lexical predictability exerts robust effects on fixation duration, but not on initial landing position during reading. *Experimental Psychology, 56,* 66–74.

Vitu, F., O'Regan, J. K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception & Psychophysics, 47,* 583–600.

Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review, 14,* 770–775.

Warren, T., McConnell, K., & Rayner, K. (2008). Effects of context on eye movements when reading about possible and impossible events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1001–1010.

White, S. J., Rayner, K., & Liversedge, S. (2005). The influence of parafoveal word length and contextual constraint on fixation durations and word skipping in reading. *Psychonomic Bulletin & Review, 12,* 466–471.

Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology, 97,* 259–268.
Yang, J. (2013). Preview effects of plausibility and character order in reading Chinese transposed words: Evidence from eye movements. *Journal of Research in Reading, 36,* S18–S34.
Yang, S. N. (2006). An oculomotor-based model of eye movements in reading: The competition/interaction model. *Cognitive Systems Research, 7,* 56–69.
Yang, S. N., & McConkie, G. W. (2001). Eye movements during reading: A theory of saccade initiation times. *Vision Research, 41,* 3567–3585.
Yen, M.-H., Tsai, J.-L., Tzeng, O. J.-L., & Hung, D. L. (2008). Eye movements and parafoveal word processing in reading Chinese. *Memory & Cognition, 36,* 1033–1045.
Zola, D. (1984). Redundancy and word perception during reading. *Perception & Psychophysics, 36,* 277–284.

PART 2

Word Identification

CHAPTER

5 Visual Word Recognition in the Bayesian Reader Framework

Sachiko Kinoshita

Abstract

Visual word recognition is traditionally viewed as a process of activating a lexical representation stored in long-term memory. Although this activation framework has been valuable in guiding research on visual word recognition and remains the dominant force, an alternative framework has emerged in the last decade. The Bayesian Reader framework, proposed by Norris (2006, 2009; Norris & Kinoshita, 2012a), regards the decision processes involved in a task as integral to explaining visual word recognition, and its central tenet is that human readers approximate optimal Bayesian decision-makers operating on noisy perceptual input. This chapter focuses on two issues fundamental to visual word recognition—the role of word frequency and the representation of letter order—and describes how the Bayesian Reader framework provides a principled account of the recent findings related to these issues that are challenging to the activation framework.

Key Words: visual word recognition, interactive activation, masked priming, letter-order coding, word frequency, Bayesian Reader

In research on visual word recognition, the term *lexical access* is used to describe the state where a match has been found between the visual input and a representation stored in the reader's internal lexicon. Balota (1990) called it a "magic moment" (p. 9) at which the word has been recognized as familiar but its meaning is not yet retrieved; in the E-Z Reader model of eye movements in reading (Reichle, Pollatsek, Fisher, & Rayner, 1998), it is regarded as the point at which attention is shifted from the current word to the next word.

In explaining the lexical access process, the notion of activation has played and continues to play a dominant role. The idea was formalized in the interactive-activation (IA) model put forward by McClelland and Rumelhart (1981), which has served as the basis for many subsequent computational models of visual word recognition. The basic idea is that words are represented as nodes in a network, and a word is said to be recognized when the activation level in a word node reaches a threshold. Each word node is connected to the letter units and to other word nodes. The word node's activation level is boosted as it receives activation from the letter units the word contains, and it is inhibited by competing word nodes. This basic architecture and the notion of activation have been adopted in many computational models of visual word recognition, such as the dual-route cascaded (DRC) model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), the multiple readout (MROM) model (Grainger & Jacobs, 1996), and the spatial coding model (SCM) (Davis, 2010).

While acknowledging that activation has been an "extremely valuable metaphor," Norris and McQueen (2009) have questioned the explanatory value of the concept of activation. They note that "beyond the general notion that bigger is better, activation does not directly determine the behavior of these models. In particular, neither reaction time nor response probabilities can be derived directly from activation values without additional assumptions" (p. 357). These comments were made in the context of models of spoken word recognition, but the same could be said of models of visual word recognition. In this chapter I will describe an alternative, the Bayesian Reader framework, proposed by Norris (2006, 2009; Norris & Kinoshita, 2008, 2012a). In brief, the Bayesian Reader framework views the human reader as accumulating noisy evidence from the perceptual input and as making optimal decisions as required by the task in accordance with Bayes' theorem. I will discuss recent research findings on two issues that are fundamental to visual word recognition—the role of word frequency and the representation of letter order—within the Bayesian Reader framework, and show how this framework provides a more principled and coherent account of the findings than is possible within the activation framework.

Word Frequency

Word frequency is the single most powerful determinant of the ease of word identification: Words that appear frequently in the language are recognized more readily than words that appear less frequently. This holds across different languages and different laboratory tasks used to study word identification, such as the lexical decision task (in which participants are asked to decide whether a letter string is a word) and the read-aloud task (also referred to as the speeded naming or pronunciation task). Word frequency (log frequency) is the single most important predictor of reaction time (RT, correlation ~-.6) and accuracy (~.4) in the large-scale lexical decision databases containing tens of thousands of words (see Yap & Balota, this volume, regarding the development of these databases), such as the ELP (English Lexicon Project; Balota et al., 2007), BLP (British Lexicon Project; Keuleers, Lacey, Rastle, & Brysbaert, 2012), DLP (Dutch Lexicon Project; Keuleers, Dieppendaele, & Brysbaert, 2010), and FLP (French Lexicon Project; Ferrand et al., 2010). It is also a main predictor of pronunciation latency, accounting for 7% to 10% of variance in a large-scale word naming database in English (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Spieler & Balota, 1997). In perceptual identification (e.g., Howes & Solomon, 1951) and eye-movement studies of sentence reading (Rayner, Ashby, Pollatsek, & Reichle, 2004; Schilling, Rayner, & Chumbley, 1998) word frequency also has a large effect; that is, high-frequency words have a lower threshold of identification in perceptual identification, are more likely to be skipped, and are fixated for a shorter duration than low-frequency words in sentence reading. There has been a renewed focus on the study of this most powerful variable in visual word recognition, from both a methodological perspective improving on the estimates of word frequency and a theoretical perspective explaining why word frequency has the effect it has on visual word recognition.

To study the role of word frequency in word recognition requires accurate estimates of how often a word occurs in a language. Brysbaert and colleagues led the charge in improving the frequency measures. Brysbaert and New (2009) identified a number of problems with Kučera and Francis's (1967) frequency norms, which had been the main resource for visual word recognition researchers working in English. Brysbaert and New noted that because of its limited corpus size (1 million word tokens), the estimates of frequency, particularly for low-frequency words, were unreliable, and also that word frequency based on the Kučera and Francis norms accounted for a smaller amount of variance in ELP lexical decision data (both accuracy and RT) than other estimates of word frequency based on much larger text corpora (e.g., Hyperspace Analogue of Language, with 131 million words; Lund & Burgess, 1996; Celex, with 16.6 million words; Baayen, Piepenbrock, & Gulikers, 1995). Brysbaert and New provided a new frequency norm based on subtitles of 8,388 US films and television episodes with 51 million words, $SUBTLEX_{US}$ frequency, and showed that it provided the best correlation with the ELP data. More recently, Van Heuven, Mandera, Keuleers, and Brysbaert (2014) provided $SUBTLEX_{UK}$, based on British English, and showed that these word frequency norms explained 3% more variance in the lexical decision times of the British Lexicon Project than the $SUBTLEX_{US}$ word frequencies. With these improved frequency estimates, Keuleers et al. (2012) made the important observation that lexical decision RT is a continuous function of log frequency throughout the frequency range, and is not limited to a difference between very high-frequency words (over 100 per million) and lower frequency words.

These improved methods for estimating the frequency of occurrence provide a valuable resource, but they do not tell us why high frequency words are recognized more readily than low-frequency words. Furthermore, both in perceptual identification (where a single word is presented briefly for a variable duration and the presentation time required for a given level of accuracy—threshold time—is the dependent measure) and speeded response tasks like lexical decision and word naming, the function that relates the dependent variable (threshold time in perceptual identification, accuracy and RTs in the speeded tasks) to the frequency of occurrence is not linear, but logarithmic (e.g., Keeulers et al., 2012). Few have attempted to explain why this should be the case. The original interactive activation model of visual word recognition (McClelland & Rumelhart, 1981) built in log frequency as the resting level of lexical representation, and subsequent models based on the IA framework such as the DRC model (Coltheart et al., 2001) and MROM (Grainger & Jacobs, 1996) followed this practice, but they do not provide a theoretical justification for this assumption.

Another important finding is that in lexical decision the size of word frequency effect is modulated by the type of nonword foils (e.g., Lupker & Pexman, 2010; Ratcliff, Gomez, & McKoon, 2004). Specifically, the size of the word frequency effect in lexical decision is reduced when the nonword foils are less wordlike, making word-nonword discrimination easier. In Ratcliff et al. (2004), the difference in mean RT between the same high- versus low-frequency words was 75 ms when the nonword foils were pronounceable pseudowords generated by replacing the vowels of existing words with other vowels (e.g., *SPONT*), but the difference shrank to 38 ms when the nonwords were random letter strings. Similarly, in Lupker and Pexman (2010), the word frequency effect for the same set of words was larger when the nonword foils resembled specific words—when nonwords were pseudohomophones (nonwords that sound like a real word when pronounced, e.g., *BRANE*) or when the nonwords were generated by transposing internal letters of an existing word (e.g., *JUGDE*). It is difficult to see within the IA framework why the nature of nonwords should affect the resting level of lexical representations.

The Bayesian Reader

Norris (2006, 2009) has argued that the effect of frequency on word recognition tasks and the specific form of function relating word frequency to RTs and identification threshold naturally follow from the assumption that in visual word recognition readers behave as optimal Bayesian decision-makers, with word frequency serving as an index of prior probability.

Bayes' theorem, shown in the following equation, provides the optimal procedure for combining noisy data with knowledge of prior probability:

$$P(H \mid D) = P(H) \times P(D \mid H) / P(D)$$

where $P(H|D)$ is the probability for the hypothesis H given the data D (i.e., the posterior probability in the light of data), $P(H)$ is the prior probability for the hypothesis H, $P(D|H)$ is the likelihood of observing the data D when the hypothesis H is true, and $P(D)$ is the probability of observing the data (which reduces to a constant in cases where the same data are used).

An often-used example to illustrate the application of Bayes' theorem is that of medical diagnosis: One wants to compute the probability that a patient has disease X given that the patient tests positive in a test that has a hit rate of, say, 95% (i.e., the test correctly produces a positive result for the disease 95% of the cases when the disease is present) and a false alarm rate of say 10% (i.e., the test incorrectly produces a positive result on 10% of the cases when the disease is not present). Application of Bayes' theorem shows that when the evidence provided by the test is uncertain (i.e., the hit rate is not 100% and the false alarm rate is not 0%), the probability that the patient testing positive on the test really has the disease is much lower if the disease has a lower base rate (i.e., when the disease is rare). For example, if the disease occurs in 1 in 1,000 people in the population, the probability that the patient testing positive on the test has the disease is .009 (= .001 × .95/[(.001 × .95) + ((1 −.001) × .1)]), whereas if the disease is common, occurring in 1 in 5 people in the population, the probability that a patient testing positive on the test has the disease is .70 (= .2 × .95/[(.2 × .95) + ((1 −.2) × .1)]). To put it simply, application of Bayes' theorem provides a correction for false alarms—because there is a substantial probability that the test will produce a false positive result, the rarer the disease, the more likely that the positive result is a false alarm.

Similarly, in the Bayesian Reader model of visual word recognition, readers accumulate evidence from the visual input via noisy perceptual sampling, and make decisions as required by the task. The decision in perceptual identification would be which word the input corresponds to; in lexical decision, the decision is whether the input is a word or a nonword. Note that the readers need not be instructed to express the decision by an overt response, for example, by means of a button press. In sentence reading, for example, the decision required is when to move the eyes from one word to the next, and it is indexed by the fixation duration. The key idea in the Bayesian Reader model's explanation of word frequency effect is that, when the evidence is uncertain (because the perceptual evidence is noisy and there is a possibility of misperception, equivalent to a false alarm in diagnostic tests), it is optimal to take into consideration prior probability in accordance with Bayes' theorem. To put it another way, when there is ambiguity in the perceptual data, word frequency effectively alters the weighting of the evidence. In psychophysics, Piéron's law (e.g., Stafford & Gurney, 2004) states that the time to reach a decision is an exponentially decreasing function of the strength of evidence: the stronger the evidence, the faster the time to reach a decision. In line with this view, Norris (2006) showed that in the Bayesian Reader simulation of perceptual identification and lexical decision, the model RT correlated linearly with log frequency. Norris (2009) also showed that the Bayesian Reader model correctly simulated that word frequency affects the rate of evidence accumulation, and the effect of word frequency in lexical decision varies with the type of nonword foils (which affects the ease of word-nonword decision), as Ratcliff et al. (2004) have shown.

The conceptualization of word frequency as a measure of prior probability within a Bayesian framework offers a potential resolution to various recent debates and questions concerning the role of word frequency in visual word recognition. One debate concerns the question of whether contextual diversity provides more explanatory power than word frequency. Contextual diversity is operationalized as the number of passages (text samples) a word occurs in. Although it is highly correlated with the frequency of occurrence, contextual diversity would be lower for words equated on frequency of occurrence across text samples if the word occurs in more specialized, restricted contexts (e.g., *neutron, joust*). Adelman, Brown, and Quesada (2006) showed that contextual diversity is a better predictor of lexical decision and pronunciation latencies in the ELP data, accounting for 1% to 3% more variance than word frequency (based on Kučera and Francis's [1967] corpus, Touchstone Applied Science Associates corpus, and British National Corpus), and the finding was replicated by Brysbaert and New (2009) based on their $\text{SUBTLEX}_{\text{US}}$ corpus. Using low-frequency words (for which frequency and contextual diversity measures diverge more), Plummer, Perea, and Rayner (2014) extended the finding to eye-movement measures of word recognition in sentence reading, showing that contextual diversity is a better predictor than word frequency. Norris (2009) argued that the slightly better performance of contextual diversity would be expected from the view that word frequency provides a poor estimate of prior probability of words presented in isolation (without context) if the distribution of the occurrence of words is not uniform (i.e., if words vary in contextual diversity, with some words occurring only in restricted contexts and others occurring across diverse contexts). Although the words were not presented in isolation in Plummer et al.'s (2014) study, the same point applies, as the sentence context was neutral: The sentences were constructed so that the critical word was not predictable at the point it occurred (e.g., "In class, Howard learned the role of the *neutron* in an atom's structure and function."). This explanation of the relationship between word frequency and contextual diversity finds support in an analysis reported by Brysbaert and New (2009). Brysbaert and New observed that many words with higher word frequency/contextual diversity ratios (i.e., words that occur in specific contexts) were words that could be a proper name (e.g., *prince, drake*), just the sort of words that would be expected to occur in specific contexts. When they recalculated word frequency only for the cases when the word started with a lowercase letter (which would generally exclude the cases where the word is used as a proper name), the advantage of contextual diversity as a predictor of lexical decision latency and accuracy over word frequency was attenuated. In other words, one way to think about contextual diversity is that it provides a better estimate of frequency across different contexts. In sum, the Bayesian Reader offers a principled and coherent explanation of the role of word frequency and contextual diversity—both are construed as a measure of prior probability.

A recent methodological advance in studying the role of word frequency is the size of corpus used in estimating the frequency of occurrence. Brysbaert and New (2009) suggested that whereas the frequency estimates of high-frequency words are relatively stable, words that occur less than 10 times per million require a large corpus of at least 16 million words for reliable estimates of frequency. Kuperman and Van Dyke (2013) confirmed this observation regarding the stability of frequency estimates through a sampling study. However, they pointed out that frequency counts from a large corpus tend to overestimate the occurrence of rare words in smaller samples by assigning larger-than-zero frequencies to a large percentage of words that are not part of an individual's vocabulary. That is, it is not necessarily the case that the larger the corpus, the better. Rather, objective frequency counts of low-frequency words based on large text corpora may actually overestimate the prior probability of these words an individual reader knows, particularly for those with a small vocabulary. This may explain why Brysbaert and New (2009) found that the amount of variance in the lexical decision RT accounted for by log frequency increased with the corpus size used to estimate the frequency, but the gains made by the increase in corpus size leveled off at about 16 million words—a corpus size of 32 million or 88 million did not fare better (see their Table 2, p. 980). That the objective frequency counts based on very large text corpora may overestimate subjective frequency, especially for those with a small vocabulary, would be a point to consider when investigating the role of vocabulary size in word recognition tasks (see Andrews, this volume; Yap & Balota, this volume). Here again, the Bayesian Reader framework's conceptualization of word frequency as a measure of subjective prior probability—which would naturally vary with the reader's vocabulary—provides a useful framework.

Representing Letter Order

The second issue for which the Bayesian Reader has served as a useful theoretical framework relates to the representation of letter order. In this section, I will first describe the issues with representing letter order that are problematic for the models of visual word recognition based on the IA framework. I will then describe the two main approaches taken to the problem: one that preserves the notion of activation but proposes a new representation that mediates between the letter level and the word level and the other based on the notion of noisy perception, consistent with the Bayesian Reader framework. I will then evaluate these approaches with respect to the findings observed with the masked priming paradigm.

PROBLEMS WITH THE SLOT CODE

A word on a page consists of a string of letters. (This is true of all current writing systems, including unpointed Hebrew in which vowels are not represented, or in Chinese, in which a character, which can be conceived of as a letter, denotes a morpheme.) Because most words in an adult reader's vocabulary comprise more than a single letter/character, it is necessary to code the order of the letters. In the last decade, there has been a spirited debate regarding how to represent the order of letters in a word (see also Frost, this volume; Perea, this volume). This focus on the representation of letter order stems from the recognition of a shortcoming in the way in which letter order is represented in the classic visual word recognition models. (In the interest of brevity, I will focus here only on the slot-code representation. For a more extensive coverage of other earlier views on the coding of letter order, readers are referred to, e.g., Davis, 2006).

The original IA model (McClelland & Rumelhart, 1981), and models based on the IA framework such as the DRC (Coltheart et al., 2001) and MROM (Grainger & Jacobs, 1996) use the slot-coding scheme. In the original IA model (whose vocabulary consisted of four-letter words only), there are separate slots for each possible letter position within a word, and letter identities are associated with specific slots. For example, the word *TIME* would be represented as $T_1I_2M_3E_4$, with the letter *T* associated with the position 1 slot, letter *I* in position 2, and so on. In contrast, the word *ITEM* would be represented as $I_1T_2E_3M_4$. This means that the letters *T, I, M*, and *E* in *TIME* and *ITEM* are effectively different letters (T_1 in *TIME,* T_2 in *ITEM*, and so on). The slot-coding scheme thus allows anagrams like *TIME* and *ITEM* to be distinguished. This is important in alphabetic writing systems, in which the number of letters is severely limited, and hence there are very many anagrams. (Shillcock, Ellison, & Monaghan, 2000, reported that one-third of three- and four-letter words in English are anagrams.) However, the slot-coding scheme is challenged by phenomena showing that readers are reasonably tolerant of distortions of canonical order of the letters in a word.

The transposed-letter (hereafter TL) priming effect (e.g., Forster, Davis, Schoknecht, & Carter, 1987; Kinoshita & Norris, 2009; Perea & Lupker, 2003) refers to the finding that in masked priming (see the section "Masked Priming" for more detail about the procedure), a prime generated by transposing two adjacent letters in a word (e.g., *jugde*) facilitates the recognition of the base word (*JUDGE*) almost as much as the word itself, and more than a prime generated by replacing the corresponding letters with other letters not in the word (two-substituted-letters [2SL] prime, e.g., *junpe*). In both the TL prime and the 2SL prime, the slots corresponding to the third and fourth letters have the wrong letter identities. According to the slot-coding scheme, a TL prime and a 2SL prime are equally similar to the base word (*JUDGE*), and hence should produce equal priming effects. Note that the visual similarity of the substituted letter (e.g., *G* is more similar to *C* than to *P*) has little impact on the size of priming (see Kinoshita, Robidoux, Mills, & Norris, 2014).

A related problem with the slot-coding scheme is that it cannot capture the similarity between letter strings differing in length that contain the same sequence of letters. Specifically, DRC uses a beginning-anchored slot-coding scheme. According to this scheme, the similarity between two letter strings differing in length can be captured if the letter sequences overlap at the beginning of a word (e.g., *STAR* and *START*), but not at the end of a word (e.g., *PRAY* and *SPRAY*). Such schemes also cannot capture the similarity between the letter strings that differ by deletion/addition of letters in the middle of the word (e.g., *aprt-APRICOT; journeal-JOURNAL*) that maintain the general order of letters. The fact that such primes produce robust facilitation—referred to as the relative position priming effect (Grainger, Grainier, Farioli, van Assche, & van Heuven, 2006; Van Assche & Grainger, 2006)—presents a further problem for the slot-coding scheme.

Alternatives to the Slot Code: Two Approaches

In the last decade, many models have been developed to meet the challenge posed by the phenomena such as the TL priming and relative position priming effects. These models can be broadly classified into those that posit a representation that mediates between the letter level and the word level and those that explain the phenomena in terms of perceptual noise. Various versions of open bigram models are main examples of the first approach, and the overlap model (Gomez, Ratcliff, & Perea, 2008), and the noisy slot Bayesian Reader model (Norris, Kinoshita, & van Casteren, 2010) and its successor, the noisy channel model (Norris & Kinoshita, 2012a) are the main examples of the latter approach, and I will focus on these models here.

Open Bigram Models

Open bigrams (OBs) are ordered letter pairs (bigrams) that can be contiguous or noncontiguous. For example, the word *CAT* contains the contiguous OBs *CA, AT*, and the noncontiguous OB *CT*. The key claim of OB models is that a word is represented as a collection of OBs. For example, *CAT* is represented as {*AT, CT, CA*}. That is, OBs represent an intermediate level between the letter level and the word level. Grainger and Whitney (2004) suggested that OBs provide a natural explanation for experimental data demonstrating TL priming and relative-position priming effects. Specifically, the amount of priming is assumed to be a function of the number of OBs shared by the letter strings. For example, if all OBs are represented, *JUDGE* contains the following 10 OBs: *JU, JD, JG, JE, UD, UG, UE, DE, GE*, and *DG*. The TL prime *jugde* shares all of the OBs except for *DG*. In contrast, the 2SL prime *junpe* shares only 3 OBs with the target (*JU, JE*, and *UE*). Accordingly, the TL prime produces a greater priming effect than the 2SL prime. Relative-position priming effects are also explained in terms of the large number of OBs shared between the superset and subset primes (e.g., *aprt-APRICOT; journeal-JOURNAL*) and the target.

There are variations among the OB models. Earlier versions of OB models considered all OBs contained in a letter string irrespective of distance between the letter pairs, but all of the more recent versions limit the separation to two intervening letters (e.g., in *JUDGE JE* is not counted). Dehaene, Cohen, Sigman, and Vinckier (2005) motivated this assumption on neurobiological grounds, based on the notion of a neuronal hierarchy along the ventral visual pathway along which the receptive size increases by a factor of 2 to 3 at each stage. In their local combination detector model, bigram detectors respond to two-letter combinations in specific order if the first letter of a pair is less than two letters away from the second (e.g., the bigram neuron *AM* will fire in response to *HAM* and *ATOM* but not to *ALARM* or *ATRIUM*).

The OB models also differ in terms of whether all OBs are weighted equally. In Grainger and Van Heuven's (2003) parallel OB model, all bigrams are weighted equally, irrespective of the number of letters intervening between the letter pairs. In SERIOL (sequential encoding regulated by inputs to oscillations within letter units; Whitney, 2001, 2008), bigrams are weighted differentially according to the separation between the letter pair. In the most recent version (Whitney, 2008), the adjacent bigrams are weighted 1.0, OBs spanning one intervening letter .8, and OBs spanning two intervening letters .4. In addition, the bigrams involving the initial or final letter in the string and an edge character (edge bigrams) are also weighted 1.0, thus according greater similarity between two letter strings that share the initial or final letters.

Finally, the OB models differ in terms of whether the OBs are activated for the reversed order of the letters. Grainger et al.'s (2006) overlap OB model allows this possibility by incorporating the noisy letter-position assumption as proposed by Gomez et al. (2008) in their overlap model. The overlap OB model "attempts to code for contiguous letter sequences (bigrams) only . . . however, the noisy coding of letter position implies that noncontiguous letter sequences (open bigrams) are formed as well as transposed bigrams" (Grainger et al., 2006, p. 883). The weighting of contiguous and noncontiguous bigrams in this model is similar to that of SERIOL, in that the weighting of bigrams is graded according to the number of intervening letters: Adjacent bigrams are weighted 1.0, OBs spanning one letter are weighted .607, OBs spanning two intervening letters are weighted .135, and reversed contiguous bigrams (e.g., *BA* in *TABLE*) are weighted .135.

Overlap Model

The key assumption of the overlap model (Gomez et al., 2008) is that in visual perception, the representation of location is not a point, but is distributed over space (as suggested by Logan, 1996, and Ratcliff, 1981). According to the overlap model, the identities of the letters in a letter string are assumed to be normally distributed over position. For example, in the string *TRAIL, A* will be associated with position 3 but to a lesser degree with position 2 and 4 and, depending on the degree of spread, with position 1 and 5 as well. It is essentially a noisy-slot model.

Noisy-Channel Model

Norris and Kinoshita's (2012a) noisy-channel model is based on the Bayesian Reader framework (Norris, 2006). As described earlier, the key assumption of the Bayesian Reader framework is that readers make near optimal decisions based on noisy (uncertain) evidence accumulated from the perceptual input. Bayes' theorem provides the optimal procedure for combining uncertain evidence with knowledge of prior probability. While the original Bayesian Reader model (Norris, 2006) incorporated noise in letter identity information, it did not incorporate uncertainty in the letter-position information or in whether (any) letter was present/absent (i.e., the number of letters was certain). It was effectively a slot model. Norris et al. (2010) showed through model simulations that incorporating the noisy position assumption allowed the model to capture the TL priming effect. The noisy-channel model (Norris & Kinoshita, 2012a) further extended the noise assumption to the presence/absence of letters in the visual input (i.e., the human visual perception system—the noisy channel—could either insert a spurious letter or delete a letter), which allowed the model to simulate relative position priming with superset and subset primes (i.e., primes generated by inserting/adding or deleting letter[s] from the target, e.g., *journeal-JOURNAL; aprt-APRICOT*).

MASKED PRIMING

Masked priming has been the most commonly used experimental paradigm in studying the coding of letter order. In the procedure developed by Forster and Davis (1984), standardly used in visual word recognition studies, the sequence of events in a trial consists of (1) a forward mask, typically a string of #s presented for 500 ms; (2) a prime presented briefly (usually 50 ms); and (3) a target, to which a response is required, presented until the participant's response. Typically, the prime is presented in lowercase and the target in uppercase, so that even when they share letters in the same position, there is usually little physical overlap and the target effectively masks the prime. At the prime duration of 50 ms, with the target serving as a backward mask, and the forward mask making the detection of the onset of the prime difficult, participants are generally unaware of the presence of the prime, let alone its identity. It is generally agreed therefore that the influence of the prime on the target is automatic, in that it is free of response strategies that are possible with visible primes.

Using the masked priming procedure, researchers have accumulated a large data set of orthographic priming effects produced by a variety of primes—substituted-letter prime, where a letter in a target is substituted by a different letter (e.g., *mudge-JUDGE*); TL prime, where letters in the target are transposed (e.g., *jugde-JUDGE*); and relative-position primes, where the relative order of the letters is maintained despite either an addition of letter(s) (superset prime, e.g., *journeal-JOURNAL*) or deletion of a letter(s) (subset prime, e.g., *aprt-APRICOT*). Most of these orthographic priming data have come from European languages that use the Latin alphabet, such as English, Spanish, and French, using the lexical decision task. In contrast to these languages, it is important to note that TL priming effects are absent in Hebrew (e.g., Velan & Frost, 2009). I will return to this point later.

Given that all models are generally able to fit the basic TL priming and relative position priming effects, more specific differences predicted by the different models are reported for the purpose of theory adjudication. In discussing such data, several points should be noted about masked priming.

Masked Priming Data Are Noisy

One point is that masked priming data are noisy. Consider the so-called relative position constraint. Peressotti and Grainger (1999) reported that transposition of medial letters in a subset prime (e.g., *gdrn-GARDEN*) eliminated priming. This finding has been contrasted with the robust priming effects observed with full-length primes with letter transposition (e.g., *gadren-GARDEN*), and it has been taken to argue that "letter order is highly important when the prime is comprised of a restricted subset of the target's letters" but that "when all letters of the target are present in the prime, maintenance of letter order is less important" (Whitney, Bertrand, & Grainger, 2012, p. 110; see Grainger & Whitney, 2004, for a similar claim).

Peressotti and Grainger's (1999) conclusion was based on a statistically significant 20-ms priming effect by the subset prime with canonical letter order (e.g., *grdn*) and a statistically nonsignificant 5-ms priming effect produced by a TL-subset prime (e.g., *gdrn*), but they did not test whether the size of the two effects differed statistically. Stinchcombe, Lupker, and Davis (2011) did make this comparison and concluded that, contrary to the claim by Whitney et al. (2012) and Grainger and Whitney (2004), there is no reliable evidence that the letter transposition has a more detrimental effect in subset primes than in full-length primes. In Stinchcombe et al.'s data, although the relative size of priming was numerically similar to that of Peressotti and Grainger (1999), there was no statistical difference in the size of priming produced by a subset prime with the canonical letter order and the TL subset prime.

As attested by the preceding example, given that the upper limit of masked priming effect in a lexical decision task is about 50 ms (found with identity primes), a difference in the size of priming produced by different primes is necessarily limited. Caution is needed in using the difference (or the absence of difference) in the size of priming to draw theoretical conclusions. In recognition of this problem, a group of researchers (Adelman et al., 2014) recently collected masked priming data from over 1,000 participants from the United Kingdom, Australia, the United States, and Canada, using 420 six-letter word targets. In this dataset, a priming effect of 2.9 ms is significant at the 5% level. Such a dataset would be useful in countering the problem of limited statistical power associated with small datasets as well as testing the replicability of a finding reported in smaller scale studies. Indeed, Adelman et al. noted that an effect that was argued by Davis and Bowers (2006; see also Davis, 2006) to be one of the benchmark effects that any model of orthographic input coding must explain—namely, that a substitution neighbor (e.g., *azkle-ANKLE*) produces more priming than a prime involving transposition of the substituted letter (dubbed the neighbor once-removed, e.g., *akzle-ANKLE*)—was not replicated in the masked priming database. Other instances of failure to replicate include morphological boundary effects. Initial reports (e.g., Christianson, Johnson, & Rayner, 2005; Duñabeitia, Perea, & Carreiras, 2007) indicated that TL priming is eliminated when the transposed letters straddle a morphological boundary (e.g., *accidenatl-ACCIDENTAL*). Later studies (e.g., Rueckl & Rimzhim, 2011; Sánchez-Gutiérrez & Rastle, 2013), however, found equally robust TL priming effects across and within morphological boundaries. These findings highlight the need for caution in using small differences in the size of priming produced by different primes, and call for the need for replication before drawing theoretical implications.

The Use of Match Scores

Another point to note in using masked priming data for theory adjudication concerns the use of match scores. Match scores are an index of orthographic similarity between two letter strings, and range between 0 for no overlap (as would be for two letter strings that share no letters anywhere in the strings, e.g., *NOBLE* and *DRIFT*) and 1 for a perfect overlap. Match scores carry the explanatory burden in models that explain orthographic priming effects in terms of overlap of specialized orthographic representations such as OBs. (A useful resource for calculating match scores for these models has been provided by Davis, and at the time of writing this chapter is available at www.pc.rhul.ac.uk/staff/c.davis/Utilities/MatchCalc/.) In contrast, in the noisy-channel model (and the noisy-slot model) orthographic similarity is not a static quantity, but varies dynamically with time (i.e., with greater perceptual sampling, the perceptual information becomes more certain). In this perspective, a TL prime (e.g., *jugde* for *JUDGE*) produces greater priming effect than an SL prime (e.g., *junpe*) because in the limited time that the prime is available, information about letter identity is more certain (and hence the two wrong letters in *junpe* indicate that it is different from the base word *JUDGE*) than the relative order of two adjacent letters. But with unlimited exposure there would be no uncertainty in the perceptual information, and readers can readily decide with certainty that both *jugde* and *junpe* differ from the base word *JUDGE*.

Earlier studies (e.g., Davis & Bowers, 2006; Whitney, 2008) sought to adjudicate between theories by comparing the size of masked priming effect against the predictions derived from match scores. The assumption underlying this approach is summarized well in the words of Stinchcombe et al. (2011): "Everything else being equal, primes that are more similar to their targets should produce greater facilitation in a masked priming experiment. Hence, one can generally determine whether, and to what degree, a given model predicts priming effects in a masked priming task" (p. 478). More recently, however, it has been recognized that everything is not equal: The size of orthographic priming effects in lexical decision is not a simple function of match scores (e.g., Guerrera & Forster, 2008; Kinoshita & Norris, 2009; Lupker & Davis, 2009), but it is modulated by the lexical characteristics of the target. For example, the letter transposition manipulation produces the same match scores for nonword-word target pairs (e.g., *fiath-FAITH*) and nonword-nonword target pairs (e.g., *biath-BAITH*), but produces no priming for the nonword targets. Similarly, match scores are uninfluenced by the neighborhood density (the number of orthographically similar words or neighbors; see Perea, this volume, for details), but masked priming effects are reduced for words that have many neighbors.

Evaluation: The Same-Different Task

Given the preceding considerations in using the masked priming data for theory adjudication, one approach that has been taken is to computationally implement the lexical decision task and to test the fit between the model's performance and the human data (see Davis, 2010, and Norris & Kinoshita, 2012a, 2012, for this approach). This has not yet been done with the many versions of the OB models. Another approach, one discussed in detail here, is to compare the match score predictions against a task in which the masked priming effects are not modulated by lexical characteristics of the stimuli, unlike the lexical decision task. The same-different task was developed to serve this purpose (see Kinoshita & Norris, 2012; Norris & Kinoshita, 2008). In this task, a reference item is shown for 1 second immediately before the prime, and the participant is instructed to decide whether the target is the same as or different from the reference item. The reference item and target are in different cases so that the match cannot be based on physical identity. (See Figure 5.1 for a trial sequence.)

According to the Bayesian Reader account of masked priming proposed by Norris and Kinoshita (2008), priming reflects the evidence contributed by the prime toward the decision required to the target. Within this view the computations involved in the same-different task and the lexical decision task are similar, differing primarily in the set of items that the target is matched against. In the same-different task, the match involves the target against the referent; in lexical decision, instead of a single referent, the target is matched against words in the lexicon (to be more precise, the words in the neighborhood of the target). This view is contrasted with the activation framework, according to which the masked priming effect reflects the prime

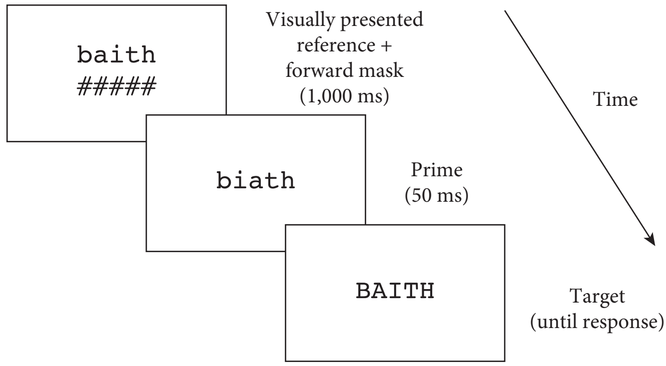


Fig. 5.1 Trial sequence in the masked priming same-different task.

preactivating the corresponding representation in the mental lexicon. Such a view predicts no priming for nonword targets, for which there is no representation in the lexicon.

Contrary to the activation account but as expected from the Bayesian Reader account, Norris and Kinoshita (2008) showed that the same nonword targets (e.g., *BAITH*) for which priming was absent in lexical decision produced robust identity priming effects in the same-different task. Using this task, Kinoshita and Norris (2009, Experiment 3) showed that nonwords produce robust TL priming effects (e.g., *biath-BAITH*), no smaller in size than the effect observed with word targets (e.g., *fiath-FAITH*). This finding of TL priming effects with nonword stimuli has since been replicated many times, for example by Perea and Acha (2009) in Spanish, and even with nonlinguistic symbols (García-Orza, Perea, & Muñoz, 2010). These findings indicate that the origin of TL priming effect is perceptual. Specifically, consistent with the assumption of models such as the overlap model (Gomez et al., 2008), the noisy-slot Bayesian Reader model (Norris et al., 2010), and the noisy-channel model (Norris & Kinoshita, 2012a), the perception of position of letters in a multiletter string is noisy. Hence, in the limited time the prime is available, the information regarding the precise order of letters is uncertain. In contrast, according to OB models, TL priming effects are not expected for nonword targets because nonwords—which, by definition, are not represented in the lexicon and do not have word nodes—are not coded by OBs.

The use of the same-different task has not been limited to demonstrating TL priming effects with nonwords. Contrasting the masked priming effects obtained with this task with the lexical decision task (and other reading tasks) is informative in showing how the uncertainty in letter position interacts with the lexical environment. This is most clearly illustrated with Hebrew. Across a range of reading tasks including reading of sentences with rapid serial presentation of words, masked priming lexical decision, and sentence reading measuring eye movements, Velan and Frost (2007, 2009, 2011; Velan, Deutsch, & Frost, 2013) have demonstrated that Hebrew readers, unlike readers of European languages that are written in the Latin alphabet, are not tolerant of letter transposition. Frost (2012; see also Frost, this volume) took these data to argue that "letter-order insensitivity… is a variant and idiosyncratic characteristic of some languages, mostly European, reflecting a strategy of optimizing encoding resources, given the specific structure of words" (p. 263, abstract).

The call for a theory of reading to consider cross-language differences is timely and well taken, but such differences need not imply that the origin of letter-order insensitivity is not universal. As noted earlier, models such as the overlap model and the noisy-channel model explain letter-order insensitivity as arising from the noisy perception of letter position, which is universal across languages. Consistent with this view, Kinoshita, Norris, and Siegelman (2012) showed that the same Hebrew stimuli used by Velan and Frost (2009) that produced no TL priming effect in lexical decision produce robust TL priming effects in the same-different task. This task dissociation—the finding of TL priming effect in the same-different task and its absence in lexical decision for Hebrew words—is explained within Norris and Kinoshita's (2008) account of masked priming described earlier in terms of the set of items the target is matched against in the two tasks: namely, the single referent item in the former and the words in the neighborhood of the target in the lexicon in the latter. The same account also provides a natural explanation of why TL priming effects in lexical decision are found in European languages but are absent in Hebrew (see Norris & Kinoshita, 2012b). In the Semitic morphology, a word is composed of a triconsonantal root that generally carries the meaning and a phonological word pattern in which it is embedded (e.g., the word *tizmoret* 'orchestra' consists of the root *Z.M.R.* and the phonological word pattern *ti- -o-et* with each dash indicating the position of a root consonant).[1] Because of this the lexical space in Hebrew is very dense, as transposition of two letters in a root will typically produce a different root, and hence another word. In English (and other European languages) transposing two adjacent letters will generally produce a nonword; that is, the closest word may be still the word the TL prime was derived from (as in *jugde-JUDGE*). Identifying words in Hebrew will therefore require readers to accumulate more evidence about letter order than in English.

What is suggested in the preceding discussion is that in reading, noisy perception interacts with the linguistic environment, so that during lexical access—that is, when attempting to find items in the lexicon that match the perceptual input—letter-order insensitivity phenomena arise in European languages where transposition neighbors are scarce. This view that perception interacts with the linguistic environment does not imply that "the 'perceptual system' *fully* completes its task, and only then does the 'linguistic system' come into play to produce differential effects of transposition" (p. 312), as Frost (2012; see also Velan et al., 2013) incorrectly attributed to Norris and Kinoshita (2012b). Elsewhere (e.g., Kinoshita & Norris, 2009, p. 14) we have been explicit that "the *evolving* [emphasis added here] prelexical orthographic representation"—in which the precise letter order is uncertain—is "the input to the lexical access process."

Kinoshita and Norris (2013) used the masked priming same-different task to test the core assumption of the OB models. The key tenet of OB models is that letter order in a word is coded by the presence of ordered letter pairs; that is, the letters in an OB must be in the right order (e.g., *CAT* does not contain the OB *TC*). A clear prediction of this assumption is that priming should be observed only for bigram primes that are in the correct order. Contrary to this prediction, robust priming was observed with reversed-order bigram primes (e.g., *fo-OF*). Reversed-order bigram primes also produced priming when the letters spanned three intervening letters in the target (e.g., *sb-ABOLISH*), thus ruling out the overlap OB model that incorporates the noisy-position assumption and hence allows the activation of reversed-order bigrams, but only if the letters are contiguous. These data challenge the core assumption of OB models that the tolerance to distortion of letter order arises from the coding of letter order in terms of the presence of ordered letter pairs.

A feature of OB models is that they postulate two levels of orthographic representations: OBs, and letters from which OBs are constructed. In the overlap model (Gomez et al., 2008) and the noisy-channel model (Norris & Kinoshita, 2012a), there is only one level of orthographic representation—letters. These latter models assume that the ambiguity in letter order originates in noisy perception of letter position. Gomez et al. (2008) have questioned the motivation for OBs thus: "In the open bigram models, there is accurate information about letter position, but this is discarded to produce a noisy representation of letter order in the form of OBs. One can ask why the system cannot access this accurate information about position" (p. 590). The extant masked priming data provide little support for the OB representation, and suggest instead that phenomena such as TL priming and relative priming effects originate in noisy perception of location of letters.

Conclusion

This chapter provided a selective review of recent research findings related to two issues that are fundamental to understanding the processes involved in visual word recognition—the role of word frequency and the representation of letter order—discussed within the Bayesian Reader framework. The Bayesian Reader approach differs from the activation framework that has been dominant in the visual word recognition research in at least two ways. First, it recognizes explicitly that the location information output by the visual perceptual system, which serves as the input to the word recognition process, is noisy. Second, the decision process involved in a task is seen as an integral part of visual word recognition, and the reader is viewed as making optimal decisions in accordance with Bayes' theorem based on noisy perceptual evidence. Within this framework, recent observations related to the role of word frequency—the logarithmic function relating word frequency to various measures of word recognition tasks, and its relationship to contextual diversity and subjective frequency—find a natural explanation in terms of word frequency serving as an index of prior probability. The Bayesian Reader framework also provides a coherent account of the empirical phenomena demonstrating the flexibility in letter-order coding, as seen in various masked priming effects such as the TL priming effect and the relative-position priming effects. As reviewed in this chapter, there is a growing body of evidence that these phenomena originate in the noisy perception of position of letters within a letter string. The Bayesian Reader provides a coherent account of these phenomena in terms of how the evolving orthographic representation interacts with the reader's lexicon.

Note

1 In the example I use the Latin alphabet letter having the same sound as the corresponding Hebrew letter, and use uppercase letters to indicate the root letters.

References

Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science, 17,* 814–823.

Adelman, J. S., Johnson, R. L., McCormick, S. F., McKague, M., Kinoshita, S., Bowers, J. S., ... Davis, C. J. (2014). A behavioral database for masked form priming. *Behavior Research Methods, 46,* 1052–1067

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* (Release 2). [CD-ROM]. Philadelphia, PA: University of Pennsylvania. Linguistic Data Consortium.

Balota, D. A. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. Flores-d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 9–32). Hillsdale, NJ: Erlbaum.

Balota, D., Cortese, M., Sergent-Marshall, S., Spieler, D., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283–316.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41,* 977–990.

Christianson, K., Johnson, R. L., & Rayner, K. (2005). Letter transposition within and across morphemes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1327–1339.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. C. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Davis, C. J. (2006). Orthographic input coding: A review of behavioral evidence. In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 180–206). Hove, England: Psychology Press.

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review, 117,* 713–758.

Davis, C. J., & Bowers, J. S. (2006). Contrasting five theories of letter position coding. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 535–557.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences, 9,* 335–341.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2007). Do transposed-letter similarity effects occur at a morpheme level? Evidence for morpho-orthographic decomposition. *Cognition, 105,* 691–703.

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods, 42,* 488–496.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 680–698.

Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology, 39,* 211–251.

Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences, 35,* 263–329.

García-Orza, J., Perea, M., & Muñoz, S. (2010). Are transposition effects specific to letters? *Quarterly Journal of Experimental Psychology, 63,* 1603–1618.

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review, 115,* 577–600.

Grainger, J., Granier, J. P., Farioli, F., van Assche, E., & van Heuven, W. J. (2006). Letter position information and printed word perception: The relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 865–884.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103,* 518–565.

Grainger, J., & van Heuven, W. J. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *The mental lexicon* (pp. 1–24). New York, NY: Nova Science.

Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? *Trends in Cognitive Sciences, 8,* 58–59.

Guerrera, C., & Forster, K. I. (2008). Masked form priming with extreme transposition. *Language and Cognitive Processes, 23,* 117–142.

Howes, D. H., & Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology, 41,* 401–410.

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology, 1,* 174.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British lexicon project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods, 44,* 287–304.

Kinoshita, S., & Norris, D. (2009). Transposed-letter priming of prelexical orthographic representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 1–18.

Kinoshita, S., & Norris, D. (2012). Task-dependent masked priming effects in visual word recognition. *Frontiers in Psychology, 3,* 178.

Kinoshita, S., & Norris, D. (2013). Letter order is not coded by open bigrams. *Journal of Memory and Language, 69,* 135–150.

Kinoshita, S., Norris, D., & Siegelman, N. (2012). Transposed-letter priming effect in Hebrew in the same-different task. *Quarterly Journal of Experimental Psychology, 65,* 1296–1305.

Kinoshita, S., Robidoux, S., Mills, L., & Norris, D. (2014). Visual similarity effects on masked priming. *Memory & Cognition, 42,* 821–833.

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Kuperman, V., & Van Dyke, J. A. (2013). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 802–823.

Logan, G. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review, 103,* 603–649.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, and Computers, 28,* 203–208.

Lupker, S. J., & Davis, C. J. (2009). Sandwich priming: A method for overcoming the limitations of masked

priming by reducing lexical competitor effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 618–639.
Lupker, S. J., & Pexman, P. (2010). Making things difficult in lexical decision: The impact of pseudohomophones and transposed-letter nonwords on frequency and semantic priming effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1267–1289.
McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.
Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review, 113,* 327–357.
Norris, D. (2009). Putting it all together: A unified account of word recognition and reaction-time distributions. *Psychological Review, 116,* 207–219.
Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology: General, 137,* 434–455.
Norris, D., & Kinoshita, S. (2012a). Reading through a noisy channel: Why there's nothing special about the perception of orthography. *Psychological Review, 119,* 517–545.
Norris, D., & Kinoshita, S. (2012b). Orthographic processing is universal: It's what you do with it that's different. Comment on Frost, R. Towards a universal model of reading. *Behavioral and Brain Sciences, 35,* 296–297.
Norris, D., Kinoshita, S., & van Casteren, M. (2010). A stimulus sampling theory of letter identity and order. *Journal of Memory and Language, 62,* 254–271.
Norris, D., & McQueen, J. M. (2009). Shortlist B: A Bayesian model of continuous speech perception. *Psychological Review, 115,* 357–395.
Perea, M., & Acha, J. (2009). Does letter position coding depend on consonant/vowel status? Evidence with the masked priming technique. *Acta Psychologica, 130,* 127–137.
Perea, M., & Lupker, S. J. (2003). Transposed-letter confusability effects in masked form priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art* (pp. 97–120). Hove, England: Psychology Press.
Peressotti, F., & Grainger, J. (1999). The role of letter identity and letter position in orthographic priming. *Perception and Psychophysics, 61,* 691–706.
Plummer, P., Perea, M., & Rayner, K. (2014). The influence of contextual diversity on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 275–283.
Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review, 88,* 552–572.
Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111,* 159–182.
Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 720–732.
Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157.
Rueckl, J. G., & Rimzhim, A. (2011). On the interaction of letter transpositions and morphemic boundaries. *Language and Cognitive Processes, 26,* 482–508.
Sánchez-Guitérrez, C., & Rastle, K. (2013). Letter transpositions within and across morphemic boundaries: Is there a cross-language difference? *Psychonomic Bulletin & Review, 20,* 988–996.
Schilling, H. E. H., Rayner, K., & Chumbley, J. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26,* 1270–1281.
Shillcock, R., Ellison, T. M., & Monaghan, P. (2000). Eye-fixation behavior, lexical storage, and visual word recognition in a split processing model. *Psychological Review, 107,* 824–851.
Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science, 8,* 411–416.
Stafford, T., & Gurney, K. N. (2004). The role of response mechanisms in determining reaction time performance: Piéron's law revisited. *Psychonomic Bulletin & Review, 11,* 975–987.
Stinchcombe, E., Lupker, S. J., & Davis, C. J. (2011). Transposed-letter priming effects with masked subset primes: A re-examination of the "relative position priming constraint." *Language and Cognitive Processes, 27,* 475–499.
Van Assche, E., & Grainger, J. (2006). A study of relative-position priming with superset primes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 399–415.
Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67,* 1176–1190.
Velan, H., Deutsch, A., & Frost, R. (2013). The flexibility of letter-position flexibility: Evidence from eye movements in reading Hebrew. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1143–1152.
Velan, H., & Frost, R. (2007). Cambridge University versus Hebrew University: The impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review, 14,* 913–918.
Velan, H., & Frost, R. (2009). Letter-transposition effects are not universal: The impact of transposing letters in Hebrew. *Journal of Memory and Language, 61,* 285–320.
Velan, H., & Frost, R. (2011). Words with and without internal structure: What determines the nature of orthographic and morphological processing? *Cognition, 118,* 141–156.
Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review, 8,* 221–243.
Whitney, C. (2008). Comparison of the SERIOL and SOLAR theories of letter-position coding. *Brain and Language, 107,* 170–178.
Whitney, C., Bertrand, D., & Grainger, J. (2012). On coding the position of letters in words: A test of two models. *Experimental Psychology, 59,* 109–114.

CHAPTER

6 Neighborhood Effects in Visual Word Recognition and Reading

Manuel Perea

Abstract

This chapter discusses research on how words that are orthographically (or phonologically) similar to a printed word influence the speed and accuracy of its encoding. The relevant set of words (the word's neighbors) was previously defined to be those lexical units differing from the target stimulus by a single letter in a given position. Recent evidence has revealed that a better definition of a word's neighborhood includes lexical units of different length and lexical units created by transpositions. The study of a word's neighborhood has revealed that that the activation of neighbors may interfere with the processing of the target words in word-identification tasks and during sentence reading, supporting the basic claims of interactive activation models. Some challenges to the current definitions of the sets of word neighborhoods are also examined, in particular the need to include differences between how consonants and vowels are encoded during word processing.

Key Words: word recognition, computational models, letter-position coding, consonant/vowel status, lexical decision, similarity metric

The examination of the nature of the underlying mechanisms that associate a printed word with its correct lexical unit (i.e., the process of lexical access) is one of the most basic issues in the research on reading. There are several reasons for this relevance. First, lexical access is a central component of sentence reading (Besner & Humphreys, 1991). Second, many reading disorders may originate from a deficient process of word identification (e.g., see Castles & Coltheart, 1993).

There is some agreement that when we identify a word in an alphabetic language, there is an early stage at which a number of similarly spelled lexical units to the printed stimulus (i.e., *neighbors*) are partially activated (or accessible). That is, during the process of visual word recognition there is a collection of lexical candidates that are similar (in some sense) to a given word and these candidates influence the ease with which the stimulus word is encoded or perceived. During the course of word processing, according to these models, these lexical candidates are progressively deactivated until only one lexical unit remains active (i.e., the perceived word) (e.g., search model: Murray & Forster, 2004; Bayesian Reader model: Norris, 2006; interactive-activation model: McClelland & Rumelhart, 1981, and its successors—multiple read-out model: Grainger & Jacobs, 1996; dual-route cascaded model: Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; spatial coding model: Davis, 2010). It is important to note here that there are other models that have a completely different metaphor for word identification, as in parallel distributed processing (PDP) models (Seidenberg & McClelland, 1989; see also Woollams, this volume).

The present chapter first examines the different metrics that have been proposed to define the set of a word's neighbors. It then examines whether a word's neighbors help or hinder during the process of word identification. Finally, it examines the limitations of current neighborhood metrics and suggests potential alternatives.

The Initial Definitions of the Set of a Word's Neighbors

One preliminary issue when studying the role of a word's neighbors during lexical access is to precisely define a word's neighborhood. If we use the characterization offered by Landauer and Streeter (1973, p. 120), "a similarity neighborhood will be defined as the set of words in the language from which a given stimulus word is indistinguishable after a specified loss of information about the stimulus word," then all lexical units in this set different from the target stimulus (specified by this criterion) are neighbors. Importantly, the influence of a word's neighbors on lexical access may be examined on two basic dimensions: (1) What is the influence of the number of the lexical units that compose the neighborhood (*neighborhood size*) on the processing of a given word? and (2) What is the influence of the frequency of the lexical units that compose the neighborhood (*neighborhood frequency*) on the processing of a given word?

The Initial Definitions of a Word's Neighborhood

In their seminal paper on word neighborhoods, Havens and Foote (1963) assumed that the set of a word's competitors (they used the term "competitors" rather than "neighbors") was composed of more frequent lexical units that shared all the letters but one with the target word, such that the differing letter should be an internal letter. Furthermore, this differing letter had to be visually similar to the original letter in terms of the ascending/neutral/descending pattern (e.g., *s* is a neutral letter, *d* is an ascending letter, and *p* is a descending letter). For instance, *list* would be a close neighbor of *lint*, as it is a high-frequency word that differs in a middle letter and the differing letter (*s*) has the same visual shape as the original letter (*n*; i.e., both are neutral letters). In contrast, lexical units such as *lift* and *line*, which also share three letters in the same position with *lint*, would not be close competitors. Havens and Foote (1963) did not provide a specific definition or weighting of the different grades of competitors, though.

Landauer and Streeter (1973) defined a word's neighborhood as the set of one-letter substitution neighbors. That is, two words are neighbors if they share the letters in all but one position. Albeit somewhat rudimentary, this is the definition that was later used in the influential study of Coltheart, Davelaar, Jonasson, and Besner (1977), and the number of orthographic neighbors of a word has commonly been termed *Coltheart's N*. Unlike Havens and Foote's (1963), this definition does not take into account letter shape. This assumption is consistent with claims that word identification is mediated by abstract (i.e., case-independent) letter units rather than by visual similarity (or letter shape), at least for skilled readers (see Rayner, McConkie, & Zola, 1980, for early evidence of the role of abstract letter/word representations in reading). In addition, the one-letter substitution characterization considers that all neighbors are equal regardless of whether the different letter is an external (i.e., initial or final) letter or a middle letter. This assumption was perhaps made more to keep the metrics as simple as possible rather than on the basis of empirical data. Indeed, there is evidence that shows that, as anticipated by Havens and Foote (1963), one-letter substitution neighbors that differ in an internal letter position are more perceptually similar to the target stimulus than lexical units that differ in an external letter position (see Perea, 1998). Both the Havens and Foote (1963) and Landauer and Streeter (1973) definitions also implicitly assume that the cognitive system initially encodes the number of letters of the target stimulus without noise, so that different-length lexical units (e.g., *house* and *hose*) are not part of the word's neighborhood. Furthermore, both Havens and Foote (1963) and Landauer and Streeter (1963) assume that, during word processing, letter-position coding operates without noise, so that *trail* would not be activated upon presentation of the word *trial* (i.e., they would not form part of the same neighborhood).

Thus, in the one-letter substitution definition from Landauer and Streeter (1973) and Coltheart et al. (1977), *clam* has *slam, cram, clad, clan, clap, claw*, and *clay* as orthographic neighbors. The idea here is that the number of orthographic neighbors (Coltheart's N = 7 in the case of *clam*) provides an initial index of the size (or density) of a word's neighborhood. That is, there are words with a large neighborhood (i.e., high-N words) such as *pale* (N = 20), and words with a small neighborhood (i.e., low-N words) such as *trek* (N = 1; its only neighbor is *tree*). Since the Coltheart et al. (1977) experiment, a number of experiments have examined the size of the orthographic neighborhood in a wide range of behavioral, eye-tracking, and neurophysiological paradigms (see Andrews, 1989, 1992, 1997). Another way to examine the role of a word's neighbors in visual word recognition has been to examine not the number of neighbors per se but their frequency (see Grainger, O'Regan, Jacobs, & Segui, 1989). This is in line with the initial proposal of Havens and Foote (1963) of considering the pivotal role of higher-frequency competitors during visual word recognition. In this line of research, the basic comparison is between a set of words with strong competitors (i.e., higher-frequency neighbors) and a set of words without strong competitors (i.e., no higher-frequency neighbors).

The First Expansion of the Neighborhood: Uncertainty in Letter-Position Coding

Research in the 1950s showed that participants could easily reproduce the base word upon the brief presentation of transposed-letter nonwords (e.g., *avitaion*; the base word is *aviation*) (see Bruner & O'Dowd, 1958), thus suggesting that nonwords created by transposing two letters resembled their original base words to a large degree. In more systematic research, Chambers (1979) and O'Connor and Forster (1981) examined the intricacies of letter-position encoding in chronometric word/nonword discrimination tasks (i.e., lexical decision), and found that transposed-letter nonwords like *mohter* produced a sizable number of "word" responses (more than 20%). More recently, Vergara-Martínez, Perea, Gomez, and Swaab (2013) found that a relatively late electrophysiological component such as the N400 (a peak that occurs around 400 ms after stimulus presentation) was evoked similarly for high-frequency words and for their transposed-letter counterparts (e.g., *mother* and *mohter*) in two visual word recognition tasks (lexical decision and semantic categorization), whereas this did not occur for replacement-letter nonwords (e.g., *mopher*). The fact that nonwords created from transposing adjacent letters (e.g., *mohter*) are highly wordlike suggests that the encoding of letter position and letter identity do not go hand in hand and that letter-position coding is quite flexible

The results just described pose problems for the orthographic coding scheme of the interactive-activation model (McClelland & Rumelhart, 1981). The model was initially implemented with a basic vocabulary of words of four letters, in which letter position was assumed to be processed in the correct order. In the model, *slat* and *salt* would be no closer than *slat* and *scar* (i.e., two letters in common in the two pairs), so that the presence of letter transposition effects rules out this orthographic coding scheme. Additional research from Andrews (1996) and a myriad of experiments in the past decade (see Perea & Lupker, 2003, 2004) have helped refine the ideas of how letter-position coding is attained during visual word recognition. As a result, a number of models with a more flexible coding scheme have been proposed (see Frost, this volume; Kinoshita, this volume). The important issue here is that transposed-letter neighbors are also activated during visual word recognition and reading (e.g., *trial* would influence the processing of *trail*), so that a word's neighborhood should also include these lexical units within the set of candidates (see Acha & Perea, 2008b; Johnson, 2009, for evidence of an inhibitory effect of transposed-letter neighbors during normal reading).

The Second Expansion of the Neighborhood: The Issue of Word Length

In a word identification task with masked stimuli, Grainger and Segui (1990) noted that a number of errors involved the addition or deletion of a letter (e.g., *votre* 'your' instead of *vote* 'vote'; *cuir* 'leather' instead of *cuire* 'to cook'). Grainger and Segui indicated that "competing units in the word-recognition process need not be of the same length" (p. 195), which echoed the research in auditory word recognition, where a word's neighborhood is typically defined in terms of words in which one phoneme is substituted, added, or removed (see Luce & Pisoni, 1998). Later, more systematic research provided evidence that the processing of a printed word may be affected by lexical units that differ in the number of letters: This is the case of addition-letter neighbors (the addition-letter neighbor *slate* may influence the processing of the target word *slat*; see Davis, Perea, & Acha, 2009; Davis & Taft, 2005; de Moor & Brysbaert, 2000) and deletion-letter neighbors (the deletion-letter neighbor *sat* may influence the processing of the target word *slat*; see Davis et al., 2009; Perea & Gomez, 2010).

Finally, syllable neighbors (i.e., lexical units that share a syllable in the same position with the target word, in particular the initial syllable) may also be activated. Syllable neighbors may be particularly relevant in those languages in which the syllable plays a major role in word identification (e.g., Spanish; see Carreiras, Álvarez, & de Vega, 1993; Perea & Carreiras, 1998). These neighbors may have the same length as the target word (e.g., *laca* and *lago* in Spanish) or may not (*laca* and *lavar*). In sum, a word's orthographic neighborhood can be composed of different types of neighbors: one-letter substitution neighbors (*slam* and six others for *slat*); transposed-letter neighbors (*salt*); addition-letter neighbors (*slate* and two others); deletion-letter neighbors (*sat*); and (possibly) syllabic neighbors.

One issue here is whether we can obtain a single, combined metric of a word's orthographic neighborhood. Davis (2005) proposed the use of *N** as the sum of all the one-letter substitution neighbors (i.e., Coltheart's N), transposed-letter neighbors, addition-letter neighbors, and deletion-letter neighbors. For example, in the case of *clam,* N* is 12 + 1 + 3 + 1 = 17. Although using Davis's combined set of neighbors as a metric may be considered a good initial approach, it is not free from shortcomings. One shortcoming is that it assigns the same weight to all types of neighbors, but there is evidence that some types of neighbors may be more equal than others (e.g., Davis et al., 2009; Duñabeitia, Perea, & Carreiras, 2009). Another shortcoming is that this measure tends to be zero (or close to zero) for relatively long words—as also occurs with Coltheart's N.

To overcome this latter limitation, a number of researchers have an alternate measure, *OLD20* (Yarkoni, Balota, & Yap, 2008). This measure is based on the Levenshtein distance between two words. The Levenshtein distance, a common measure in information theory, is the minimum number of single-letter changes (replacements, additions, deletions) required to change one word into the other. For instance, the Levenshtein distance between *hose* and (its one-letter substitution neighbor) *nose* is 1 (i.e., the replacement of *h* with *n*). Similarly, the Levenshtein distance between *hose* and (its addition-letter neighbor) *horse* is also 1 (i.e., addition of *r*). The OLD20 measure is defined as the mean distance, in terms of these single-letter changes, from each word relative to its 20 closest Levenshtein neighbors. Thus, as occurs with N*, the OLD20 is a measure of the size of the orthographic neighborhood rather than a measure of the frequency of its members. Two advantages of this measure over N* and N are the following: (1) While they all apply to long words, OLD20 is less likely to be 0; and (2) OLD20 allows a (more realistic) graded view of a word's neighbors (i.e., it is not only measuring whether two words are neighbors but also measuring how perceptually close two words are). Indeed, the OLD20 measure is rapidly becoming the most-employed neighborhood measure in research on visual word recognition (Grainger, Dufau, Montant, Ziegler, & Fagot, 2012; see Vergara-Martínez & Swaab, 2012, for electrophysiological evidence of OLD20 effects).

Despite its importance, there is one arguable limitation of the Levenshtein distance when describing a word's neighborhood: It weights single substitutions more heavily than letter transpositions. That is, in this metric, *train* and *trail* (Levenshtein distance 1; i.e., replacing *n* with *l*) are more closely related than *trial* and *trail* (Levenshtein distance 2; i.e., replacing *i* with *a* and *a* with *i*). However, there is evidence that transposed-letter neighbors have a special status within a word's neighborhood so that they are more closely related than substitution-letter neighbors (see Duñabeitia et al., 2009; Gómez et al., 2008; Perea & Fraga, 2006). Although this issue does not affect the majority of orthographic systems, where the number of transposed-letter word pairs is usually very small (see Acha & Perea, 2008b; Andrews, 1996), it may be a relevant factor in those languages with a large set of transposed-letter word pairs (e.g., Semitic languages like Arabic and Hebrew; see Perea, Abu Mallouh, & Carreiras, 2010; Velan & Frost, 2007).

In this section, I have focused on orthographic measures of a word's neighborhood. Parallel measures have been proposed for phonological neighborhoods. Following the logic of the Landauer and Streeter (1973) definition, Yates, Locker, and Simpson (2004) defined "phonological neighbors as words that could be formed by changing only one phoneme of the target word" (p. 453). That is, the POLD20 is analogous to the OLD20 measure except that it deals with phonological rather than orthographic neighbors. While in a number of languages orthographic and phonological neighbors typically coincide (e.g., in Spanish), this is not always the case (e.g., in English). Given that fewer studies have manipulated phonological neighborhoods than orthographic neighborhoods, the following section will focus primarily on the effects of orthographic neighborhoods. Nonetheless, recent experiments on the effects of phonological neighborhoods will be reviewed at the end of the section.

Do Neighbors Help or Hinder the Process of Word Identification?

The initial experiments on the role of a word's orthographic neighbors using response time tasks such as the lexical decision task (i.e., "is the stimulus a real word or not?") tested one basic assumption of the interactive activation model: the idea of competition at the lexical level via inhibitory links among the nodes that represented the lexical units. Specifically, in a lexical decision experiment, Grainger et al. (1989) compared the word identification times of a set of words with no higher-frequency (one-letter substitution) neighbors and a set of words with at least one higher-frequency (one-letter substitution) neighbor. Consistent with the predictions of the interactive-activation model, they found that words with a higher-frequency neighbor produced longer word identification times than the words with no higher-frequency neighbors. This finding is not limited to laboratory word identification tasks, as it has been also replicated and generalized to sentence reading. In particular, fixation times on words with higher-frequency neighbors are longer (and/or there are more regressions back to the target word) than the parallel measures for control words with no higher-frequency neighbors, and this has been reported using different types of neighbors: one-letter substitution neighbors (Perea & Pollatsek, 1998; Slattery, 2009), transposed-letter neighbors (Acha & Perea, 2008b; Johnson, 2009), addition-letter neighbors (Davis, Perea, & Acha, 2009), and deletion-letter neighbors (Davis et al., 2009). Importantly, the sentence reading experiments have revealed that the effects of these higher-frequency neighbors tend to occur in relatively late measures (i.e., once the reader has left the target word, such as the fixation duration following the target word or the percentage of regressions back to the target word) rather than in early fixation measures (e.g., the initial fixation on the target word). This outcome is consistent with models of eye movement control (e.g., E-Z-Reader model; see Reichle, Rayner, & Pollatsek, 2003; Reichle & Sheridan, this volume; see also Johnson, Staub, & Fleri, 2012, for evidence of transposed-letter neighborhood effects in word reading by using response time distributions).

The previous paragraph offered evidence in favor of competitive effects at the lexical level during visual word recognition. However, Andrews (1989, 1992) found that low-frequency words with many (one-letter substitution) neighbors produced faster latencies in the lexical decision task than low-frequency words with few (one-letter substitution) neighbors. This finding seems to be at odds with the existence of competition at the lexical level, since one would have expected that having many neighbors would lead to more lexical competition (via inhibitory links), which, in turn, would lead to longer word identification times for high-N words. Indeed, simulations with the interactive activation model cannot capture that pattern of effects (see Grainger & Jacobs, 1996).

To explain this apparent discrepancy, Grainger and Jacobs (1996) argued that the facilitative effect of the number of neighbors was due to task-specific factors in the lexical decision task. In particular, in their multiple read-out model, Grainger and Jacobs expanded the interactive-activation model so that a "word" response in lexical decision could be generated not only on the basis of a word unit reaching a given level of activation (i.e., the original criterion in the interactive activation model) but also on the basis of a global activity criterion on the basis of the summed activation of the orthographic neighbors (the so-called Σ-criterion). This new model was able to capture simultaneously the facilitative effect of number of neighbors and the inhibitory effect of neighborhood frequency that occurs in lexical decision (Grainger & Jacobs, 1996; but see Wagenmakers, Ratcliff, Gomez, & McKoon, 2008). One prediction from the Grainger and Jacobs (1996) model is that, when the same words that produce a facilitative effect of neighborhood size in lexical decision are employed in a situation in which actual word identification is required (e.g., in sentence reading), the effect should become inhibitory. This prediction was later confirmed by Pollatsek, Perea, and Binder (1999).

Taken together, the evidence described here is consistent with models that assume that there is lexical competition among the neighboring units activated upon word presentation, such as the interactive-activation model (or its successors). One limitation of the experiments that use two different sets of stimuli, such as those described here (e.g., words with higher-frequency neighbors vs. words with no higher-frequency neighbors; words from large neighborhoods vs. words from small neighborhoods) is that the control of some of the characteristics of the two sets of stimuli is not straightforward. One complementary way to examine the role of lexical competition during lexical access is to employ a priming procedure—in particular, *masked priming* (Forster & Davis, 1984; see also Grainger, 2008, for a review). The procedure of the masked priming technique is straightforward: after a 500-ms forward pattern mask (#####), the priming stimulus is presented briefly (around 30–50 ms), just prior to the target. Participants are required to make a response to the target stimulus (i.e., lexical decision, semantic categorization, or naming). Although the trace of the masked prime is (essentially) inaccessible to conscious report, the prime is capable of affecting the recognition of the target in lexical decision and other tasks. Unsurprisingly, the strongest positive priming effect is obtained when the prime is the same word as the target (faster response times to *house-HOUSE*, than to *ocean-HOUSE*), but masked priming effects also occur when the prime and target share an orthographic, phonological, morphological, or even a semantic relationship. The nature of masked priming is at an abstract level of representation, as masked priming effects are the same magnitude for pairs that are nominally and physically the same in lowercase and uppercase (e.g., *kiss-KISS*) and for pairs that are nominally (but not physically) the same in lowercase and uppercase (e.g., *edge-EDGE*) (see Bowers, Vigliocco, & Haan 1998; Perea, Jiménez, & Gómez, 2014). In contrast to single-word (or reading) experiments, the target materials in priming conditions (or in any other within-item manipulation) are held constant across the priming conditions. This avoids the problems of attempting to control for potential confounds in the selected stimuli (see Forster, 2000, for discussion), and it also allows for a within-item rather than a less powerful between-item analysis.

A number of masked priming experiments using the lexical decision task have provided converging evidence in favor of competition at the lexical level from a word's neighbors. Masked form priming effects on word targets usually differ depending on whether the prime stimulus is a word or not. Specifically, while the influence of word neighbor primes on target processing tends to be inhibitory, the influence of nonword neighbor primes tends to be facilitative (e.g., Carreiras & Perea, 2002; Davis & Lupker, 2006; Duñabeitia et al., 2009; Nakayama, Sears, & Lupker, 2011; Segui & Grainger, 1990). This applies to experiments using one-letter substitution neighbors and transposed-letter neighbors. This outcome fits quite well with the idea that a neighboring word prime exerts an inhibitory influence on the processing of the target word (via inhibitory links at the lexical level), while a neighboring nonword prime produces sublexical facilitation.

Further evidence that a word's neighbors may hinder its processing comes from the interaction between masked form/repetition priming and neighborhood density. Consider the effect of form priming with nonword primes (e.g., *honse-HOUSE* vs. *minve-HOUSE*). While form priming occurs for target words with few neighbors (low-N words), it is absent for target words with many neighbors (high-N words; Forster, Davis, Schoknecht, & Carter 1987; see also Perea & Rosa, 2000). Furthermore, a parallel effect occurs in repetition priming. The magnitude of masked repetition priming is larger for low-N words than for high-N words (Perea & Rosa, 2000). That is, high-N words benefit less from a repeated presentation of the same word than low-N words. In a series of three masked form/repetition priming lexical decision experiments testing with three different sets of stimuli varying in overall word frequency, Perea and Forster (in preparation) found that repetition priming in English was greater for low-N words (47, 49, and 52 ms) than for high-N words (31, 34, and 34 ms). Likewise, low-N words showed a significant form priming effect of approximately 25–30 ms in the three experiments (26, 25, and 29 ms), which is a bit less than half the size of the repetition priming effects for these same items—the form priming effect for high-N words was negligible across the three experiments (-2, 2, and 7 ms). Since the prime duration was 50 ms in these masked priming experiments, this means that low-N words, but not high-N words, obtained full benefit from the identity prime (i.e., a presumed advantage of around 50 ms of the identity over the unrelated priming condition; see Gomez, Perea, & Ratcliff, 2013, for a discussion of the nature of masked repetition priming effects) via inhibitory links from the preactivated lexical units in large neighborhoods. The basic conclusion from these experiments is that a high-N word receives less processing benefit from its previous masked presentation than a low-N word. Therefore, the modulation of masked form/repetition priming provides converging evidence in favor of those models that assume that there is competition at the lexical level.

The previous paragraphs focused on the impact of orthographic neighbors in visual word recognition. A less studied issue has been the impact of phonological neighbors in visual word recognition. Indeed, most of the current (implemented) models of visual word recognition focus on the orthographic level of processing (e.g., spatial coding model, Davis, 2010). However, several studies have examined the influence of a word's phonological neighbors in visual word recognition and reading while controlling for the word's orthographic neighbors. Yates et al. (2004) reported that words with many phonological neighbors were responded to faster in a (visual) lexical decision task than the words with few phonological neighbors. Subsequently, Yates, Friend, and Ploetz (2008) examined whether this facilitative effect could be generalized to a normal reading situation.

In particular, Yates et al. conducted a sentence reading experiment in which a target word (with many/few phonological neighbors) was embedded in each sentence. The results were mixed. While they found shorter first-fixation times on the target words with many phonological neighbors than on the words with few phonological neighbors, this facilitative effect vanished in other eye movement measures such as gaze durations (i.e., the sum of all fixations on the target word before leaving it) and total fixation times (i.e., the sum of all fixations on the target word including regressive fixations). Thus, while there were some hints that at some processing level, a word's phonological neighbors may have had a facilitative influence on the target word, the evidence was not decisive. Clearly, an important topic for future research is to examine in detail the impact of both orthographic and phonological neighbors in visual word recognition and reading across a range of languages.

Does the Consonant/Vowel Status Matter in a Word's Neighborhood?

A neglected issue in most studies on neighborhood effects is the distinction between consonants and vowels. The reason is that most influential models of visual word recognition assume that there is no distinction between the consonant/vowel status of printed letters (e.g., interactive-activation model, McClelland & Rumelhart, 1981; spatial coding model, Davis, 2010; Bayesian Reader model, Norris, 2006; open bigram model, Grainger & van Heuven, 2003; SERIOL model, Whitney, 2001; overlap model, Gomez et al., 2008). Therefore, in these models, neighbors that differ in one vowel such as *list* and *lost* are perceptually as close as neighbors that differ in one consonant such as *list* and *lift*.

However, a large body of research has revealed that, in various languages, consonants and vowels are not processed in exactly the same way (see Caramazza, Chialant, Capasso, & Miceli, 2000; Mehler, Peña, Nespor, & Bonatti, 2006). In particular, it has been claimed that consonants are more relevant than vowels for access to the mental lexicon, whereas vowels are more relevant for conveying grammatical information (Mehler et al., 2006). Indeed, when using shortcuts in text messages, we tend to omit the vowels rather than the consonants and the resulting words can be easily reproduced (see Perea, Acha, & Carreiras, 2009, for eye-tracking evidence). With respect to the specific issue of consonants and vowels and orthographic neighborhoods, an important piece of evidence is the masked priming lexical decision experiment of New, Araujo, and Nazzi (2008). The two critical priming conditions were a consonant-preserving condition (e.g., *duvo-DIVA; apis-OPUS*) and a vowel-preserving condition (e.g., *rifa-DIVA; onub-OPUS*). For adult readers, consonant-preserving primes facilitated target processing to a larger degree than vowel-preserving primes. Indeed, the response times in the vowel-preserving priming condition did not differ significantly from those of an unrelated priming condition (e.g., *rufo-DIVA; anib-OPUS*). In a recent series of experiments, Soares, Perea, and Comesaña (2014) replicated the New et al. finding in another language (Portuguese) with adult readers and also extended the finding of a consonant/vowel difference to developing readers (fifth-grade children).

Another piece of information relevant to the importance of the consonant/vowel status of letters comes from the masked priming lexical decision experiments with nonword partial primes conducted by Duñabeitia and Carreiras (2011). They found that partial primes composed of consonants were more effective than partial primes composed of vowels (i.e., faster response times to *csn-CASINO* than to *aio-CASINO*). Furthermore, letter transposition effects differed for consonant and vowel transpositions: *caniso* and *casino* are perceptually closer than *anamil* and *animal*, as deduced from the fact that a target word like *CASINO* is identified more rapidly when preceded by the transposed-letter nonword *caniso* than when preceded by the replacement-letter nonword *caviro*, whereas the parallel difference is absent for the transposition/replacement of two vowels (i.e., similar word identification times for *anamil-ANIMAL* and *anomel-ANIMAL*; Perea & Lupker, 2004; see also Lupker, Perea, & Davis, 2008).

Therefore, the current measures of a word's neighborhood should be expanded to account for the consonant/vowel distinction. As stated earlier, current computational models of visual word recognition do not account for these consonant/vowel differences. One straightforward option would be to give a differential weight to consonantal modifications in OLD20 distance metrics. With the advent of big databases of identification times for thousands of words in different languages (e.g., Balota et al., 2007), it should be easy to test whether a modified OLD20 (or POLD20) measure that weights changes in consonants and vowels differently offers better fits than the current OLD20 measure. At the same time, it may be important to examine whether assigning higher weights to external letters than to internal letters may also provide a better fit. In turn, between the external letters, the beginning letter may also be assigned higher weights than the end letter.

Conclusions and Future Directions

Experimentation on the impact of a word's neighborhood during lexical access in laboratory word identification tasks (either in single-presentation or masked priming paradigms) and in silent reading (via eye tracking) has provided evidence of competition at the lexical level, thus providing empirical support to the claims of interactive-activation models. Despite the limitations of neighborhood metrics, the basic findings that were obtained with the one-letter substitution neighbors in the 1970s and 1980s have been extended—with appropriate adjustments—to other types of neighbors.

One final issue that deserves some comment is to what degree a word's neighborhood during reading is influenced by properties of the visual-attentional system that were ignored in the models that were discussed. As indicated earlier, the OLD20 measure has the limitation that letter transpositions involve two steps while a single addition, deletion, or replacement only involve one step, and evidence reveals that transposed-letter neighbors are very close to the target word (i.e., closer than one-letter substitution neighbors). This phenomenon may be related to how the visual system encodes letter position: Perceptual uncertainty regarding letter position has been posited to originate from noise in encoding position at the visual level (Gomez et al., 2008). As such, it also appears when coding sequences of geometrical objects (García-Orza, Perea, & Estudillo, 2011) and when reading musical notes in a staff (Perea, García-Chamorro, Centelles, & Jiménez, 2013). Indeed, when the same materials that in the visual modality produce a transposed-letter effect (e.g., *cholocate* is error-prone when presented visually) are presented in a tactile modality such as Braille, the transposed-letter effect vanishes (Perea, García-Chamorro, Martín-Suesta, & Gómez, 2012). Therefore, research in modalities other than the visual, such as research in Braille, may be informative to find out which aspects of the reading process, including the definition of a word's neighbors, are modality-independent and which aspects are modality-specific (see Perea, Jiménez, Martín-Suesta, & Gómez, 2014 for a comparison of sentence reading in sighted vs. Braille readers).

An important issue for further research is how a word's neighborhood evolves in developing readers. Castles, Davis, Cavalot, and Forster (2007; see also Acha & Perea, 2008a; Soares et al., 2014) have claimed that the organization of the neighborhood varies as a function of reading skill across primary school children. In their *lexical tuning* hypothesis, Castles et al. indicated that the orthographic recognition system is initially coarsely tuned and that it becomes more and more precise with increased reading skill. Consistent with this view, Castles et al. (2007) found large effects of masked form priming close in size to those of identity priming in beginning readers. In older children they found the expected advantage in effect size of identity priming over form priming that occurs in adult readers. Therefore, the definition of a word's neighborhood in children may reflect more flexible coding of letter identity and letter position. More research should examine in detail the relationship between reading level and word neighborhoods.

Another relevant issue is how a word's neighborhood is affected by the existence of two (or multiple) lexicons in bilinguals. There is evidence that, in bilinguals, presentation of a word activates similarly spelled words in the bilingual's two languages, as predicted by interactive-activation models. In particular, the bilingual activation model (Dijkstra, van Heuven, & Grainger, 1998) can successfully deal with many of the intricacies of bilingual word recognition (see Grainger, Midgley, & Holcomb, 2010 for a review of recent research).

Most of the research summarized in this chapter has been conducted in languages that employ the Latin script. In languages that employ the second most widely used alphabetic script in the world, Arabic (e.g., Arabic, Persian, Urdu, Uyghur), the specific shape of each letter form depends on whether it is connected to the neighboring letters. Arabic is a semicursive script that is read from right to left in which, for instance, the shape of the letter *nūn* (*n* in the Buckwalter transliteration) differs depending on whether it is connected to both contiguous letters (middle form: ن), when it is only connected to the previous letter (initial form: ن), when it is only connected to the following letter (final form: ن), and when it is not connected to the neighboring letters (isolated form: ن). While some letters in Arabic can connect with the following letter, others cannot, thus potentially creating graphemic chunks, as in the word شراع ('sail', *$rAE* in the Buckwalter transliteration; $ = /ʃ/, r = /r/, A = /aː/, and E = /ʕ/ in IPA notation), in which the two initial letters are connected, and the two final letters are isolated—note that, as indicated earlier, Arabic is read from right to left. The position-dependent allography of the words' constituent letters in Arabic script influences the structure of a word's neighborhood (see Friedmann & Haddad-Hanna, 2012, for evidence in Arabic; see also Yakup, Abliz, Sereno, & Perea, 2014 for evidence in Uyghur). These two studies revealed that the words لهمت ('slowed', *tmhl* in the Buckwalter transliteration) and لمهت ('neglect', *thml*), which share the letter-position allographs (note that the transliterations of the phonemes *m* [ـمـ] and *h* [ـهـ] are both in their middle form positions in the two words) are orthographically closer than the words عارش ('sail', transliterated as *$rAE*) and عراش ('street', *$ArE*) that do not share the letter-position allographs (the transliteration of the phoneme $ArE is in isolated form in *$rAE* [ا] and final form in *$ArE* [ا], whereas the transliteration of the phoneme $ArE is in final form in *$rAE* [ر] and in isolated form in *$ArE* [ر]).

Further research should also examine how a word's neighborhood is characterized in alphabetical languages that employ tones as markers, such as Thai (see Winskel & Perea, 2014, for an examination of orthographic/phonological effects of tone markers in Thai). Importantly, the nature of Thai orthography, in which words are not separated by blank spaces, may also lead to letter-coding processes that differ from those in Indo-European languages (see Winskel, Perea, & Peart, 2014). For example, during sentence reading, the degree of disruption of reading transposed-letter nonwords is similar for internal and initial transposed-letter nonwords in Thai (Winskel, Perea, & Ratitamkul, 2012), whereas it is more disruptive for the initial letter position than for internal positions in Indo-European languages (see White, Johnson, Liversedge, & Rayner, 2008).

Finally, it seems likely that there is some sort of neighborhood effect (or effects) in nonalphabetic languages like Chinese and Japanese. That is, because research in alphabetic languages has shown that there are clear inhibitory effects in encoding words coming from competing similar words that are neighbors, it seems likely that there will be similar effects in nonalphabetic languages. This means that an important question is how a word's neighborhood can best be defined in these nonalphabetic languages. Consider the case of Chinese. There is the complex issue in Chinese of what a word is, and there is far from universal agreement as to which two- to four-letter sequences of Chinese characters are words. As a result, research on Chinese neighborhood effects has begun by exploring neighborhood effects in Chinese characters. Given that the majority of Chinese characters can be decomposed into a semantic radical that provides a clue to meaning and a phonetic radical that provides a clue to pronunciation, a preliminary way to examine a character's neighborhood in Chinese is by separating phonetic radical neighborhoods (i.e., characters that share the phonetic radical) and semantic radical neighborhoods (i.e., those that share the semantic radical; see Li, Bi, Wei, & Chen, 2011 for recent research on phonetic radical neighborhoods). An alternative way to define a word's neighborhoods in Chinese is by taking into account similarity at the stroke level. In particular, Wang, Jing, Weijin, Liversedge, and Paterson (2014) defined stroke neighbors in Chinese as characters that could be formed by substituting, adding, or deleting one or more character strokes. Their rationale was that a character's strokes could be considered analogous to letters in words, whereas radicals could be considered more analogous to morphemes. Importantly, Wang et al. (2014) found an inhibitory stroke neighborhood effect in masked priming and normal reading. Thus, these data paralleled the effects reported in alphabetic languages (see also Nakayama et al., 2011, for a similar finding in the Japanese syllabic script Kana). Although further research is needed to establish firm conclusions regarding the nature of lexical competition during visual word recognition, the data so far from Chinese and Japanese suggests that these processes may be common across alphabetic and nonalphabetic languages.

Acknowledgment

Preparation of this chapter was partially supported by Grant PSI2011-26924 from the Spanish Government.

References

Acha, J., & Perea, M. (2008a). The effects of length and transposed-letter similarity in lexical decision: Evidence with beginning, intermediate, and adult readers. *British Journal of Psychology, 99,* 245–264.

Acha, J., & Perea, M. (2008b). The effect of neighborhood frequency in reading: Evidence with transposed-letter neighbors. *Cognition, 108,* 290–300.

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 802–814.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 234–254.

Andrews, S. (1996). Lexical retrieval and selection processes: Effects of transposed-letter confusability. *Journal of Memory and Language, 35,* 775–800.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4,* 439–461.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459.

Besner, D., & Humphreys, G. W. (1991). Basic processes in word recognition and identification: An overview. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 1–9). Hillsdale, NJ: Erlbaum.

Bowers, J. S., Vigliocco, G., & Haan, R. (1998). Orthographic, phonological, and articulatory contributions to masked letter and word priming. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1705–1719.

Bruner, J. S., & O'Dowd, D. (1958). A note on the informativeness of parts of words. *Language and Speech, 1,* 98–101.

Caramazza, A., Chialant, D., Capasso, R., & Miceli, G. (2000). Separable processing of consonants and vowels. *Nature, 403,* 428–430.

Carreiras, M., & Perea, M. (2002). Masked priming effects with syllabic neighbors in the lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1228–1242.

Carreiras, M., Alvarez, C. J., & de Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language, 32,* 766–780.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47,* 149–180.

Castles, A., Davis, C., Cavalot, P., & Forster, K. I. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked form priming and transposed-letter effects. *Journal of Experimental Child Psychology, 97,* 165–182.

Chambers, S. M. (1979). Letter and order information in lexical access. *Journal of Verbal Learning and Verbal Behavior, 18,* 225–241.

Coltheart, M., Davelaar, E., Jonasson, J., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance* (Vol. 6, pp. 535–555). Hillsdale, NJ: Erlbaum.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods, 37,* 65–70.

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review, 117,* 713–758.

Davis, C. J., & Lupker, S. J. (2006). Masked inhibitory priming in English: Evidence for lexical inhibition. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 668–687.

Davis, C. J., & Taft, M. (2005). More words in the neighborhood: Interference in lexical decision due to deletion neighbors. *Psychonomic Bulletin & Review, 12,* 904–910.

Davis, C. J., Perea, M., & Acha, J. (2009). Re(de)fining the orthographic neighbourhood: The role of addition and deletion neighbours in lexical decision and reading. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1550–1570.

De Moor, W., & Brysbaert, M. (2000). Neighborhood-frequency effects when primes and targets are of different lengths. *Psychological Research, 63,* 159–162.

Dijkstra, T., van Heuven, W. J. B., & Grainger, J. (1998). Simulating cross-language competition with the bilingual interactive activation model. *Psychologica Belgica, 38,* 177–196.

Duñabeitia, J. A., & Carreiras, M. (2011). The relative position priming effect depends on whether letters are vowels or consonants. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 1143–1163.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2009). There is no clam with coats in the calm coast: Delimiting the transposed-letter priming effect. *Quarterly Journal of Experimental Psychology, 62,* 1930–1947.

Forster, K. I. (2000). The potential for experimenter bias effects in word recognition experiments. *Memory & Cognition, 28,* 1109–1115.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 680–698.

Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology, 39,* 211–251.

Friedmann, N., & Haddad-Hanna, M. (2012). Letter position dyslexia in Arabic: From form to position. *Behavioural Neurology, 25,* 193–203.

García-Orza, J., Perea, M., & Estudillo, A. (2011). Masked transposition effects for simple vs. complex non-alphanumeric objects. *Attention, Perception, & Psychophysics, 73,* 2573–2582.

Gomez, P., Perea, M., & Ratcliff, R. (2013). A diffusion model account of masked vs. unmasked priming: Are they qualitatively different? *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1731–1740.

Gómez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review, 115,* 577–601.

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes, 23,* 1–35.

Grainger, J., Dufau, S., Montant, M., Ziegler, J. C., & Fagot, J. (2012). Orthographic processing in baboons (*Papio papio*). *Science, 336,* 249–255.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103,* 518–565.

Grainger, J. Midgley, K. J., & Holcomb, P. J. (2010). Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). In M. Kail & M. Hickman (Eds.), *Language acquisition across linguistic and cognitive systems* (pp. 267–284). Philadelphia, PA: Benjamins.

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. *Perception & Psychophysics, 45,* 189–195.

Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Perception & Psychophysics, 47,* 191–198.

Grainger, J., & van Heuven, W. J. B. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *Mental lexicon: Some words to talk about words* (pp. 1–23). Hauppauge, NY: Nova Science.

Havens, L. L., & Foote, W. E. (1963). The effect of competition on visual duration threshold and its independence of stimulus frequency. *Journal of Experimental Psychology, 65,* 6–11.

Johnson, R. L. (2009). The quiet clam is quite calm: Transposed-letter neighborhood effects on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 943–969.

Johnson, R. L., Staub, A., & Fleri, A. (2012). Distributional analysis of the transposed-letter neighborhood effect on naming latency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 1773–1779.

Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior, 12,* 119–131.

Li, Q. L., Bi, H. Y., Wei, T. Q., & Chen, B. G. (2011). Orthographic neighborhood size effect in Chinese character naming: Orthographic and phonological activations. *Acta Psychologica, 13,* 35–41.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19,* 1–36.

Lupker, S. J., Perea, M., & Davis, C. J. (2008). Transposed letter priming effects: Consonants, vowels and letter frequency. *Language and Cognitive Processes, 23,* 93–116.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88,* 375–407.

Mehler, J., Peña, M., Nespor, M., & Bonatti, L. (2006). The "soul" of language does not use statistics: Reflections on vowels and consonants. *Cortex, 42,* 846–854.

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review, 111,* 721–756.

Nakayama, M., Sears, C. R., & Lupker, S. J. (2011). Lexical competition in a non-Roman, syllabic script: An inhibitory neighbor priming effect in Japanese Katakana. *Language and Cognitive Processes, 26,* 1136–1160.

New, B., Araujo, V., & Nazzi, T. (2008). Differential processing of consonants and vowels in lexical access through reading. *Psychological Science, 19,* 1223–1227.

Norris, D. (2006). The Bayesian Reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review, 113,* 327–357.

O'Connor, R. E., & Forster, K. I. (1981). Criterion bias and search sequence bias in word recognition. *Memory & Cognition, 9,* 78–92.

Perea, M. (1998). Orthographic neighbours are not all equal: Evidence using an identification technique. *Language and Cognitive Processes, 13,* 77–90.

Perea, M., & Fraga, I. (2006). Transposed-letter and laterality effects in lexical decision. *Brain and Language, 97,* 102–109.

Perea, M., & Gomez, P. (2010). Does LGHT prime DARK? Masked associative priming with addition neighbors. *Memory & Cognition, 38,* 513–518.

Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter confusability effects in masked associative priming. *Memory & Cognition, 31,* 829–841.

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language, 51,* 231–246.

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 767–777.

Perea, M., & Rosa, E. (2000). Repetition and form priming interact with neighborhood density at a short stimulus-onset asynchrony. *Psychonomic Bulletin & Review, 7,* 668–677.

Perea, M., Abu Mallouh, R., & Carreiras, M. (2010). The search of an input coding scheme: Transposed-letter priming in Arabic. *Psychonomic Bulletin & Review, 17,* 375–380.

Perea, M., Acha, J., & Carreiras, M. (2009). Eye movements when reading text messaging (txt msgng). *Quarterly Journal of Experimental Psychology, 62,* 1560–1567.

Perea, M., & Forster, K. I. (in preparation). The neighborhood density constraint on form/repetition priming.

Perea, M., García-Chamorro, C., Centelles, A., & Jiménez, M. (2013). Position coding effects in a 2D scenario: The case of musical notation. *Acta Psychologica, 143,* 292–297.

Perea, M., García-Chamorro, C., Martín-Suesta, M., & Gómez, P. (2012). Letter position coding across modalities: The case of Braille readers. *PLoS ONE, 7,* e45636.

Perea, M., Jiménez, M., Martín-Suesta, M., & Gómez, P. (2014). Letter position coding across modalities: Braille and sighted reading of sentences with jumbled words. *Psychonomic Bulletin & Review*.

Pollatsek, A., Perea, M., & Binder, K. (1999). The effects of neighborhood size in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1142–1158.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology, 12,* 206–226.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The EZ Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26,* 445–476.

Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 65–76.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Slattery, T. J. (2009). Word misperception, the neighbor frequency effect, and the role of sentence context: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1969–1975.

Soares, A. P., Perea, M., & Comesaña, M. (2014). Tracking the emergence of the consonant bias in visual-word recognition: Evidence with developing readers. *PLoS ONE, 9,* e88580.

Velan, H., & Frost, R. (2007). Cambridge University vs. Hebrew University: The impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review, 14,* 913–928.

Vergara-Martínez, M., Perea, M., Gomez, P., & Swaab, T. Y. (2013). ERP correlates of letter identity and letter position are modulated by lexical frequency. *Brain and Language, 125,* 11–27.

Vergara-Martínez, M., & Swaab, T. Y. (2012). Orthographic neighborhood effects as a function of word frequency: An event-related potential study. *Psychophysiology, 49,* 1277–1289.

Wagenmakers, E. J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the

lexical decision task. *Journal of Memory and Language, 58,* 140–159.
Wang, J., Tian, J., Han, W., Liversedge, S. P., & Paterson, K. B. (2014). Inhibitory stroke neighbour priming in character recognition and reading in Chinese. *Quarterly Journal of Experimental Psychology*.
White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word-beginning letters. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1261–1276.
Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review, 8,* 221–243.
Winskel, H., & Perea, M. (2014). Does tonal information affect the early stages of visual-word processing in Thai? *Quarterly Journal of Experimental Psychology, 67,* 209–219.
Winskel, H., Perea, M., & Peart, E. (2014). Testing the flexibility of the modified receptive field (MRF) theory: Evidence from an unspaced orthography (Thai). *Acta Psychologica, 150,* 55–60.
Winskel, H., Perea, M., & Ratitamkul, T. (2012). On the flexibility of letter position coding during lexical processing: Evidence from eye movements when reading Thai. *Quarterly Journal of Experimental Psychology, 64,* 1522–1536.
Yakup, M., Abliz, W., Sereno, J., & Perea, M. (2014). How is letter position coding attained in scripts with position-dependent allography? *Psychonomic Bulletin & Review*.
Yarkoni, T., Balota, D. A., & Yap, M. J. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15,* 971–979.
Yates, M., Friend, J., & Ploetz, D. M. (2008). The effect of phonological neighborhood density on eye movements during reading. *Cognition, 107,* 685–692.
Yates, M., Locker, L., & Simpson, G. B. (2004). The influence of phonological neighborhood on visual word perception. *Psychonomic Bulletin & Review, 11,* 452–457.

CHAPTER

7

Cross-Linguistic Perspectives on Letter-Order Processing: Empirical Findings and Theoretical Considerations

Ram Frost

Abstract

The processing of letter order has profound implications for understanding how visually presented words are processed and how they are recognized, given the lexical architecture that characterizes a given language. Research conducted in different writing systems suggests that letter-position effects, such as transposed-letter priming, are not universal. The cognitive system may perform very different types of processing on a sequence of letters depending on factors that are unrelated to peripheral orthographic characteristics but related to the deep structural properties of the printed stimuli. Assuming that identical neurobiological constraints govern reading performance in any language, these findings suggest that neurobiological constraints interact with the idiosyncratic statistical properties of a given writing system to determine the preciseness or fuzziness of letter-position coding. This chapter reviews the evidence for this interaction and discusses the implications for theories of reading and for modeling visual word recognition.

Key Words: visual word recognition, transposed-letter effect, letter-position coding, morphological processing, learning models

The processing of letter order in visual word recognition has become in the last decade the focus of heated debates, extensive research, and formidable modeling efforts. Whereas the first three decades of reading research centered on providing a general framework of lexical architecture (e.g., Forster, 1976; McClelland & Rumelhart, 1981; Morton, 1969) and mapping the processing of orthographic, phonological, semantic, and morphological information (e.g., Frost, 1998; Grainger & Ferrand, 1994; Marslen-Wilson, Tyler, Waksler, & Older, 1994), the last decade has seen an increased preoccupation with the front end of visual word recognition—mainly, the coding of letter order. This research effort has been largely driven by consistent findings showing that readers are surprisingly tolerant of letter transpositions, so that they are only slightly affected by manipulation of letter order in terms of speed and reading accuracy. This finding has profound implications for understanding how visually presented words are processed and how they are recognized. The present chapter examines letter order effects across different writing systems. As a first step, the evidence regarding the coding of letter-position in various orthographies will be outlined. This evidence will then be discussed in terms of its theoretical implications for modeling visual word recognition and for understanding reading.

Transposed-Letter Effects

A large set of experimental findings demonstrates that readers are relatively resilient when it comes to the jumbling of letters within words. Apparently, the original demonstration of the effect of letter transpositions on skilled readers (or lack thereof) belongs to Rawlinson (1976; see Rawlinson, 1999, for a reference to it). In an unpublished dissertation, he showed that letter randomization in the middle of words had little effect on the ability of skilled readers to understand the printed text. The first published investigation of this phenomenon was the study by Forster, Davis, Schoknecht, and Carter (1987), who showed that letter transposition in a briefly presented masked prime (*anwser*) results in significant facilitation in recognizing the target (*ANSWER*). Moreover, Forster et al. (1987) observed that transposed-letter (TL) primes (in most cases, primes with two adjacent middle letters of the target that are transposed) produced priming as large as identity primes (*answer-ANSWER*). This finding suggested that at least in the initial stage of visual word recognition, exact letter order is not registered by the cognitive system. However, perhaps the most popular demonstration of how reading is immune to letter transposition is the following paragraph that made its way around the Internet:

> Aoccdrnig to rscheearch at Cmabrigde Uinervtisy, it deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a total mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe.

This text does not refer to any true research project that was conducted at the University of Cambridge, and probably the author who drafted the preceding paragraph chose the name "Cambridge University" at random. Since its first appearance in 2003, the original English text has been translated into dozens of languages, demonstrating how reading is universally resilient to jumbled letters. The name "Cambridge" stuck, and the phenomenon of being able to read words with jumbled letters without much effort is thus often labeled "the Cambridge University" effect.

The manner in which readers treat the constituent letters of printed words subsequently became the focus of extensive systematic research in a variety of European languages such as English (e.g., Perea & Lupker, 2003), French (Schoonbaert & Grainger, 2004), Spanish (Perea & Carreiras, 2006a, 2006b; Perea & Lupker, 2004), and Basque (Duñabeitia, Perea, & Carreiras, 2007), and also in Japanese Kana, where characters represent syllables rather than phonemes (Perea & Perez, 2009). In general, these studies have shown that TL nonword primes (*caniso*) facilitate recognition of word targets (*CASINO*) relative to nonword primes with substituted letters (*cabino*) regardless of orthographic letter similarity. The facilitation caused by TL primes was shown even when the order of several letters within the word is compromised, distorting orthographic structure substantially (*snawdcih-SANDWICH*; Guerrera & Forster, 2008). The evidence from studies using TL primes converges with other forms of priming where absolute letter order is not maintained. For example, primes consisting of a subset of the target's constituent letters, in which the relative but not the absolute position of letters is maintained (*blck-BLACK*), produce significant priming (Grainger et al., 2006; Humphreys, Evett, & Quinlan, 1990; Peressotti & Grainger, 1999). Similar effects have been demonstrated with superset priming, where primes contain more letters than the target (*juastice-JUSTICE*; Van Assche & Grainger, 2006).

Research on eye movements has shown that letter transpositions result in some cost in terms of fixation time-measures on target words during reading (Johnson et al., 2007; Rayner et al., 2006). This cost, however, seems relatively small in magnitude (about 11 ms). Using rapid serial visual presentation (RSVP), Velan and Frost (2007, 2011) have shown that subjects' ability to detect words that had letter transpositions embedded in English sentences was particularly low (d' = 0.86, d' is a measure of sensitivity for detecting a signal in noise, where d' = 0 reflects chance level performance), and about one-third of the subjects were at chance level in perceiving even one of three transpositions in the sentence.

Several studies have examined whether the relative insensitivity to letter position is modulated by morphological (e.g., Christianson, Johnson, & Rayner, 2005; Duñabeitia et al. 2007; Duñabeitia, Perea, & Carreiras, 2014), or phonological factors (e.g., Acha & Perea 2010; Perea & Carreiras 2006a, 2006b). The results are mixed. For example, whereas some studies report reduced TL priming effects while crossing morphemic boundaries (*faremr-FARMER*), other studies show identical TL effects when morphemic boundaries are crossed (e.g., Beyersmann, McCormick, & Rastle, 2013; Sanchez-Gutierrez & Rastle, 2013; see Duñabeitia et al., 2014, for a discussion). Similarly, some studies examined the interaction of letter-position coding with phonological factors such as consonant versus vowel processing, showing some difference in the magnitude of TL effects for the two types of letters (e.g., Perea & Lupker, 2004). Other studies, however, showed no phonological effects in letter-position coding, reaching the conclusion that transposed letter effects are orthographic in nature (e.g., Acha & Perea, 2010; Perea & Carreiras 2006a, 2006b).

Taken together, these findings suggest that readers display substantial flexibility regarding the coding of letter position. Thus the recognition of printed words appears to be primarily determined by correctly registering the identity of constituent letters, whereas the registry of their exact position within a given word is fuzzy. This finding seems compatible with neurobiological constraints related to noisy registry of information regarding absolute location within the visual system. These constraints concern, among other things, characteristics of receptive fields in the visual cortex, spatial acuity and how it decreases with eccentricity, and neural firing rates (see Dehaene, Cohen, Sigman, & Vinckier, 2005; Whitney & Cornelissen, 2008). Consequently, researchers in the area of visual word recognition have offered in recent years a large number of computational models aimed at reproducing TL effects.

Modeling Letter-Order Coding

The old generation of computational models, such as the interactive-activation model (IAM) (McClelland & Rumelhart, 1981) initially used a nonflexible coding of letter position to differentiate words like *ACT* and *CAT*, where *ACT* is represented and coded as having A in the first position, C in the second, and T in the third. However, general concerns regarding rigid positional coding were acknowledged in early discussions of the IAM model (Rumelhart & McClelland, 1982). One problem with absolute-position coding is that it is not neurologically plausible. No less important, it cannot adequately explain word recognition in different orthographic contexts—for example, the recognition of the base word *ACT* in the morphologically complex word *REACT*, when *A* is now in third position rather than first. Some proposals for alternative coding schemes using context units (for example, representing the word *FROM* by **FR, FRO, ROM*, and *OM**, where * represents the word boundary) were therefore subsequently offered (e.g., Seidenberg & McClelland, 1989). However, the major shift in coding letter position in computational models of visual word recognition has occurred in the last decade. This modeling effort has centered on producing letter-coding schemes and computational solutions that are context sensitive, realistic in terms of visual system constraints, and fit the data regarding readers' relative insensitivity to letter position (e.g., the SERIOL model, Whitney, 2001; the open-bigram model, Grainger & van Heuven, 2003; the SOLAR and the spatial coding model, Davis, 1999, 2010; the Bayesian Reader model, Norris, Kinoshita, & van Casteren, 2010; the overlap model, Gomez, Ratcliff, & Perea, 2008).

The computational principles according to which these different models are structured to fit the data regarding letter-position insensitivity are quite different. For example, the SERIOL model (Grainger & Whitney, 2004; Whitney, 2001, 2008; Whitney & Cornelissen, 2008) is based on serial activation of letter detectors that fire serially in a rapid sequence. The model assumes that the firing sequence serves as input to a layer of open bigram units, which do not contain precise information about letter contiguity but preserve information about relative position. Thus the word *CART* would be represented by activation of the bigram units *#C, CA, AR, RT*, and *T#* (where # represents a word boundary) and also *CR, AT*, and *CT* (the open bigrams). A transposition prime, such as CRAT, shares most of these units, namely *#C, CA*, AT, *RT, CT*, and *T#*, resulting in almost identical priming as the identity prime. Thus the simple introduction of noncontiguous bigram detectors into the model suffices to reproduce TL priming.

Other models obtain letter-position flexibility by assuming noisy slot-based coding. For example, the overlap model (Gomez et al., 2008) posits a letter-order scheme in which the positions of letters in a word are not fixed but are represented as overlapping Gaussian distributions so that the probability of assigning a given position to the different letters decreases with eccentricity. This results in inherent position uncertainty, so that information regarding order of letters is noisier (and therefore slower) than information about letter identity. Similarly, to accommodate TL effects, Kinoshita and Norris (2009), Norris and Kinoshita (2008), and Norris et al. (2010) have implemented as part of their computational model a noisy letter-position scheme in which, in the limited time the prime is presented, information regarding order of letters as well as information about letter identity is ambiguous. In a similar vein, a combination of noisy retinal mapping of letter coding with either contiguous bigram detectors (Dehaene et al., 2005) or location-specific letter detectors (Grainger et al., 2006) has been suggested as well to account for letter-position flexibility.

Although all of the models described here deal in one way or another with fuzzy letter-position coding, they differ in the scope of phenomena they describe. Whereas context-sensitive coding models such as open bigrams are meant from the outset to fit the data regarding TL effects, models like the extended Bayesian Reader model (Norris et al., 2010; and see Norris & Kinoshita, 2012), the overlap model (Gomez et al., 2008), and the spatial coding model (Davis, 2010) offer a rather broad and comprehensive view of visual word recognition and reading, aiming to produce a neurologically plausible model. Discussions regarding the descriptive adequacy of all of these models have centered mainly on their ability to predict the reported TL priming effects and reproduce the empirical data regarding readers' resiliency to different types of letter jumbling. For example, almost all of the 20 simulations offered to validate the recent spatial coding model (Davis, 2010) deal in some way with TL priming effects. Underlying this extensive modeling effort is the implicit supposition that TL effects are universal and reflect the hardwired constraints of the visual system in coding exact position of orthographic units that are sequentially aligned. This supposition, however, is not supported by recent cross-linguistic research conducted in non-European languages, mainly Hebrew and Arabic.

Orthographic Processing in Semitic Languages

Hebrew and Arabic have an alphabetic orthography, where letter strings represent phonemes, similar to European languages. However, both Hebrew and Arabic have a morphological system in which all verbs and most nouns and adjectives are composed of two basic derivational morphemes: a *root* and a *word pattern*. The root usually consists of three consonants, while the word pattern consists of either vowels or a combination of vowels and consonants. The aspect of Hebrew morphology which is relevant to the present context concerns the manner by which these two morphemes are combined. Unlike languages with concatenated morphology, the root and the word pattern are not attached to each other linearly; rather, they are intertwined. The nonlinear structure often obscures the phonological (and the orthographic) transparency of the two morphemes. For example, the Hebrew word /ti**lbos**et/ (written t**lb**w**s**t, 'costume') is a derivation of the root **l.b.s**. This root is mounted on the phonological pattern /ti$\mathbf{C}_1\mathbf{C}_2$o$\mathbf{C}_3$et/ (each C indicates the position of a root consonant). The root **l.b.s.** alludes to the concept of wearing, whereas the phonological pattern /ti$\mathbf{C}_1\mathbf{C}_2$o$\mathbf{C}_3$et/ is often (but not always) used to form feminine nouns. The merging of the root with the word pattern forms the word meaning 'costume.' Other phonological word patterns may combine with the same root to form different words with different meanings that can be either closely or remotely related to the notion of wearing (e.g., /ma**lbus**/ 'clothing', /**lebi**-**s**ah/ 'wearing'), and other roots may combine with the word pattern /ti$\mathbf{C}_1\mathbf{C}_2$o$\mathbf{C}_3$et/ to form feminine nouns (e.g., /ti**zmor**et/, /ti**fzor**et/).

Although Semitic languages have alphabetic orthographies as do European languages, TL priming is not obtained in Hebrew or Arabic. The first demonstration of letter-coding rigidity rather than insensitivity was reported by Velan and Frost (2007). In this study, Hebrew-English balanced bilinguals were presented with sentences in English and in Hebrew, half of which had transposed-letter words (three jumbled words in each sentence) and half of which were intact. The sentences were presented on the screen word-by-word via RSVP so that each word appeared for 200 ms. Following the final word, subjects had to say the entire sentence and report whether they had detected letter transpositions in the sentence. The results showed a marked difference in the effect of letter transposition in Hebrew compared with English. For English materials, the report of words was virtually unaltered when sentences included words with transposed letters, and reading performance in sentences with and without jumbled letters was quite similar. This outcome concurs with the Cambridge University effect and all findings regarding letter-position flexibility reported in English or other European languages (e.g., Duñabeitia et al., 2007; Perea & Carreiras, 2006a, 2006b; Perea & Lupker, 2003, 2004; Schoonbaert & Grainger, 2004). For Hebrew materials, however, letter transpositions were detrimental to reading, and performance in reading sentences that included words with jumbled letters dropped dramatically.

Perhaps the most revealing finding of the Velan and Frost (2007) study concerns subjects' ability to perceptually detect the transposition of letters in Hebrew as compared with English, as revealed by the sensitivity measure *d prime*. As described earlier, at the rate of presentation of 200 ms per word in RSVP, subjects' sensitivity to detection of transposition with English material was particularly low. In contrast, subjects' sensitivity to detecting the transposition with Hebrew material was exceedingly high (*d'* = 2.51, in contrast to *d'* = 0.86 in English), and not a single subject was at chance level in the perceptual task. Since *d'* mainly taps the early perceptual level of processing, this outcome suggests a difference in the characteristics of orthographic processing in Hebrew and English.

The significant sensitivity of Hebrew readers to letter transpositions raises the question of whether the typical TL priming effects obtained in European languages are obtained in Hebrew. The answer seems, again, straightforward. Hebrew TL primes do not result in faster target recognition relative to letter substitution, even though they do in English, Dutch, French, and Spanish. More important, if jumbling the order of letters in the prime results in a letter order that alludes to a different root than that embedded in the target, significant inhibition rather than facilitation is observed (Velan & Frost, 2009). This double dissociation between Hebrew and European languages regarding the effect of letter transposition suggests that letter-position encoding in Hebrew is far from flexible. Rather, Hebrew readers seem to display remarkable rigidity regarding letter order. Identical results have been demonstrated in Arabic (Perea, Abu Mallouh, & Carreiras, 2010).

Assuming that the neural circuitry of the visual system is identical for readers of Hebrew and English, what is the origin of these cross-linguistic differences in TL priming effects? The results from Hebrew and Arabic demonstrate that the cognitive system may perform different types of processing on a sequence of letters depending on factors that are unrelated to peripheral orthographic characteristics but related to the deep structural properties of the printed stimuli. Such findings bear major implications for understanding orthographic processing and the coding of letter order.

Why Are Hebrew or Arabic Readers So Sensitive to Letter Order?

Visual word recognition in Hebrew has been extensively investigated in an array of experimental paradigms such as masked priming, cross-modal priming, and the monitoring of eye movements (Deutsch, Frost, & Forster, 1998; Deutsch, Frost, Pollatsek, & Rayner 2000; Deutsch, Frost, Peleg, Pollatsek, & Rayner 2003; Deutsch, Frost, Pollatsek, & Rayner 2005; Feldman, Frost, & Pnini, 1995; Frost, Forster, & Deutsch, 1997; Frost, Deutsch, & Forster, 2000; Frost, Deutsch, Gilboa, Tannenbaum, & Marslen-Wilson, 2000; Frost, Kugler, Deutsch, & Forster, 2005; Velan, Frost, Deutsch, & Plaut, 2005). One consistent finding is that root-letter primes facilitate both lexical decision and naming of target words that are derived from these roots. Similarly, eye-movement studies have demonstrated that a parafoveal preview of the root letters results in shorter eye fixations on targets that are root derivations. Taken together, these findings suggest that the root morpheme serves as an organizing unit in the mental lexicon of Hebrew readers (e.g., Deutsch et al., 1998; Frost et al., 1997, 2005) and is, therefore, the target of lexical search. This is because Semitic roots systematically convey the shared meaning of all words derived from them (see Frost, 2006, 2009, for discussion). Thus the orthographic code generated for Hebrew does not seem to consider all of the constituent letters equally. As reliable facilitation is obtained whenever primes consist of the root letters, irrespective of what the other letters are (see also Perea et al., 2010, for Arabic), the orthographic coding scheme of Hebrew print appears to rely mainly on the three letters that carry root information.

These considerations suggest a simple explanation for extreme rigidity of letter encoding for Semitic words. Hebrew has about 3,000 roots (Ornan, 2003), which form the variety of Semitic Hebrew words. Since these triconsonantal entities are conveyed by the 22 letters of the alphabet, for simple combinatorial reasons, it is inevitable that several roots share the same three letters. To avoid the complications of homophony, Semitic languages alter the order of consonants to create different roots so that typically, three or four different roots can share the same set of three consonants (and thereby three letters). For example, the consonants of the root **l.b.s** ('to wear') can be altered to produce the root **b.s.l** ('costume'), **s.l.b** ('to combine'), and **b.l.s** ('detective'). If the orthographic processing system has to pick up the root information from the distal letter sequence, letter order cannot be flexible; it must be extremely rigid. Moreover, for a system to efficiently differentiate between roots sharing the same letters but in a different order, inhibitory connections must be set between different combinations of the same letters, each of which represents a different meaning.

A convincing demonstration that orthographic processing and the coding of letter position in alphabetic orthographies are entirely dependent on the type of morphological information carried by individual letters can be shown, again, in Semitic languages. Both Hebrew and Arabic have a large set of base words that are morphologically simple. That is, they do not have the typical Semitic structure, since they are not root-derived and thus resemble words in European languages. Such words have infiltrated Hebrew and Arabic throughout history from adjacent linguistic systems such as Persian or Greek, but native speakers of Hebrew or Arabic are unfamiliar with their historical origin. Velan and Frost (2011) found that morphologically simple words revealed the typical form priming and TL priming effects reported in European languages. In fact, Hebrew-English bilinguals did not display any differences between processing these words and processing English words. In contrast, whenever Semitic words that are root-derived were presented to the participants, the typical letter-coding rigidity emerged. For these words, form priming could not be obtained, and transpositions resulted in inhibition rather than in facilitation. Velan, Deutsch, and Frost (2013) extended these findings and examined the time course of processing letter transpositions in Hebrew, assessing their impact on reading the different types of Hebrew words (Semitic vs. non-Semitic). By monitoring eye movements, Velan et al. (2013) found that letter transposition resulted in dramatic reading costs for words with Semitic word structure and much smaller costs for non-Semitic words, even at the first fixation. This result suggests that Hebrew readers differentiate between Semitic and non-Semitic forms at the very early phases of visual word recognition, so that letters are differentially processed across the visual array, given their morphological structure and their contribution to recovering semantic meaning.

Writing Systems Modulate Coding of Letter Order

Both the reported cross-linguistic studies, as well as studies within one language—Hebrew—demonstrate that flexible letter-position coding is not a general property of the cognitive system or a property of a given language. In other words, it is not the coding of letter position that is flexible, but the reader's strategy in processing it. Letter-transposition effects are therefore not universal: They appear or disappear given the overall phonological and morphological structure of the language. If we assume that significant noise exists in registering the exact position of sequential visual stimuli and that identical neurobiological constraints govern reading performance in any language, what determines whether the coding of letter position will be flexible or rigid?

Important insights are provided by recent simulations reported by Lerner, Armstrong, and Frost (2014). Lerner et al. investigated how a simple domain-general connectionist architecture performs in tasks such as letter transposition and letter substitution, having learned to process words in the context of different linguistic environments. They constructed a multilayer connectionist network that maps orthographic inputs to semantic outputs. The network was trained to map inputs to outputs for English and for Hebrew stimuli using backpropagation. The network's behavior was then tested in response to new pseudowords that had letter transpositions or letter substitutions. The study involved two different artificial linguistic environments, English-like and Hebrew-like. In the English-like environment, letter sets were associated with one meaning only. Hence, there were no anagrams. In the Hebrew-like environment, letter sets were associated with more than one meaning by switching letter order, so that there were many anagrams, mimicking the characteristics of Semitic morphology. Lerner et al. (2014) found that the relatively simple domain-general learning model produced the cross-linguistic differences in TL effects reported in Hebrew and English, when it was trained on the English-like and Hebrew-like linguistic environments. Thus, independent of the noise involved in registering letter position in all languages, flexibility and inflexibility in coding letter order is shaped by the number of anagrams in the language.

In a subsequent simulation, Lerner and his colleagues trained their model on a random sample of real English words and real Hebrew words and then tested the model's performance in mapping orthography to meaning for nonwords that had letter transpositions. Large TL priming effects emerged for English, and much smaller effects emerged for Hebrew. Interestingly, the cross-linguistic differences were modulated by the differences in the number of anagrams in English and Hebrew at various word lengths. Lerner et al. consequently argued that readers of European languages can essentially rely on letter identity information alone to activate a correct semantic representation. This is because most words in European languages have different letter sets, and anagrams such as *clam-calm* are the exception rather than the rule. Readers of European languages learn, therefore, to rely primarily on this superior source of information. In Hebrew, however, anagrams are the rule rather than the exception. Therefore, error-driven learning cannot correctly activate semantic representations without considering position information. Readers of Semitic languages thus learn to rely on positional information despite the inherent fuzziness of positional input representations.

This account provides an alternative framing to TL effects. Rather than focusing on hardwired neurobiological constraints related to how the brain encodes the position of letters in printed words in all orthographies, the explanation of TL effects shifts and is considered to be an emergent behavior occurring only when the brain learns languages with specific statistical properties. The theoretical implications of this approach are far reaching, as discussed in what follows.

Universal Principles of Letter-Position Coding

Two principles thus far set the theoretical foundations for explicating letter-position coding in different writing systems. The first is that, from a pure neurobiological perspective, the registry of letter position in any language is inherently noisier than that of letter identity (e.g., Martelli, Burani, & Zoccolotti, 2012; Perea & Carreiras, 2012; see Gomez et al., 2008, for discussion). Thus, accurate position information requires greater resources and more extensive processing than identity information. The second principle is that this neurobiological constraint interacts with the idiosyncratic statistical properties of a given writing system to determine the preciseness or fuzziness of letter-position coding, as revealed by TL priming effects. These statistical properties differ between writing systems as they are shaped by the language's morphological structure and the characteristics of its phonological structure.

In a recent review (Frost, 2012a) I argued that writing systems have evolved to optimally represent the languages' phonological spaces and their mapping into semantic meaning, so that basic principles related to optimization of information can account for the variety of human writing systems and their different characteristics. These principles are important for understanding reading, because they provide critical insight regarding how the cognitive system picks up the information conveyed by print. This view has the flavor of a Gibsonian ecological approach (Gibson 1986) and assumes that to be efficient, the cognitive system that processes language must be tuned to the structure of the linguistic environment in which it operates. In the present context, it explains letter-position coding in terms of the interaction of neurobiological mechanisms with language properties. What, then, are the relevant features that eventually impact letter-position coding?

The discussion so far leads to the conclusion that the manner in which letters combine to form words in a given orthography has an impact on their processing. The cognitive system thus performs different operations on different sequences of letters given the deep structural properties of the printed stimuli, which are language-dependent (see Frost, 2012a, 2012b, for an extensive discussion). For example, European languages impose few rigid constraints on the phonological internal structure of base words, so that in principle most phonemes (and therefore letters) could be located in any position within the spoken word (albeit not necessarily with equal probability, see for example Kessler & Treiman, 1997). Most importantly, base words in European languages are morphologically simple, since morphological complexity (inflections and derivations) proceeds by linearly adding affixes to a base morpheme. For these languages, the individual letters composing base words contribute to meaning retrieval equally. In contrast, Semitic base words are necessarily morphologically complex (a root embedded in a word pattern morpheme), and the number of word patterns is relatively small. Readers are thus repeatedly presented with word-pattern letter sequences with very high distributional properties, while the contribution of root letters to meaning recovery exceeds that of the frequently repeated word pattern letters. Note that on the average, printed words in Hebrew or Arabic have fewer letters than printed words in European languages (most vowels are not conveyed in print; see Bentin & Frost, 1987; Frost, Katz, & Bentin, 1987). Because words with similar word patterns are differentiated only by root letters, and different roots inevitably share the same subset of letters, Hebrew orthographic lexical space is exceedingly dense. Often, several words share the same set of letters but in a different order. To become an efficient reader, an individual must pick up and implicitly assimilate these statistical properties related to phonology and morphology. This results in learning to use precise letter-position coding for Semitic words and relax this criterion for non-Semitic words.

The focus on Semitic and European languages in the present discussion is meant to outline a general principle rather than discuss the specificities of one writing system or another. The brief review of findings regarding letter-transposition effects in various writing systems leads to important insights regarding the principles by which a theory of visual word recognition should be constructed. The review so far suggests that orthographic processing cannot be researched, explicated, or understood without considering the manner in which orthographic structure represents phonological, semantic, and morphological information in a given writing system. This is because any orthographic effect obtained in a given language, such as sensitivity to letter order, is an emerging product of the full linguistic environment of the reader, not just of the structure of the graphemic sequence (cf. Norris & Kinoshita, 2012). The statistical properties of letter distributions in the language and their relative contribution to meaning have to be picked up by readers of the language, and the transitional probabilities of letter sequences have to be implicitly assimilated. They can be understood only while considering the full phonological and morphological structure of the language. In a nutshell, a theory of reading should be a theory of the interaction of the reader with his or her linguistic environment. This approach also sets clear guidelines regarding how future computational models of letter coding should be constructed.

Implications for Modeling Letter-Position Coding

Most recent models of visual word recognition were set to fit the emerging data regarding letter-transposition effects. The logic of this modeling approach in cognitive science follows a series of steps. First, a body of findings is identified (for example, TL priming); second, new computational mechanisms are hypothesized to fit the particular set of data (for example, structuring open bigrams in the model); third, the model's performance is evaluated by its quantitative fit to specific existing empirical findings; and finally, the model's architecture becomes a theoretical construct in explaining behavior. This strategy, labeled by Rueckl (2012) the backward engineering approach to modeling, has some merits. It provides testable predictions and it generates potential hypotheses regarding the source of behavior. It has, however, serious limitations. The first is that it has inherently a narrow scope: The models often lack generalization because their architectures are tailored to produce and fit a narrow predetermined set of effects. The second limitation of the backward engineering approach is that it often has narrow support: The models primarily reflect the modeler's intuitions about the source of a behavior with little independent empirical support. The fitting of the data in itself is taken as proof that the mechanism of behavior has been well understood, so that often the mechanism is presented as a tautological redescription of the data. Finally, the approach has a narrow horizon. It often fails to explore and predict new types of phenomena that could emerge from general computational principles.

Regarding the specific domain of letter-position coding, the current discussion results in the conclusion that flexibility or rigidity in coding letter-position emerges as a learning principle that is tuned to the statistical properties of a linguistic environment, thereby resulting in strong or weak TL effects or lack thereof. Therefore, for a model to produce differential behavior as a function of the statistical properties of the particular language, the model has to be able to pick up the statistical properties of the language. This suggests that only models that are based on domain-general learning mechanisms have the potential ability to produce cross-linguistic differences in TL priming (see Frost, 2012a, 2012b, for a detailed discussion). The main advantage of these models is they are not aimed to fit a predetermined set of data; rather, they are designed to pick up, through simple learning principles, the distributional characteristics of the input. Once a behavior has emerged, the models provide a relatively transparent explanation as to why it has evolved (see Lerner et al., 2014). In general, the emphasis on domain-general principles, and in particular on simple learning mechanisms, is compatible with a broad range of learning and processing phenomena, allowing the model to have a potentially wider scope, capturing, explaining, and predicting empirical phenomena observed in any language. Models of this kind have had significant success in producing the cross-linguistic differences in letter-position coding. For example, using naïve discriminative learning (Baayen et al., 2011), which maps form directly into meaning without hidden layers mediating the mapping, Baayen (2012) compared the sensitivity to letter order and the costs of letter transposition in English and biblical Hebrew, for cases in which words from the two languages were aligned with their meanings (text taken from the book of Genesis, or random selections of words from the database of phrases from the British National Corpus). Baayen demonstrated that the predictive value of pairs of contiguous letters (correlated with order information in the model) was significantly higher in Hebrew than in English, thereby showing that greater sensitivity to letter order emerges in Semitic languages when nothing but abstract discriminant learning principles are considered. In the same vein, Lerner et al. (2014) have shown that a connectionist neural network trained on Hebrew and English words produces the observed cross-linguistic differences.

Summary and Concluding Remarks

Most current discussions of orthographic processing and letter-order coding have focused on characterizing the mechanisms involved in the front end of word perception, considering letters as visual entities and assuming identical processing principles across orthographies. The extreme version of this approach regards printed words as two-dimensional objects that are treated by the visual system like any other visual object, so that the linguistic information carried by individual letters is irrelevant (e.g., Grainger & Hannagan, 2012; Norris & Kinoshita, 2012; Ziegler et al., 2013). Here, I advocate an opposite approach. By considering cross-linguistic differences in letter-position effects, I argue that the statistical properties embedded in a writing system given its morphological structure and phonological constraints govern early orthographic processing in a given language. Consequently, a theory of visual word recognition should consider both the neurobiological constraints of the information processing system and the linguistic environment on which it operates. Letters in words are not simply visual objects. Their specific alignment one next to the other in a particular language reflects phonological, semantic, and morphological considerations, which are themselves the object of perception, as exemplified by cross-linguistic differences in processing letter position. The brain adapts to a writing system in the course of literacy acquisition, so that it is hard to discern the demarcation line between vision per se and language. However, a strictly bottom-up feed-forward approach according to which the perceptual system fully completes its task, to feed its output to the linguistic system, is not very probable, nor it is supported by the data. Visual word recognition necessarily involves processing a noisy information channel, but the manner by which readers deal with this noise is language-dependent.

Acknowlegments

This chapter was supported in part by the Israel Science Foundation (Grant 159/10 awarded to Ram Frost), and in part by the National Institute of Child Health and Human Development (RO1 HD067364 awarded to Ken Pugh and Ram Frost, and HD-01994 awarded to Haskins Laboratories).

References

Acha, J., & Perea, M. (2010). Does kaniso activate CASINO? Input coding schemes and phonology in visual-word recognition. *Experimental Psychology, 57,* 245–251.

Baayen, R. H. (2012). Learning from the Bible: Computational modeling of the costs of letter transpositions and letter exchanges in reading Classical Hebrew and Modern English. *Lingue & Linguaggio, 2,* 123–146.

Baayen, R. H., Milin, P., Durdevic, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review, 118,* 428–481.

Beyersmann, E., McCormick, S. F., & Rastle, K. (2013). Letter transpositions within morphemes and across morpheme boundaries. *Quarterly Journal of Experimental Psychology, 66,* 2389–2410.

Bentin, S., & Frost, R. (1987). Processing lexical ambiguity and visual word recognition in a deep orthography. *Memory & Cognition, 15,* 13–23.

Christianson, K., Johnson, R. L., & Rayner, K. I. (2005). Letter transpositions within and across morphemes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1327–1339.

Davis, C. J. (1999). *The Self-Organizing Lexical Acquisition and Recognition (SOLAR) model of visual word recognition*. Unpublished doctoral dissertation, University of New South Wales.

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review, 117,* 713–758.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences, 9,* 335–341.

Deutsch, A., Frost, R. & Forster, K. I. (1998) Verbs and nouns are organized and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1238–1255.

Deutsch, A., Frost, R., Peleg, S, Pollatsek, A., & Rayner, K. (2003). Early morphological effects in reading: Evidence from parafoveal preview benefit in Hebrew. *Psychonomic Bulletin & Review, 10,* 415–422.

Deutsch, A., Frost, R., Pollatsek A., & Rayner, K. (2000). Early morphological effects in word recognition in Hebrew: Evidence from parafoveal preview benefit. *Language and Cognitive Processes, 15,* 487–506.

Deutsch, A., Frost, R., Pollatsek, A. & Rayner, K. (2005). Morphological parafoveal preview benefit effects in reading: Evidence from Hebrew. *Language and Cognitive Processes, 20,* 341–371.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2007). Do transposed-letter similarity effects occur at a morpheme level? Evidence for ortho-morphological decomposition, *Cognition, 105,* 691–703.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2014). Revisiting letter transpositions within and across morphemic boundaries. *Psychonomic Bulletin & Review, 21,* 1557–1575.

Feldman, L. B., Frost, R., & Pnini, T. (1995). Decomposing words into their constituent morphemes: Evidence from English and Hebrew. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21,* 947–960.

Forster, K. I. (1976). Accessing the mental lexicon. *New Approaches to Language Mechanisms, 30,* 231–256.
Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology, 39A,* 211–251.
Frost, R. (1998). Towards a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123,* 71–99.
Frost, R. (2006). Becoming literate in Hebrew: The grain-size hypothesis and Semitic orthographic systems. *Developmental Science, 9,* 439–440.
Frost, R. (2009). Reading in Hebrew vs. reading in English: Is there a qualitative difference? In K. Pugh & P. McCardle (Eds.), *How children learn to read: Current issues and new directions in the integration of cognition, neurobiology and genetics of reading and dyslexia research and practice* (pp. 235–254). New York, NY: Psychology Press.
Frost, R. (2012a). Towards a universal model of reading. *Behavioral and Brain Sciences, 35,* 269–173.
Frost, R. (2012b). A universal approach to modelling visual word recognition and reading: Not only possible, but also inevitable. *Behavioral and Brain Sciences, 35,* 310–329.
Frost, R., Deutsch, A., & Forster, K. I. (2000). Decomposing morphologically complex words in a nonlinear morphology. *Journal of Experimental Psychology Learning, Memory, and Cognition, 26,* 751–765.
Frost, R., Deutsch, A., Gilboa, O., Tannenbaum, M., & Marslen-Wilson, W. (2000). Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory & Cognition, 28,* 1277–1288.
Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 829–856.
Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 104 –115.
Frost, R., Kugler, T., Deutsch, A., & Forster, K. I. (2005). Orthographic structure versus morphological structure: Principles of lexical organization in a given language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1293–1326.
Gibson, J. J. (1986). *The ecological approach to visual perception*. Hillsdale, NJ: Erlbaum.
Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review, 115,* 577–601.
Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes, 23,* 1–35.
Grainger, J., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language, 33,* 218–233.
Grainger, J., Granier, J. P., Farioli, F., Van Assche, E., & Van Heuven, W. (2006). Letter position information and printed word perception: The relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 865–884.
Grainger, J., & Hannagan, T. (2012). Explaining word recognition, reading, the universe, and beyond: A modest proposal. *Behavioral and Brain Sciences, 35,* 288–289.
Grainger, J., & Van Heuven, W. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *The mental lexicon* (pp. 1–24). New York, NY: Nova Science.
Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? *Trends in Cognitive Sciences, 8,* 58–59.
Grainger, J., & Ziegler, J. C. (2011). A dual-route approach to orthographic processing. *Frontiers in Language Sciences, 2,* 54.
Guerrera, C., & Forster, K. I. (2008). Masked form priming with extreme transposition. *Language and Cognitive Processes, 23,* 117–142.
Humphreys, G. W., Evett, L. J., & Quinlan, P. T. (1990). Orthographic processing in visual word recognition. *Cognitive Psychology, 22,* 517–560.
Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading. Evidence from eye-movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 209–229.
Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language, 37,* 295–311.
Kinoshita, S., & Norris, D. (2009). Transposed-letter priming of prelexical orthographic representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 1–18.
Lerner, I., Armstrong, B. C., & Frost, R. (2014). What can we learn from learning models about sensitivity to letter-order in visual word recognition. *Journal of Memory and Language, 77,* 40–58.
Marslen-Wilson, W. D., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review, 101,* 3–33.
Martelli, M., Burani, C., & Zoccolotti, P. (2012). Visual perceptual limitations on letter position uncertainty in reading. *Behavioral and Brain Sciences, 35,* 32–33.
McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.
Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165–178.
Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology: General, 137,* 433–455.
Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: Why there's nothing special about the perception of orthography. *Psychological Review, 119,* 517–545.
Norris, D., Kinoshita, S., & van Casteren, M. (2010). A stimulus sampling theory of letter identity and order. *Journal of Memory and Language, 62,* 254–271.
Ornan, U. (2003). *The final word: Mechanism for Hebrew word generation*. Haifa, Israel: Haifa University Press. (In Hebrew.)
Perea, M., Abu Mallouh, R., & Carreiras, M. (2010). The search for an input coding scheme: Transposed-letter priming in Arabic. *Psychonomic Bulletin and Review, 17,* 375–380.
Perea, M., & M. Carreiras (2006a). Do transposed-letter similarity effects occur at a prelexical phonological level? *Quarterly Journal of Experimental Psychology, 59,* 1600–1613.
Perea, M., & Carreiras, M. (2006b). Do transposed-letter similarity effects occur at a syllable level? *Experimental Psychology, 53,* 308–315.

Perea, M., & Carreiras, M. (2012). Perceptual uncertainty is a property of the cognitive system. *Behavioral and Brain Sciences, 35,* 298–300.

Perea, M., & Lupker, S. J. (2003). Transposed-letter confusability effects in masked form priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art* (pp. 97–120). New York, NY: Psychology Press.

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language, 51,* 231–246.

Perea, M., & Perez, E. (2009). Beyond alphabetic orthographies: The role of form and phonology in transposition effects in Katakana. *Language and Cognitive Processes, 24,* 67–88.

Peressotti, F., & Grainger, J. (1999). The role of letter identity and letter position in orthographic priming. *Perception & Psychophysics, 61,* 691–706.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Rawlinson, G. E. (1976). *The significance of letter position in word recognition.* Unpublished PhD thesis, Psychology Department, University of Nottingham, Nottingham, UK.

Rawlinson, G. E. (1999). Reibadailty. *New Scientist, 162,* 55.

Rayner, K., White, S. J., Johnson, R. L., & Liversedge, S. P. (2006). Raeding wrods with jubmled lettrs: There's a cost. *Psychological Science, 17,* 192–193.

Rueckl, J. G. (2012). The limitations of the reverse-engineering approach to cognitive modeling. *Behavioral and Brain Sciences, 35,* 43.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60–94.

Sánchez-Gutiérrez, C., & Rastle, K. (2013). Letter transpositions within and across morphemic boundaries: Is there a cross-language difference? *Psychonomic Bulletin and Review, 20,* 988–996.

Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes, 19,* 333–367.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134,* 584–615.

Van Assche, E., & Grainger, J. (2006). A study of relative-position priming with superset primes. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32,* 399–415.

Velan, H., Deutsch, A., & Frost, R. (2013). The flexibility of letter-position flexibility: Evidence from eye-movements in reading Hebrew. *Journal of Experimental Psychology. Human Perception and Performance, 39,* 1143–1152.

Velan, H., & Frost, R. (2007). Cambridge University vs. Hebrew University: The impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review, 14,* 913–918.

Velan, H., & Frost, R. (2009). Letter-transposition effects are not universal: The impact of transposing letters in Hebrew. *Journal of Memory and Language, 61,* 285–302.

Velan, H., & Frost, R. (2011). Words with and without internal structure: What determines the nature of orthographic processing. *Cognition, 118,* 141–156.

Velan, H., Frost, R., Deutsch, A., & Plaut, D. (2005). The processing of root morphemes in Hebrew: Contrasting localist and distributed accounts. *Language and Cognitive Processes, 29,* 169–206.

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review, 8,* 221–243.

Whitney, C. (2008). Supporting the serial in the SERIOL model. *Language and Cognitive Processes, 23,* 824–865.

Whitney, C., & Cornelissen, P. (2008). SERIOL reading. *Language and Cognitive Processes, 23,* 143–164.

Ziegler, J. C., Hannagan, T., Dufau, S., Montant, M., Fagot, J., & Grainger, J. (2013). Transposed-letter effects reveal orthographic processing in baboons. *Psychological Science, 24,* 1609–1611.

CHAPTER

8 The Nature of Lexical Representation in Visual Word Recognition

Marcus Taft

Abstract

This chapter explores how information about words is represented for the purposes of recognizing those words when reading. A description is first given of the various architectures that have been proposed to frame our understanding of lexical processing, with an emphasis on the way they portray lexical representation. The importance of morphological structure to the nature of lexical representation is highlighted, and attention is directed to specific models that attempt to capture that structure. The model that forms the major focus of the chapter, the AUSTRAL model, is one where identification of a letter string is based on information associated with an abstract level that mediates between form and function; namely, a lemma level. The incorporation of a lemma level into the lexical processing system provides a locus for morphological structure. It captures a level of lexical representation that not only underlies both visual and spoken word recognition but also is compatible with models of word production.

Key Words: hierarchical activation, lemmas, lexical representation, models of lexical processing, morphological processing, orthographic processing

It is easy for a literate English speaker to recognize that *cat* is a word while *lat* is not. Only the former is represented in the speaker's long-term memory for words (i.e., the mental lexicon), and access to this representation allows the speaker to differentiate between real words and pronounceable strings of letters that are nonwords (also called pseudowords, e.g., *frink*). Such a lexical decision judgment therefore provides a window into the nature of the lexical representations that are accessed when we read and, because of this, has been adopted widely as a laboratory-based task to explore the issue of lexical representation and retrieval.

Participants in the lexical decision task are typically presented a randomly mixed series of words and nonwords and asked to press a "yes" or "no" button as quickly but as accurately as possible in response to whether the presented letter string is a word or not. Reaction times (RTs) and error rates are measured. The RT for a word item (usually averaging around 500 ms) reflects the time it takes to access the relevant lexical information, as well as to decide that such information is sufficient to indicate that the letter string is a word. The RT for a nonword (usually longer than for a word item) reflects the amount of lexical information accessed on the basis of the letter string and the time it takes to decide that this is insufficient for a word response to be made. Different types of item are compared by including 15 or more examples of each. The two conditions making up each type vary on the factor of interest while being matched closely on as many important factors as possible (e.g., frequency of occurrence in the language in the case of real words). In the case of priming research, the same target is compared when preceded by different primes and the impact of different relationships between the prime and target can therefore be measured.

Linguists are likely to interpret the notion of lexical representation in terms of the linguistic information associated with a word (i.e., its semantic and syntactic functions). For the cognitive psychologist who is interested in how we read, however, an understanding of the nature of the functional information itself is less important than the mental representations that provide access to that information. For this reason, the issue that will be addressed in this chapter is the nature of the representation that allows identification of the word. Couched in terms of the lexical decision task, the question is what representation is accessed that allows a letter string to be identified as a particular word (or alternatively to be classified as a nonword). It is this representation that constitutes the gateway through which the incoming letter string can be associated with its functional interpretation during reading.

Models of Lexical Processing

Early Views

The notion of lexical access in visual word recognition was first explored in the late 1960s and early 1970s, with two distinct approaches being adopted.

Lexical Search

Forster (1976) outlined a model of lexical processing where all information about a word (semantic, syntactic, phonological, and orthographic) is stored in a master file that is accessed via a serial search through modality-specific peripheral access files. The orthographic access file that is used in reading is a list of words in order of their frequency of occurrence, although divided into smaller sized bins according to form-based characteristics. A visually presented word might therefore be recognized when found to match with an entry in the orthographic access file. However, Taft and Forster (1975, 1976) argued that a polymorphemic word (such as *revive, henchman*) whose stem is not a free-standing word is recognized when that stem (e.g., *vive, hench*) is accessed in the orthographic access file, with information about the whole word being subsequently extracted from the master file entry. As such, the access file includes nonwords (e.g., *vive, hench*), which means that word recognition requires access to the master file because that is the locus of information necessary to discriminate real words from nonwords.

Lexical Activation

An alternative idea that words are accessed in lexical memory via a parallel activation system was incorporated by Morton (1969, 1970) in his logogen model. The lexical entry for each word is seen as an information-collecting device (i.e., a logogen) whose activation level increases in response to relevant features being contained in the stimulus. Once enough evidence accumulates in one of the logogens for its threshold to be reached, the corresponding word becomes available for recognition and the relevant functional information associated with that word can be accessed. There are separate sets of orthographic and phonological logogens that are activated depending on the modality of the input. In addition, Morton argued for the existence of orthographic and phonological output logogens that are used for writing and speaking, respectively, although it is unclear how parallel incremental activation would work within an output system.

Lexical representations in the logogen model are the words corresponding to each logogen. However, on finding that inflected words prime recognition of their stem (e.g., *cars* primes *car*, while *card* does not), Murrell and Morton (1974) concluded that logogens actually correspond to morphemes. Since the logogen model does not specify how polymorphemic words might be represented, its notion of lexical representation is rather vague.

More Recent Approaches

The idea of serial search through lexical memory has become largely outdated, with most investigators now adopting an account that incorporates a parallel activation mechanism. Even the main proponent of lexical search has recently proposed a model that combines parallel activation with serial processing (Forster, 2012). Therefore, the nature of lexical representation will now be considered in the light of the activation frameworks that are currently most influential.

Interactive-Activation

The interactive-activation (IA) model, as first outlined by McClelland and Rumelhart (1981) and Rumelhart and McClelland (1982), elaborates on the notion of parallel activation. Words are represented in the IA model in much the same way as in the logogen account. However, there is also a layer of activation units corresponding to individual letters that feed their activation to the word level, and a layer of activation units corresponding to visual features that feed their activation to the letter level. Thus activation passes up from features to letters to words, with activated units inhibiting competitors at the same level. As a unit increases in activation, it feeds activation back down the hierarchy so that the lower-level units whose activation has been most productive at the higher level will be strengthened. From this interaction of activation throughout the system, a single word unit will eventually reach a threshold that allows the letter string to be recognized as that word.

The word units in the IA model are lexical representations in the sense that a letter string can be identified once one such unit reaches its recognition threshold. However, the term "lexical representation" need not be taken in its literal sense of "the mental portrayal of a complete word," but can be more broadly defined as "the stored information through which a word can be recognized." Given that the sublexical units (i.e., features and letters) are integral to the activation of word units in the IA model, it can be argued that they are also a part of the lexical representation. Thus when considering the nature of lexical representation, the breadth of its definition needs to be clear. To further our understanding of visual word recognition we will primarily be interested in the broadest definition, because the act of reading involves the whole procedure of getting from the letter string to its identification as a particular word.

PARALLEL DISTRIBUTED PROCESSING

Like the IA approach, the parallel distributed processing (PDP) model is a connectionist framework. However, while the former represents words as localist (i.e., specified) lexical units, the latter captures lexical information within patterns of activation distributed across sets of hidden units that mediate among the orthographic, phonological, and semantic levels of representation (e.g., Harm & Seidenberg, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989). As such, there is no explicit representation of the whole word (i.e., no lexical entry per se), only a pattern of connection weights that has been settled on through repeated experience with the letter string. Since representations in the PDP model are common to more than one word and also participate in the processing of nonwords, it can be said that lexical representations do not exist (cf. Seidenberg & McClelland, 1989, p. 560).

However, in a general sense, there must be some type of lexical representation within the PDP system given that it is possible to discriminate words from nonwords. Lexical decision judgments are made by comparing the orthographic input with the orthographic output that is generated by the hidden units on the basis of that input. This provides a measure of orthographic familiarity because the more the system encounters a letter string, the more accurately the weights within the hidden units will settle on a pattern that reflects the orthographic form of that letter string. Since a nonword has never been previously encountered, the pattern of activation generated in the hidden units will be a less accurate reflection of the orthographic input and, if this match between input and output falls below some criterial level, a nonword classification can be made. Within such an account, then, it is the settled pattern of activation within the hidden units that is equivalent to a lexical representation. While the representation of a word may be distributed across a number of units that overlap with the distribution of units representing other words in the vocabulary, the pattern of weighted activation is nevertheless unique and therefore functions as a lexical representation.

Within the PDP framework, a pattern of connection weights becomes more stable the more an orthographic form is systematically associated with an output, either phonological or semantic. So, systematic sublexical relationships that exist between print and sound are captured within the hidden units that mediate between the orthographic and phonological levels (e.g., the fact that ‹EE› is typically pronounced /i:/), while systematic sublexical relationships that exist between print and meaning are captured within the hidden units that mediate between the orthographic and semantic levels. In fact, the only sublexical systematicity between print and meaning is at the level of the morpheme and, as such, the hidden units mediating between orthography and semantics must reflect morphemic rather than submorphemic information (e.g., Gonnerman, Seidenberg, & Andersen, 2007; Rueckl & Raveh, 1999). Given the similar lack of submorphemic systematicity between phonology and semantics, it is parsimonious to amalgamate the hidden units that mediate between orthography and meaning with those that mediate between phonology and meaning, and that is what Gonnerman et al. (2007) propose.

The PDP approach, regardless of its specific computational implementation, is impressive in the way it simulates known data by capturing the statistical relationships that exist between the orthographic, phonological, and semantic domains. However, the model can be seen more as a facsimile of human reading performance than providing an understanding of it. The PDP explanation for how we read is essentially that the presented letter string sets up a pattern of neural activity that corresponds to a pattern that was previously acquired in response to that letter string. Such an account is therefore not very instructive and, accordingly, the PDP approach has not proven very successful over the years in generating new research into the reading process. For this reason, many have sought a more revealing approach to the question of lexical representation and the processes involved in reading words. This typically involves the adoption of the IA framework where the localist description gives a clearer picture of what the units that are involved in the processing of the word represent. A description of one such approach follows.

The AUSTRAL Model

The model to be outlined here in greater detail was introduced in Taft (1991) and developed further in Taft (2006). It adopts the IA framework, but replaces the word level with a level of representation that provides a link between function (semantic, syntactic, pragmatic, etc.) and form regardless of the modality of input which, as such, incorporates information about morphemic structure. Taft (2006) refers to this as the lemma level following the lead of Baayen, Dijkstra, and Schreuder (1997), who adopted the notion from the production literature (cf. Kempen & Huijbers, 1983; Roelofs, 1992). The inclusion of such a level of representation makes it the localist equivalent of the PDP model of Gonnerman et al. (2007), inasmuch as meaning is linked to both orthographic and phonological form via the same set of units that capture morphemic information.

Figure 8.1 depicts a version of this model, which will be referred to as AUSTRAL for reasons to be given shortly. The figure illustrates the way in which the word *cat* is both recognized and pronounced when visually presented, and depicts the lexical representation in its broadest sense. That is, it describes the representations required to access the word during reading. Representational units at the form level are sublexically based, while units at the function level represent componential semantic and syntactic features and any other information relevant to the word. A lemma can be seen, then, as the unit that brings these components together. So, the lemma for "cat" represents the union of the graphemes *c, a*, and *t* (in that order) and the phonemes /k/, /æ/, and

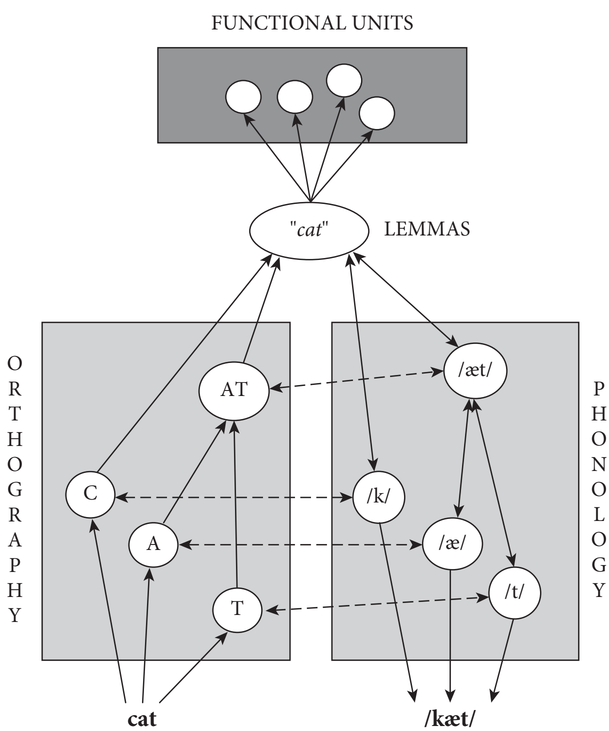


Fig. 8.1 Example of how *cat* is represented in the AUSTRAL model.

/t/ with the notion of a countable concrete noun that has the semantic features relevant to the concept of "cat."

Although neutral with regard to the way in which the functional information might be represented, the AUSTRAL model specifies the nature of the sublexical form units. In particular, there has been strong evidence that consonantal onsets are treated separately from the rest of the syllable when reading monosyllabic English words (e.g., Andrews & Scarratt, 1998; Taraban & McClelland, 1987; Treiman & Chafetz, 1987). The rest of the syllable is referred to as the *body* of the word (or orthographic *rime*), and is composed of a vowel plus consonantal coda if there is one (e.g., *str* is the onset of *street*, and *eet* is its body, comprising vowel *ee* and coda *t*). Thus the sublexical units of the AUSTRAL model form a hierarchy whereby grapheme units activate body units (e.g., the graphemes A and T activate the body AT), and the lemma is then activated through the combination of the onset and body. It is this "activation using structurally tiered representations and lemmas" that characterizes the model and creates the acronym AUSTRAL.[1]

By incorporating structurally tiered orthographic representations linked to phonological units of a corresponding structure, the AUSTRAL model embodies lexical and sublexical information in a different way from dual-route models of reading aloud such as DRC (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) and CDP++ (e.g., Perry, Ziegler, & Zorzi, 2010), where lexical and sublexical information are processed through different pathways. Like the PDP model, AUSTRAL pronounces nonwords via the same units through which words are named, except that nonwords lack the support that words receive from the lemma level. Taft (1991, 2006) discusses the way in which AUSTRAL readily explains why regular words (e.g., *pink*) are named with a shorter latency than irregular words (e.g., *pint*) and why regular words with a body that is inconsistently pronounced (e.g., *hint*) also show slower naming latencies. However, the focus of this chapter is not on the generation of sound from print but on the identification of visually presented words. So discussion will now be restricted to the way in which a visually presented letter string is represented for the purposes of recognition during reading.

Information about the whole word is found at the lemma level, which potentially makes it the locus of lexical representation in its narrowest sense. That is, words can potentially be discriminated from nonwords at the lemma level because only words have developed a unit linking form with function. However, this is not entirely true because, as will be explained later, there are real words that might not be represented by a lemma and there are also nonwords that might be. To elucidate, we need to consider how morphological structure is represented in the lexical processing system (see also Hyönä, this volume). As will be seen, the way in which morphemically complex words are processed is central to our understanding of lexical representation.

Morphological Processing

A morpheme is usually defined as the smallest unit of form associated with a semantic or syntactic function. Given that a lemma is a unit that encapsulates the association between form and function, it follows that the lemma level captures morphemic structure and drives the processing of morphemically complex words (e.g., *cats, unfriendly, daydream*). How are such polymorphemic words recognized?

Obligatory Decomposition in the AUSTRAL Model

In accord with its notion that the form level represents components of the whole word, the AUSTRAL model has all polymorphemic letter strings being decomposed into their apparent morphemes for recognition to take place (e.g., *un, friend*, and *ly*). This idea of obligatory decomposition has been around since Taft and Forster (1975), and has been supported in more recent times by research using the masked priming paradigm. As overviewed by Rastle and Davis (2008), many experiments have shown not only that the masked presentation of a transparently derived word (e.g., *hunter*) facilitates subsequent recognition of its stem (*hunt*) but also that the same is true of a pseudoderived word (e.g., *corner-corn*). Because no facilitation is reported when the prime does not include a putative suffix (e.g., *turnip-turn*), it is concluded that a pseudoderived word is blindly decomposed into its apparent morphemes (e.g., *corn* and *er*). Only at a later stage is this analysis overturned so that the pseudoderived word is correctly treated as a monomorphemic word. How then is a polymorphemic word recognized after the form representations of its component morphemes are accessed? Taft (2003, 2004) and Taft and Nguyen-Hoan (2010) argued that the lemmas for each of the component morphemes are activated via their form representations and then there are two possible ways in which the whole word might be recognized.

First, if the function of the polymorphemic word is entirely transparent with respect to the function of its component morphemes, as is typically the case for regularly inflected words (e.g., *cats, jumped, eating*), the whole word can be recognized purely on the basis of the functional information associated with each morpheme lemma. The stem and the affix each have their own lemma. For example, once it is known what a *cat* is and that the suffix *s* can denote the plural of a countable noun, everything that is known about the word *cats* can be determined. So, as in a printed dictionary, there is no need for a whole-word representation in the mental lexicon, because it would be redundant. These are the types of real words referred to earlier that would not be represented at the lemma level in the model, but that can nonetheless be recognized on the basis of functional information.

Second, and in contrast to words whose morphological composition is entirely transparent, a polymorphemic word that has any semantic or grammatical function that cannot be determined on the basis of its component morphemes must be represented by a whole-word lemma to provide a link to that idiosyncratic information. For example, there needs to be a lemma for *friendly* to associate it with the functional knowledge that it means more than just "characteristic of a friend." Similarly, the present participle of *meet* (as in *I am meeting her for the first time*) is entirely understandable from the combination of the functions associated with the stem and affix lemmas and, hence, does not require a lemma, but there does need to be a whole-word lemma for *meeting* when used as a gerund (as in *we'll hold a meeting tonight*) in order to understand that it specifically means "an assembly of people for the purposes of discussion." Taft (2003, 2004) and Taft and Nguyen-Hoan (2010) propose that such a whole-word lemma is activated via the lemmas for its component morphemes, creating a hierarchy of lemmas from monomorphemic to polymorphemic.

Support for the existence of such a hierarchy of lemmas is presented by Taft and Nguyen-Hoan (2010) from a masked priming experiment with ambiguous targets (e.g., *stick* meaning either "a twig" or "adhere"). The results showed that when asked to provide the meaning of the target, participants were biased by the meaning suggested by a prime that was an affixed version of the target. For example, more participants gave the "adhere" meaning of *stick* (as opposed to the "twig" meaning) when preceded by the masked prime *sticky* than when preceded by an unrelated word. Moreover, there was no such bias when the prime was only semantically related to that meaning without being a morphological variant of the target (e.g., the word *glue*). This lack of a bias toward the "adhere" meaning of *stick* when *glue* was the prime indicates that the observed bias to the "adhere" meaning when *sticky* was the prime could not have arisen solely at the semantic level. Neither could its locus be the form level because, logically, the two versions of a homograph are not differentiated at that level, being identical in form. It was therefore concluded that the locus of meaning bias must have been a level that mediates between form and semantics, namely, the lemma level. So, the lemma for *sticky* is activated via the lemma for only one version of *stick* (i.e., the "adhere" version) and, when *sticky* is presented as the prime, that version remains active when the target arrives, hence biasing the response.

Bound Morphemes

If morphemes are represented at the lemma level because they capture the correlation between form and function, this should be equally true whether the morpheme is free or bound, that is, whether or not it can stand as a word in its own right. Affixes are the typical bound morphemes (e.g., *un, y, ing*), but some stems are also bound. For example, *venge* cannot stand on its own as a word, yet it occurs in *revenge, avenge, vengeful*, and *vengeance*, which clearly have overlapping meanings. Therefore, it is argued (see Taft, 2003) that *venge* develops a lemma to capture this form-meaning correlation, through which the whole-word lemmas for *revenge, avenge*, and so on are activated.

With nonwords existing at the lemma level (i.e., when they are bound morphemes), it cannot be the case that lexical decision responses are made purely on the basis of there being a lemma corresponding to the presented letter string. While classifying a bound stem as a nonword is certainly difficult (e.g., Taft, 1994; Taft & Forster, 1975), it is nevertheless possible to do so, and the AUSTRAL model needs to explain how. The simplest explanation is that there is information linked to the bound-stem lemma that stipulates that it cannot be used as a word in its own right; information that would be particularly important when it comes to production. It might be the case that this information takes the form of a further level that represents lexical concepts, as has been proposed in relation to speech production (e.g., Levelt, Roelofs, & Meyer, 1999). That is, only free-standing words, be they monomorphemic or polymorphemic, correspond to holistic concepts. Bound morphemes do not and can, therefore, be rejected as words on that basis.

Another possibility is that lemmas vary in some way as a function of the form-meaning correlation that they capture. A systematic relationship between form and function is likely to be most obvious when the form retains its meaning within a variety of contexts, because the constancy of the relationship contrasts with the variability of the linguistic information surrounding it. As such, a real word will have a stronger correlation than will a bound morpheme. Although a bound morpheme might recur in several different contexts (i.e., with different affixes attached), these will be fewer than the number of contexts in which a real word can recur (i.e., all the sentences in which that word is encountered). Therefore, there might be a threshold of correlation above which the letter string is classified as a word. Of course, if such an argument were to be pursued, the mechanism by which a lemma is able to vary on the basis of form-meaning correlation would need greater specification.

Morphological Decomposition in Other Models

While other word recognition models that focus on morphological processing also incorporate the notion of a lemma level, they differ from AUSTRAL in a number of ways.

SCHREUDER AND BAAYEN (1995)

As in AUSTRAL, Schreuder and Baayen (1995) propose the separation of access representations from lemmas (which, prior to Baayen et al., 1997, were labeled as "concepts"; see also Taft, 1991). However, according to Schreuder and Baayen (1995), online decomposition only occurs at the earliest stages of acquiring a new polymorphemic word. Otherwise, polymorphemic words are identified through a whole-word access representation that activates either a lemma corresponding to the whole polymorphemic word or lemmas corresponding to its component morphemes, depending on how transparently related those morphemes are to the whole word. Therefore, apart from newly experienced polymorphemic words, the only decomposition that occurs in the model of Schreuder and Baayen (1995) is at the lemma level after whole-word access, and only for some words.

However, such a notion of postlexical activation of constituent morphemes (as also proposed by Burani & Caramazza, 1987, and Giraudo & Grainger, 2000, 2001) fails to explain the pseudoderived masked priming effect (e.g., *corner* priming *corn*) that was described earlier (see Rastle & Davis, 2008). That is, according to this account, at no point in its recognition is *corner* ever decomposed into *corn* and *er*, because its whole-word access representation only activates a whole-word lemma. Therefore, there is no reason for the processing of *corner* to influence the processing of *corn* other than through orthographic overlap, in which case *turnip* should equally prime *turn*. For this reason, other models have incorporated early morphological decomposition based purely on form in order to tap into a so-called *morpho-orthographic level* of representation.

DIEPENDAELE, SANDRA, AND GRAINGER (2009)

The model proposed by Diependaele, Sandra, and Grainger (2009) has two levels based on form; the morpho-orthographic level where *hunter* is represented by *hunt* and *er*, and the lexical form level where all words are represented, including *hunter, hunt*, and *corner*. Lexical form representations for polymorphemic words are activated both via the decompositional pathway that is mediated by the relevant morpho-orthographic units (e.g., *hunt* and *er*), and directly from the letter string without mediation. It is through the former pathway that *corner* will prime *corn*, since the lexical form of *corn* will be preactivated via its morpho-orthographic unit that is inadvertently activated when *corner* is blindly decomposed. The difference between the processing of a pseudoderived and truly derived word is that the lexical form of such words (e.g., *hunter, corner*) receives activation from the morpho-orthographic level when it has a true stem (e.g., *hunt*), but not when it has a pseudostem (e.g., *corn*).[2] Links between the morpho-orthographic level and lexical form level arise from feedback from a higher *morpho-semantic level* where words are represented as morpheme units (such that the lexical form unit for *hunter* activates the morpho-semantic units for *hunt* and *er*, but the lexical form unit for *corner* does not activate semantic units for *corn* and *er*). As such, the morpho-semantic units function as morpheme-based lemmas, with information about the whole polymorphemic word only being found at the lexical form level. Therefore, the model incorporates both a prelexical decomposition pathway (i.e., based on sublexical information) and a postlexical decomposition pathway (i.e., based on lexically stored information), with lexical decision centering on the existence of an intermediate lexical form representation. However, if a word can be recognized through direct access to the whole lexical form, what is the purpose of prelexical decomposition? If it somehow makes access to the lexical form of a complex word easier than whole-word access, what then is the purpose of postlexical decomposition?

Crepaldi, Rastle, Coltheart, and Nickels (2010)

Diependaele et al. (2009) differentiate the early processing of truly derived and pseudoderived words in order to capture the apparent fact that masked priming is stronger for the former than the latter. In contrast, Crepaldi, Rastle, Coltheart, and Nickels (2010) maintain that the magnitude of masked priming is not significantly different for truly derived and pseudoderived words, though only the former generate priming when unmasked and, therefore, the two types of words are only distinguished at a late semantic stage. According to Crepaldi et al. (2010), morpho-orthographic representations are activated via obligatory decomposition and these, in turn, combine to activate the form representation for the whole word (at a level referred to as the *orthographic lexicon*). While such a decompositional pathway also features in the Diependaele et al. (2009) account, Crepaldi et al. (2010) specify that activation of the whole-word form is mediated by the component morphemes regardless of semantic transparency. As seen in Figure 8.2, the morpho-orthographic representations *CORN* and *ER* activate the orthographic lexical unit for *CORNER* in exactly the same way that the morpho-orthographic representations *HUNT* and *ER* activate the orthographic lexical unit for *HUNTER*.

According to Crepaldi et al. (2010), activation from the orthographic lexicon passes to a lemma level and then on to the semantic system. However, unlike in AUSTRAL, derivationally related words do not share a lemma. This means that the relationship between the lemmas for *corner* and *corn* is exactly the same as that for *hunter* and *hunt*. It is only in the semantic system that the two types of words differ, because the latter have overlapping semantic features and the former do not. The purpose of the lemma level in the Crepaldi et al. (2010) account is solely to capture the relationship between inflectionally related words, where such words share a lemma regardless of whether the inflection is regular (e.g., *cats* and *cat*) or irregular (e.g., *fall* and *fell)*.

Irregularly Inflected Words and Whole-Word Form Representation

The major motivation for Crepaldi et al. (2010) to include an orthographic lexicon in their model is to capture the fact that real inflected words (e.g., *jumped*) can be distinguished from nonwords composed of a real stem and affix (e.g., *falled, sheeps*). The latter have no representations in the orthographic

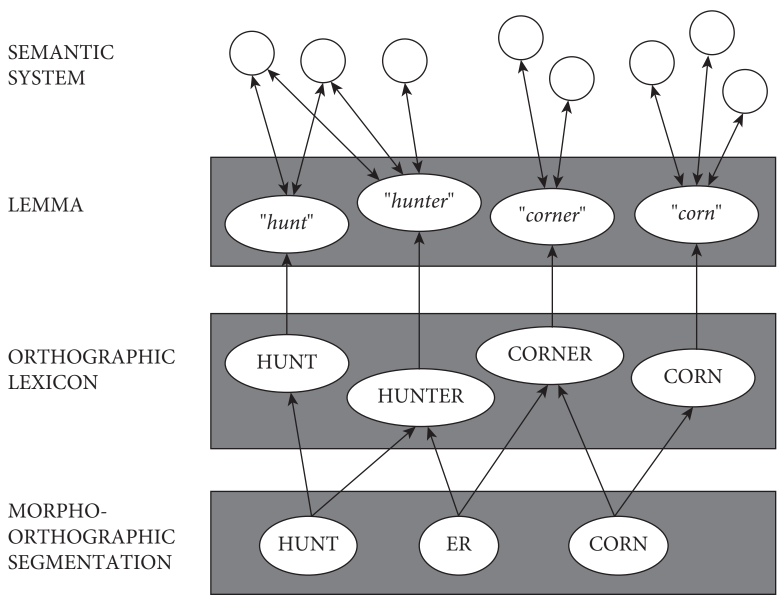


Fig. 8.2 The representation of derived and pseudoderived words according to Crepaldi et al. (2010).

lexicon, while real inflected words do. Irregularly inflected words (e.g., *fell, taught, sheep, teeth*) are not decomposed at the form level, but activate the lemma for their stem, and this is true in AUSTRAL as well (see also Allen & Badecker, 2002). Presentation of either *fell* or *fall* will directly activate the lemma for *fall*. According to AUSTRAL, the former will also activate the lemma corresponding to the past tense just as the suffix *ed* would (see Taft, 2003), because otherwise there would be no way to distinguish *fell* from *fall*. With words being identified on the basis of information associated with the lemmas in the AUSTRAL model, what stops *falled* from being recognized as a word, given that it will activate lemmas that can be combined on the basis of functional information (e.g., *fall* is a verb that can take the past tense)?

The way the AUSTRAL model can handle this is by simply having the lemma for the stem being explicitly associated with information stipulating that the word does not follow regular inflectional patterns. That is, a "yes" response could ultimately be avoided if *falled* were presented because, after decomposition, information associated with the lemma for *fall* would specify that its past tense is actually *fell* or, more generally, that the regular inflection *ed* is not appropriate for this word. So against the claim of Crepaldi et al. (2010), it is possible for the AUSTRAL model to distinguish all words and nonwords without the need for a form-based lexicon that includes all possible words, including inflected ones.

However, there is a further issue in relation to irregularly inflected words that has implications for the nature of form-based representations. In the description of the AUSTRAL model as presented here and in Taft (1991, 2006), the subsyllabic units of onset and body are depicted as the highest level of form representation. At other times, though, the model has been presented with whole-word form representations, at least when the words are monosyllabic (e.g., Taft, 2003, 2004; Taft & Nguyen-Hoan, 2010). In other words, whether or not the highest level of form representation corresponds to the whole word has been an open-ended aspect of the model. However, the proposed account of irregular word processing seems to necessitate the existence of a whole-word form representation, or at least a whole-syllable representation. The only way in which an irregularly inflected word can activate a lemma associated with its relevant syntactic function (e.g., the past-tense lemma when *fell* is presented or the plural lemma when *teeth* is presented) is if there is a whole-word form representation that can be linked to that lemma. It is the whole-word form *FELL* that is associated with the past-tense lemma, rather than either its onset *F* or its body *ELL*. If it were the body *ELL* that activated the past-tense lemma, this would happen not only when *fell* was presented, but when any other word that contains *ell* was presented, such as *yell* or *spell*. Obviously, it would be inappropriate to activate information about the past tense when the word is not actually a past-tense verb. Therefore, in a localist model such as AUSTRAL it seems necessary to allow for whole-word form representations, even if not for all words (i.e., not for regularly affixed words, or possibly even polysyllabic words; see section "Representation of Polysyllabic Words").

Graded Effects of Morphological Relatedness

There is a further aspect of morphological structure that is informative with regard to lexical representation. The transparency of whether there is a derivational relationship between an affixed word and its stem is not an all-or-none property of those words. It is clear that *hunter* and *hunt* are derivationally related (i.e., a hunter is "someone who hunts"), while *corner* and *corn* are not. However, many cases show a partial relationship, such as *archer* and *arch* where the shape of the latter is captured in the bow used by the former, or *hearty* and *heart* where a metaphorical sense of the latter seems to be contained in the former (as it also is in *heartfelt* and *wholehearted*). Speakers are sensitive to such a gradation in derivational transparency both when asked to rate semantic relatedness and in the degree of facilitation of lexical decision responses when the stem is preceded by an unmasked version of the complex word that contains it (e.g., Gonnerman et al., 2007). How is such a continuum of transparency to be incorporated into models of lexical processing?

Gonnerman et al. (2007) argue that the transparency continuum is expected within a PDP model because hidden units capture the relationship between form and meaning, and the more transparently related two words are in both form and meaning, the greater the overlap in their pattern of activation within those hidden units. This overlap provides the basis for a gradation in priming in line with derivational transparency. According to such an account, there is no need for a morpho-orthographic stage in which letter combinations that correspond to affixes are blindly stripped, because patterns of activation corresponding to different morphological structures are entirely encapsulated within the hidden units. Gonnerman et al. (2007) support such a claim by drawing on an unpublished masked priming study by Gonnerman and Plaut (2000) where pseudoaffixed words failed to prime their pseudostems (e.g., *corner* not priming *corn*). However, more recent research has clearly shown masked priming for all levels of transparency including pseudoderivations (e.g., Marslen-Wilson, Bozic, & Randall, 2008; Rastle & Davis, 2008), an outcome that seems incompatible with the PDP account as it stands (though see Rueckl & Aicher, 2008, for speculation as to how it might come about). The masked priming observed for pseudoderived words seems most readily explained by the existence of a stage of decomposition that is blind to semantic factors, namely, a morpho-orthographic stage.

Can a graded effect of transparency in ratings and unmasked priming be handled within a model that also includes morpho-orthographic processing? Certainly, the Crepaldi et al. (2010) account has no problem incorporating the idea of graded feedback from the semantic level depending on the relationship between the complex word and its stem. Such feedback is the only source of differentiation between transparently and opaquely derived words and comes into play when the complex word is more fully processed (as in the unmasked priming paradigm).

The idea of graded effects of transparency is more of an issue for the AUSTRAL account. As depicted in Figure 8.3, Taft and Nguyen-Hoan (2010) suppose that a transparently derived word (e.g., *hunter*) is activated through the lemma for its stem (*hunt*), whereas a pseudoderived word (e.g., *corner*) and its pseudostem (*corn*) compete with each other at the lemma level. The greater the opportunity for the inhibitory impact of competition to come into play, the greater its counteracting impact on the morpho-orthographic facilitation that occurs at the form level. Therefore, under unmasked conditions, pseudoderived words will show no facilitation, unlike transparent words where priming is generated at both the form and lemma levels. With such a qualitative difference between the representations of the two types of words, though, how can partially related words (e.g., *archer* and *arch*) show a graded effect of priming in the unmasked priming paradigm?

One way to capture this within the AUSTRAL framework would be to propose that lemmas that are activated via the same form representation are always linked, but with weightings that range from zero to a strongly positive value, depending on feedback based on their semantic overlap. So, the lemmas for *corn* and *corner* are linked, but with a minimal weighting. When *corner* is presented as the prime, the lemma for *corn* will be activated, but will send negligible activation to the lemma for *corner,* which is activated directly from the form level. In turn, the lemma for *corner* will provide negligible support to the lemma for *corn*, which effectively places the two lemmas in competition with each other. In contrast, the lemma for *hunt* will send positive activation to the lemma for *hunter* and vice versa, as will the lemmas for *arch* and *archer*, but to a lesser degree. When the prime is unmasked, there will be sufficient opportunity for the impact of the relationship between the two lemmas to modulate the effect of priming. In this way, the graded effect of partial transparency in the unmasked priming task can be handled within the AUSTRAL model.

With reference to Figure 8.3, the inclusion of a weighted link between the lemmas for a complex word and those for its component morphemes would mean that there is an additional link between

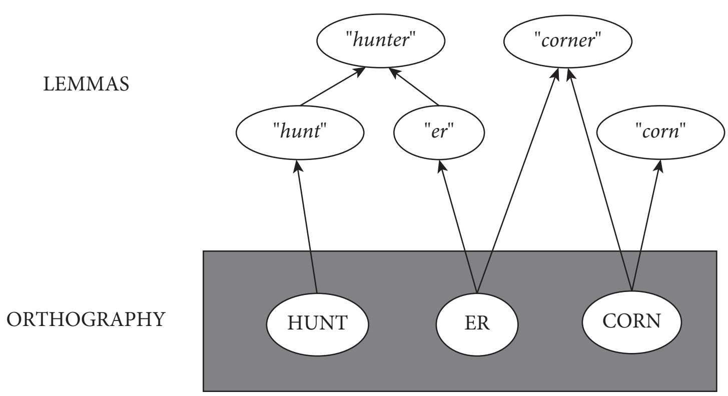


Fig. 8.3 The AUSTRAL representation of derived and pseudoderived words.

the lemma for *corn* and that for *corner*. However, being so weak, it would lose out to the pathway that links the *corner* lemma directly to the form units *CORN* and *ER*. If conceptualized in this way—namely, as a competition between a direct form-to-lemma link and a mediated lemma link—the same should hold for all complex words, including truly derived words. That is, not only is the lemma for *hunter* linked to the lemma for *hunt* but also it receives direct activation from the form units for *HUNT* and *ER*. In this case, however, the positive link between lemmas makes it a more effective pathway than the direct link from the form level and, therefore, recognition of the transparently derived word will typically be based on the mediated pathway.

The idea of competing pathways has been previously proposed in relation to morphological processing (e.g., Baayen et al., 1997; Bertram, Schreuder, & Baayen, 2000; Colé, Beauvillain, & Segui, 1989; Diependaele et al., 2009; Niswander, Pollatsek, & Rayner, 2000). In those accounts, though, the two possible pathways correspond to decomposition versus whole-word access. In contrast, form-based decomposition always occurs in AUSTRAL. It is only after decomposition that access to the whole-word representation might follow two competing pathways: one direct from the form-based components and one via the lemmas corresponding to those components. Thus flexibility can be introduced into the AUSTRAL model, with the overlap in semantic features determining which postdecompositional pathway is stronger.

Relationship Between Reception and Production

The models described in relation to the reading process address the lexical mechanisms involved in getting from print to meaning, and similar mechanisms are likely to underlie the recognition of spoken words once phonetic variation is taken into account. It would also make sense to propose that the same lexical mechanisms are involved in speech production, although running in the reverse direction. In fact, AUSTRAL owes much to the models of speech production proposed by Dell (1986) and by Levelt et al. (1999). The Dell (1989) model has a level of morpheme representation (equivalent to the lowest level of the lemma hierarchy in AUSTRAL) above which is a word level that includes derivationally complex words (potentially equivalent to the higher level of the lemma hierarchy in AUSTRAL) but does not include inflected words, which are instead activated through the functionally determined combination of the stem and affix, as in AUSTRAL. Furthermore, the highest level of form representation in the Dell model comprises syllables, which, in turn, activate their component onsets and rimes prior to activating their phonemes. This is the same as the form level of the AUSTRAL model in reverse, albeit phonologically based rather than orthographic.

The existence of a lemma level is a major feature of the WEAVER++ model of Levelt et al. (1999), but there is also a separate level that represents lexical concepts. The production of a word begins with activation of a semantically based concept, which, in turn, activates a lemma. The lemma is seen as a link to syntactic information about the word. In reception models such as AUSTRAL, little emphasis has been placed on this distinction between semantic and syntactic information. The reason for this is that when passing from the form level to the function level (as is the case in word recognition), the lemma provides direct links to both syntax and meaning and their order of activation is immaterial. As such, the distinction between the locus of lexical semantics and lexical syntax may be critical for speech production, but has little impact on lexical processing in reading.

Future Directions

Impact of Derivational Transparency in Masked Priming

The equivalence of masked priming for truly derived and pseudoderived words would seem to be critical for the Crepaldi et al. (2010) model, since the two types of words are treated in exactly the same way throughout much of their processing. However, variable results have been observed with regard to this issue. Although Rastle and Davis (2008) concluded from their overview of such research that there is no difference in priming between transparently derived and pseudoderived words, a meta-analysis indicated that priming might be greater the former than the latter (cf. Feldman, O'Connor, & Moscoso del Prado Martín, 2009; Taft & Nguyen-Hoan, 2010; but see Davis & Rastle, 2010). If this is indeed the case, it would be hard to maintain the Crepaldi et al. (2010) account. Therefore, it is important to establish whether a genuine difference between transparent and pseudoderived priming can be found under masked conditions. The AUSTRAL account is flexible with regard to this question because whether or not priming is greater for transparently derived than pseudoderived words will depend on whether competition in the latter case has the opportunity to come into play.

Dual Pathways From Form to Lemma

According to the modification to the AUSTRAL account, whereby graded effects of transparency arise from the differential use of mediated and direct links, there should also be graded effects in the impact of stem frequency on complex word recognition. If lemma activation is influenced by word frequency, recognition of a derived word will be influenced not only by its own frequency of occurrence but also by the frequency of its stem, if recognition is mediated through the lemma for that stem. The impact of stem frequency has been well documented (e.g., Baayen et al., 1997; Bertram et al., 2000; Colé et al., 1989; Niswander et al., 2000; Taft, 1979b, 2004; Taft & Ardasinski, 2006), so it would make sense to examine further whether the strength of this effect varies as a function of transparency. It should do so if semantic transparency determines which of the competing pathway succeeds, because only lemma mediation will be influenced by stem frequency. The Crepaldi et al. (2010) model would expect stem frequency effects regardless of transparency because access to the whole-word representation is the same for all types of words.

Representation of Polysyllabic Words

It can be seen from Figure 8.3 that AUSTRAL breaks down monomorphemic words like *corner* into subunits at the form level. While it is possible that this only happens when the structure of the word is morphologically complex in its appearance, it is more parsimonious to assume that all polysyllabic words are similarly broken down, even if, unlike *CORN* and *ER*, each of the form-based subunits is not associated with its own lemma. For example, the lemma for *walrus* might be activated directly from orthographic units representing the meaningless syllables *WAL* and *RUS* (e.g., Taft & Krebs-Lazendic, 2013). Where does the syllable boundary fall in cases that are potentially ambiguous in this regard? For example, *sermon* could be orthographically broken down into *ser* and *mon* in correspondence to the way it is pronounced, or, alternatively, the informativeness of its first subunit might be increased by breaking it down into *serm* and *on*, maximizing the coda of the first syllable. The latter analysis has been proposed by Taft (1979a, 1987), Taft and Kougious (2004), and Taft and Krebs-Lazendic (2013), with the maximized first syllable (e.g., *serm*) being referred to as the basic orthographic syllabic structure (or BOSS), though such an idea has by no means found wide support (e.g., Katz & Baldasare, 1983; Lima & Pollatsek, 1983; Perry, 2013).

If all polysyllabic words are represented at the form level as subunits that correspond to the maximal coda analysis, a word like *turnip* will be represented at that level as *TURN* and *IP*. As such, it might be expected that *turnip* will facilitate responses to *turn* in the masked priming paradigm despite the fact that *ip* does not have the appearance of a morpheme. The fact that such an orthographic condition has not shown masked priming effects in previous studies is addressed by Taft and Nguyen-Hoan (2010). They point out that the items used in that condition have actually been a mixture of cases where the target is the BOSS (as in *turnip-turn, brothel-broth*) and where it is not (as in *freeze-free, shunt-shun*). Therefore, it has not yet been shown that masked priming is absent when the target is specifically the BOSS of the prime (as in *turnip-turn*) and this is something that future research could pursue.

Conclusions

The purpose of this chapter has been to explore how models of visual word recognition envisage the way in which words are represented in lexical memory. As a working definition, lexical representation was taken to embrace all information required to establish that a presented letter string corresponds to a known word. If such a word has a representation that a nonword cannot have, then access to this representation should be sufficient to recognize the letter string as a word and, indeed, as that particular word. For this reason, a number of models (e.g., Crepaldi et al., 2010; Diependaele et al., 2009; Schreuder & Baayen, 1995) incorporate form-based representations for every known word, no matter what type of word it is, and this provides the basis for discriminating real words from nonwords. Nevertheless, in some accounts when the whole-word representation corresponds to a polymorphemic word, it is accessed via form representations for its component morphemes (e.g., Crepaldi et al., 2010), or at least it can be (e.g., Diependaele et al., 2009).

In contrast, the AUSTRAL model highlighted in this chapter has a form level that only represents single syllables, whether these create a whole word (e.g., the *hunt* of *hunter* or the *corn* of *corner*) or not (e.g., *er*). The recognition of a letter string as a particular word therefore takes place at a level beyond that of form. The lemma level mediates between form and function and brings together the syllables represented at the form level. Moreover, lemmas are hierarchically structured whereby a derivationally complex word (e.g., *hunter*) has its own lemma activated via the lemmas for its component morphemes (*hunt* and *er*). The lemma level therefore provides the primary locus of lexical representation, though with certain caveats attached.

One caveat is that any affixed word whose meaning is entirely predictable from its components (e.g., *jumped, cats*) is not actually represented at the lemma level, but rather is recognizable through the combination of higher level functional information associated with the lemma for each of its morphemes (i.e., semantic, syntactic, and/or idiosyncratic features of the morpheme). When a word activates such combinable functions, but has an exceptional form (e.g., *fell, teeth*), information about its distinctive characteristics can also be found at the higher functional level. This prevents a regularization of the irregular word (e.g, *falled, tooths*) from being accepted as a word.

Second, the lemma level includes representations of bound morphemes (e.g., *venge, er*) which, by definition, are not words in their own right. Therefore, there needs to be some way of discriminating such morphemes from real words. Suggestions for achieving this can be given in terms of specific information stored at the higher functional level, the existence of a concept level, or differing degrees of form-function correlation captured at the lemma level.

We see then that the lexical representations involved in reading can be conceptualized in a number of different ways ranging from orthographic units corresponding to the whole word to patterns of activation within a distributed set of units that mediate between form and function. The account favored in this chapter, AUSTRAL, is one where information about a word is associated with a localist unit that mediates between form and function (i.e., a lemma), while being activated through sublexical units at the form level. Such an account provides a concrete framework for helping understand the processes involved in reading all types of words, both in terms of retrieving their meaning and in generating their pronunciation. An understanding of the recognition of polymorphemic words is thus not just a niche topic within the domain of lexical processing but has important implications for the conceptualization of the whole lexical processing system.

Notes

1 In Taft (2006) the model was referred to as *localist-cum-distributed* (LCD). The reason for abandoning this term is that most models actually include a combination of localist and distributed characteristics, even if one dominates the other. For example, the input units of PDP models are typically localist, representing specified letters or letter groupings, while the letter units of the IA model can be seen as being distributed in the sense that more than one word is activated through the same set of units.

2 In fact, Diependaele et al. (2009) do not explicitly state that the lexical form *corner* is not activated through the morpho-orthographic unit *corn*. However, this seems a sensible conclusion based on the description they give.

References

Allen, M., & Badecker, W. (2002). Inflectional regularity: Probing the nature of lexical representation in a cross-modal priming task. *Journal of Memory and Language, 44,* 705–722.

Andrews, S., & Scarratt, D. R. (1998). Rule and analogy mechanisms in reading nonwords: Hough dou peapel rede gnew wirds? *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1052–1086.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language, 37,* 94–117.

Bertram, R., Schreuder, R, & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homophony, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 489–511.

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes, 2*, 217–227.

Colé, P., Beauvillain, C., & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language, 28,* 1–13.

Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Crepaldi, D., Rastle, K., Coltheart, M., & Nickels, L. (2010). “Fell” primes “fall,” but does “bell” prime “ball”? Masked priming with irregularly-inflected primes. *Journal of Memory and Language, 63,* 83–99.

Davis, M. H., & Rastle, K. (2010). Form and meaning in early morphological processing: Comment on Feldman, O’Connor, and Moscoso del Prado Martin (2009). *Psychonomic Bulletin & Review, 17,* 749–755.

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93,* 283–321.

Diependaele, K., Sandra, D., & Grainger, J. (2009) Semantic transparency and masked morphological priming: The case of prefixed words. *Memory & Cognition*, *37,* 895–908.

Feldman, L. B., O’Connor, P. A., & Moscoso del Prado Martín, F. (2009). Early morphological processing is morpho-semantic and not simply morpho-orthographic: An exception to form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review, 16,* 684–691.

Forster, K. I. (1976). Accessing the mental lexicon. In E. C. J. Walker & R. J. Wales (Eds.), *New approaches to language* *mechanisms* (pp. 257–287). Amsterdam, the Netherlands: North-Holland.

Forster, K. I. (2012). A parallel activation model with a sequential twist. In J. S. Adelman (Ed.), *Visual word recognition: Vol. 1. Models and methods, orthography and phonology* (pp. 52–69). Hove, England: Psychology Press.

Giraudo, H., & Grainger, J. (2000). Prime word frequency in masked morphological and orthographic priming. *Language and Cognitive Processes, 15,* 421–444.

Giraudo, H., & Grainger, J. (2001). Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin & Review, 8,* 127–131.

Gonnerman, L., & Plaut, D. C. (2000). *Semantic and morphological effects in masked priming (2000).* Department of Psychology. Paper 350. http://repository.cmu.edu/psychology/350.

Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General, 136,* 323–345.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111,* 662–720.

Katz, L., & Baldasare, J. (1983). Syllable coding in printed-word recognition by children and adults. *Journal of Educational Psychology, 75,* 245–256.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14,* 185–209.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–38.

Lima, S. D., & Pollatsek, A. (1983). Lexical access via an orthographic code? The Basic Orthographic Syllabic (BOSS) reconsidered. *Journal of Verbal Learning and Verbal Behavior, 22,* 310–332.

Marslen-Wilson, W., Bozic, M., & Randall, B. (2008). Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. *Language and Cognitive Processes, 23,* 394–421.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165–178.

Morton, J. (1970). A functional model of memory. In D. A. Norman (Ed.), *Models of human memory* (pp. 203–254). New York, NY: Academic Press.

Murrell, G. A., & Morton, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology, 102,* 963–968.

Niswander, E., Pollatsek, A., & Rayner, K. (2000). The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes, 15,* 389–420.

Perry, C. (2013). Graphemic parsing and the basic orthographic syllable structure. *Language and Cognitive Processes, 28,* 355–376.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive Psychology, 61,* 106–151.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes, 23,* 942–971.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42,* 107–142.

Rueckl, J. G., & Aicher, K. A. (2008). Are Corner and Brother morphologically complex? Not in the long term. *Language and Cognitive Processes, 23,* 972–1001.

Rueckl, J. G., & Raveh, M. (1999). The influence of morphological regularities on the dynamics of a connectionist network. *Brain and Language, 68,* 110–117

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. *Psychological Review, 89,* 60–94.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–154). Hillsdale, NJ: Erlbaum.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Taft, M. (1979a). Lexical access via an orthographic code: The Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behavior, 18,* 21–39.

Taft, M. (1979b). Recognition of affixed words and the word frequency effect. *Memory & Cognition, 7,* 263–272.

Taft, M. (1987). Morphographic processing. The BOSS re-emerges. In M. Coltheart (Ed.), *Attention and performance: Vol. 12. The psychology of reading* (pp. 265–279). London, England: Erlbaum.

Taft, M. (1991). *Reading and the mental lexicon*. Hove, England: Erlbaum.

Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes, 9,* 271–294.

Taft, M. (2003). Morphological representation as a correlation between form and meaning. In E. Assink, & D. Sandra (Eds.), *Reading complex words* (pp. 113–137). Amsterdam, the Netherlands: Kluwer.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology, 57A,* 745–765.

Taft, M. (2006). A localist-cum-distributed (LCD) framework for lexical processing. In S. M. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 76–94). Hove, England: Psychology Press.

Taft, M., & Ardasinski, S. (2006). Obligatory decomposition in reading prefixed words. *The Mental Lexicon, 1,* 183–199.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14,* 638–647.

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior, 15,* 607–620.

Taft, M., & Kougious, P. (2004). The processing of morpheme-like units in monomorphemic words. *Brain and Language, 90,* 9–16

Taft, M., & Krebs-Lazendic, L. (2013). The role of orthographic syllable structure in assigning letters to their position in visual word recognition. *Journal of Memory and Language, 68,* 85–97.

Taft, M., & Nguyen-Hoan, M. (2010). A sticky stick: The locus of morphological representation in the lexicon. *Language and Cognitive Processes, 25,* 277–296.

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26,* 608–631.

Treiman, R., & Chafetz, J. (1987). Are there onset- and rime-like units in printed words? In M. Coltheart (Ed.), *Attention and Performance: Vol. 12. The psychology of reading* (pp. 281–298). London, England: Erlbaum.

CHAPTER

9

Are Polymorphemic Words Processed Differently From Other Words During Reading?

Jukka Hyönä

Abstract

Across a variety of languages, many words comprise more than one meaning unit, or morpheme. In the present chapter, reading studies employing readers' eye movement registration are reviewed that examine how such polymorphemic words are identified. The reviewed studies have examined how compound words, derived words, and inflected words are identified. Studies are also reviewed that have investigated whether the meanings of polymorphemic words are constructed out of the meanings of their components. More generally, it is concluded that polymorphemic words are identified during reading both using whole-word representations available in the mental lexicon (the holistic route) and accessing the word identity via the component meanings (the decomposition route). Moreover, word length plays a significant role in modulating the relative dominance of the two access routes, with the decomposition route being more dominant for long polymorphemic words.

Key Words: morphology, eye movements, word identification, compound word, inflected word, derived word, meaning computation

The present chapter reviews studies that have examined how polymorphemic words are identified during reading. Does their identification qualitatively depart from that of single morpheme words? Does their morphological structure play a significant role in their identification? These questions will be discussed in relation to reading alphabetic scripts. Languages vary greatly in terms of their morphological productivity. At one end of the continuum are the *agglutinative* languages like Finnish or Turkish, where the vast majority of words are morphologically complex. At the other end of the continuum lie the *analytic* languages, such as English, that represent morphologically less productive languages. At present, eye-tracking studies on reading polymorphemic words, the topic of the present chapter, are limited to four languages: English, Finnish, German, and Dutch. In terms of morphological productivity, Finnish is the most productive language among them and English is the least productive.

Written words in alphabetic writing systems consist of letters, which make up syllables, which, in turn, form lexical items that have entries in written dictionaries. Each string of letters is associated with one or multiple meanings. Lexical forms that are associated with a meaning are called *free morphemes.* They are free, as they can stand alone in text. *Bound morphemes* comprise the other category of morphemes. As the name indicates, they are bound to free morphemes but cannot appear independently. The three main categories of bound morphemes are *derivations,* bound *stems,* and *inflections.* Derivations change the grammatical category of the affected word. For example, in English, the derivational suffix ending *-ment* makes a noun out of a verb, as in *establishment*. A derivational prefix can also be attached to the word beginning (e.g., *re* in *re-establish*). Free morphemes can have both a prefix and a suffix attached to them, as in *re-establishment*, which is an example of a polymorphemic word with one free and two bound morphemes. Bound stems are stems that cannot stand alone (e.g., *flect* in *reflect* or *vive* in *revive*; see Taft & Forster, 1975, 1976; Taft, this volume). Finally, inflections in English modify a verb's tense (e.g., *-s* in *runs*) or a noun's number (e.g., *-es* in *beaches*). They appear at the word end. In many other languages, inflections are also used to mark the syntactic status of nouns. For example, in Finnish *kettua* is the accusative (called partitive case) form (denoting that the noun is the sentence object) of *kettu* ('fox'). Finally, free morphemes can be combined to form compound words, such as *football*. Typically, compound words consist of two free morphemes, but in languages where compounding is productive they may comprise three or four free morphemes. An example from German is *Datenschutzexpert* ('data security expert'), which consists of three free morphemes.

Does a word's morphological structure have consequences for word identification in reading? Intuitively, one may entertain either a view stressing the importance of single morphemes or a view ascribing little or no significance to the word's morphological structure. In support of the latter view, one could argue that as long as the reader's mental lexicon has an entry for a morphologically complex word, such as *establishment* or *football*, there is no need to pay attention to the word's morphological structure to secure its successful identification. This is not to deny that if needed, the morphological structure may be accessed afterward, after identification is completed (Giraudo & Grainger, 2000). On the other hand, one may entertain the view that morphologically complex words are identified via their morphemes by pointing to the possibility that identification is more readily achieved via smaller units. Thus for identifying the word *re-establish*, the reader is better equipped if the word can be decomposed into its morphological components before establishing the word meaning ('establish again') by putting together the morphological components (Taft, 1979). One could further argue that such morphological decomposition is particularly pertinent when the word is morphologically highly complex. With increased complexity, the word is likely to be long (e.g., *Datenschutzexpert*), which may diminish the possibility that it can be identified with a single glance. This in turn may encourage its identification via the morphological components.

The present chapter reviews studies of morphological processing during reading. The method of registering readers' eye movements has established itself as the gold standard in studying the reading process as it evolves through time and space. The method is attractive for two main reasons. First, as eye movements are a necessary part of the visual intake of information during reading, no secondary tasks extraneous to reading itself need be introduced to study the reading processes. Related to this, the reader is free to proceed in the text as he or she wishes. Second, the method provides its user with measures that tap the time course of processing (this will be discussed later in more detail). As will become evident in the present chapter, the technique has been successfully applied to study the role of morphemes in word identification during reading.

Readers' eye movements (for more details, see Schotter & Rayner, this volume) consist of two components: fixations and saccades. During fixations the eyes stay fixed on a location in text; saccades, in turn, are fast eye movements that take the eyes from one location in text to another. Intake of visual information takes place during fixations; vision is suppressed during saccades. Fixations in reading typically last for 200 to 250 ms; a forward saccade to a new text region extends about 7 to 10 letters to the right from a previous fixation location. The area around the fixation point from which useful visual information is extracted is constrained by the limitations of the eyes' foveal area, where the visual acuity is at its best. The word identification span is estimated to be around 12 letters around the fixation (4 letters to the left and 8 letters to the right of the center of fixation; Rayner, Well, Pollatsek, & Bertera, 1982).

The eye-tracking technique provides measures of where the reader is looking at in a text or a word and for how long these individual fixations last. Durations of fixations are used to tap into the time course of information processing. The duration of the initial fixation made on a word (*first-fixation duration*) provides an estimate of early processing on that word. *Gaze duration* is the summed duration of fixations made on a word during first-pass reading; that is, before fixating away from a word. Gaze duration reflects somewhat later processing than first-fixation duration; it is often used as the primary measure to index lexical access during reading. Finally, fixations returning to a word from a subsequent word reflect later processing efforts done after a word is identified. Several such measures may be gleaned from the eye movement record. *Total fixation time* is the summed duration of all fixations made on a word (i.e., it is a composite measure of early and later processing); the measurement of *go-past time* is initiated by a regression from the target word to a preceding word and terminated by a fixation going beyond the target word. As the present chapter deals with word identification, the primary focus is on first-pass measures (i.e., first-fixation duration and gaze duration).

The chapter starts by reviewing the evidence collected for identifying compound words during sentence reading (see also the reviews of Bertram, 2011; Pollatsek & Hyönä, 2006). This is done because the study of compound words is the most well studied aspect of morphological processing in reading. It is followed by studies investigating the effects of derivations and inflections on word identification during reading. The extensive literature on the role of morphemes in isolated word recognition is not discussed here (for a comprehensive review see Amenta & Crepaldi, 2012; Taft, this volume; for the seminal work on morphological processing see Taft & Forster, 1975, 1976). Nor do I discuss parafoveal morphological processing; that is, the possibility that morphological structure is encoded before directly fixating the polymorphemic word (for a review of this topic see Hyönä, 2011).

Identifying Compound Words During Reading

Compositional or Holistic Processing—Or Both?

The eye-tracking studies examining compound word processing have used an ingenious and simple paradigm developed by Taft and Forster (1976) to investigate the possible role of morphemes in word recognition. According to its logic, if morphemes play a significant role in (compound) word identification, manipulating the frequency of the morphemes constituting a compound word should exert a significant effect on identification, after controlling for the frequency of the compound word. An effect of morpheme frequency would indicate that the morpheme is used as an entry in accessing the word in the mental lexicon. On the other hand, if morpheme frequency does not influence word identification time but word frequency does, this is taken as evidence for the view against the involvement of morphemes in compound word identification. Instead, it would support the notion that the whole-word form is used as an entry to the mental lexicon.

In what follows, results are first summarized that are obtained from the manipulation of first- and second-constituent frequency, followed by whole-word frequency. (From here on, free morphemes that are components of compound words are called constituents.) In the studied languages (English, Finnish, and German), the second constituent of a two-constituent compound word is typically the *head* that defines the general meaning of the word (though not always; see more in what follows), while the first constituent modifies the head (e.g., *football* is a type of ball you kick with the foot).

Effects of First-Constituent Frequency

The first study employing the eye-tracking technique was conducted by Hyönä and Pollatsek (1998), who manipulated the frequency of the first constituent in two-constituent, rather long (12–14 letters) Finnish compound words while controlling for the second-constituent frequency and the whole-word frequency. (Compounding is highly productive in Finnish, as in German, but unlike English, there are never spaces between the constituents.) The studies examining the effects of first-constituent frequency are summarized in Table 9.1. Here the effect refers to the difference between the low-frequency and high-frequency conditions. Hyönä and Pollatsek obtained a highly reliable first-constituent frequency effect in gaze duration. The effect was also significant in the first-fixation duration on the word, which revealed an early effect (these long compound words were typically read with two fixations). However, the bulk of the effect appeared later (during the second and third fixations). Similarly, Pollatsek and Hyönä (2005) observed a reliable first-constituent frequency effect in gaze duration for long, semantically transparent and opaque compounds (however, the effect on first-fixation duration was not significant).

Bertram and Hyönä (2003) replicated the first-constituent frequency effect for long Finnish compound words but not for shorter ones (7–9 letters). For long compounds the effect was reliable for both first-fixation and gaze duration, whereas neither effect reached significance for short compounds. These results were taken to suggest a difference in reading long versus short compound words. The final Finnish study was that of Pollatsek, Bertram, and Hyönä (2011), who established a reliable first-constituent effect

Table 9.1 Effect of First-Constituent Frequency (Difference in Milliseconds Between the Low-Frequency and High-Frequency Conditions) in Gaze Duration (GD) and First-Fixation Duration (FFD) for Two-Constituent Compound Words.

Study	Language	Word Length (letters)	Difference in GD (ms)	Difference in FFD (ms)
Long compounds				
Hyönä & Pollatsek (1998)	Finnish	12+	87*	9*
Bertram & Hyönä (2003)	Finnish	12+	70*	16*
Pollatsek & Hyönä (2005), Exp. 1	Finnish	12+	47*	2
Juhasz (2008)	English	10.9	–25	–3
Pollatsek et al. (2011): Existing compounds	Finnish	12+	62†	20*
Pollatsek et al. (2011): Novel compounds	Finnish	12+	153*	29*
Overall			**66**	**12**
Shorter compounds				
Bertram & Hyönä (2003)	Finnish	7.5	11†	–3
Juhasz et al. (2003)	English	9	8	11†
Andrews et al. (2004)	English	8.5	21† 27†	8† 7†
Juhasz (2007)	English	9.1	36*	8†
Inhoff et al. (2008): Headed compounds	English	9.1	40*	20*
Inhoff et al. (2008): Tailed compounds	English	9.1	0	5
Juhasz (2008)	English	6.6	31†	12†
Overall			**22**	**9**

* = significant.
† = either $.05 < p_{1,2} < .1$, or p_1 or $p_2 > .1$.

for both long existing and long novel compound words, the effect being greater for novel compounds (for novelty effects, see the section titled "Meaning Computation in Reading Existing and Novel Polymorphemic Words").

Most of the studies in English have employed shorter compound words than the ones used in the studies in Finnish. The results have not been as clear as those reviewed earlier for Finnish. Juhasz, Starr, Inhoff, and Placke (2003) found a marginally significant first-constituent frequency effect in first-fixation duration for 9-letter English compounds, but the effect was not significant in gaze duration. Similarly, Andrews, Miller, and Rayner (2004) observed a marginally significant first-constituent frequency effect in first-fixation and gaze duration for two-constituent English compound words that were 6 to 11 letters long (average length 8.5 letters). Juhasz (2007) manipulated first-constituent frequency separately for semantically transparent and opaque English compounds, whose length ranged from 8 to 11 letters. (For effects of semantic transparency, see the section "Meaning Computation in Reading Existing and Novel Polymorphemic Words.") She found a significant effect of first-constituent frequency in first-fixation duration (however, it was nonsignificant in the item analysis) and gaze duration. Inhoff, Starr, Solomon, and Placke (2008) manipulated first-constituent frequency separately for *headed* and *tailed* compound words. Headed compounds are words where the overall meaning is more closely related to the first constituent (e.g., *humankind*), whereas in tailed compounds (the more frequent type), the overall meaning is more closely related to the second constituent (e.g., *handbook*). Inhoff et al. (2008) obtained reliable effects of the first constituent manipulation for headed compounds (both in first-fixation and gaze duration), but not for tailed compounds. Finally, Juhasz (2008) obtained results that were quite different from those of Bertram and Hyönä (2003). She did not find an early effect for long compound words (10–13 letters); instead the effect in first-fixation duration (and gaze duration) was nearly reliable for short compound words (6–7 letters). Moreover, first-constituent frequency did not produce an effect for long compound words even in gaze duration (in fact, the trend was in the opposite direction from that of prior studies).

To sum up the results for the effects of first-constituent frequency (see Table 9.1), the effect on gaze duration is robust for long Finnish compounds. The effect appears early in processing, as indexed first-fixation duration. These data indicate that long Finnish compound words are processed via the decomposition route. On the other hand, the only Finnish study examining the identification of short compounds provided evidence for the view that holistic processing is more involved in their reading than the decomposition route. The results are not as straightforward for English. In most studies in English the compound words have been shorter than in the Finnish studies. The general trend is that the first-constituent frequency affects processing, yet the effects are often statistically marginal. The most striking difference between the English and Finnish studies is that the only English experiment employing longer compounds (Juhasz, 2008) failed to find any evidence for compositional processing. It should be borne in mind, however, that compounding is less productive in English than in Finnish, which has consequences especially for creating representative samples of long English compound words. Before drawing further conclusions, let us next look at the results for the manipulation of second-constituent frequency.

Effects of Second-Constituent Frequency

Table 9.2 summarizes the data from studies examining second-constituent frequency effects. The first eye-tracking study was that of Pollatsek, Hyönä, and Bertram (2000), who found a reliable effect in gaze duration for long Finnish compounds; however, the effect was far from significant in first-fixation duration. Similar results were obtained in English. Juhasz et al. (2003) demonstrated a reliable second-constituent frequency effect in gaze duration for 9-letter-long English compounds; however, the effect was absent in first-fixation duration. Similar effects were reported by Andrews et al. (2004) and Juhasz (2007). Finally, Inhoff et al. (2008) found a significant effect in gaze duration

Table 9.2 Effect of Second-Constituent Frequency (Difference in Milliseconds Between the Low-Frequency and High-Frequency Conditions) in Gaze Duration (GD) and First-Fixation Duration (FFD) for Two-Constituent Compound Words.

Study	Language	Word Length (letters)	Difference in GD (ms)	Difference in FFD (ms)
Pollatsek et al. (2000)	Finnish	12+	95*	1
Juhasz et al. (2003)	English	9	27*	8
Andrews et al. (2004)	English	8.5	15†	4
Juhasz (2007)	English	9.1	16†	8
Inhoff et al. (2008): Headed compounds	English	9.1	5	5
Inhoff et al. (2008): Tailed compounds	English	9.1	46*	17†

* = significant.
† = either $.05 < p_{1,2} < .1$, or p_1 or $p_2 > .1$.

for tailed compounds but not for headed compounds. Moreover, the effect of second-constituent frequency for tailed compounds was marginal even for first-fixation duration.

To sum up the findings for the manipulation of second-constituent frequency, the evidence is clearly in favor of compositional processing, as most of the reported studies displayed a reliable effect in gaze duration. The effect is not as early as that of the first-constituent frequency, as indexed by the absence of an effect in first-fixation duration. An interesting exception is the effect in first-fixation duration obtained for tailed compound words. This suggests that meaning-defining constituents play a more significant role in identification than the nondominant constituents (i.e., those that do not define the word's core meaning) early during compound word processing.

EFFECTS OF WHOLE-WORD FREQUENCY

As argued earlier, effects of whole-word frequency can be used to examine the activation of the holistic route in compound word identification, as long as other variables, such as first- and second-constituent frequency, are controlled for. Table 9.3 summarizes the studies examining effects of whole-word frequency. Pollatsek et al. (2000) manipulated whole-word frequency for long, two-constituent Finnish compounds while controlling for the first-and second-constituent frequencies. There was a reliable effect in gaze duration, but the effect did not reach significance in first-fixation duration. Analogous results were obtained by Bertram and Hyönä (2003) for another set of long Finnish compound words. Juhasz (2008) found a sizable effect of whole-word frequency in gaze duration for long English compounds that also reached significance in first-fixation duration. Two other studies, one in Finnish (Bertram & Hyönä, 2003) and one in English (Juhasz, 2008) that examined processing of short compound words obtained similar results: a reliable effect both in first-fixation and gaze duration.

The preceding results lend consistent support for the view that the holistic route is in operation during compound word identification. Moreover, the Finnish studies suggest that its activation is somewhat delayed for long compounds but immediate for short compounds. However, the English study (Juhasz, 2008) does not suggest such a qualitative difference in identifying long versus short compound words. The theoretical implications of the studies reviewed so far will be discussed next.

THEORETICAL IMPLICATIONS OF THE REVIEWED COMPOUND WORD STUDIES

The evidence suggests that both the decomposition route and the holistic route are in operation

Table 9.3 Effect of Whole-Word Frequency (Difference in Milliseconds Between the Low-Frequency and High-Frequency Conditions) in Gaze Duration (GD) and First-Fixation Duration (FFD) for Two-Constituent Compound Words.

Study	Language	Word Length (letters)	Difference in GD (ms)	Difference in FFD (ms)
Long compounds				
Pollatsek et al. (2000)	Finnish	12+	82*	5†
Bertram & Hyönä (2003)	Finnish	12+	79*	4†
Juhasz (2008)	English	10.6	167*	20*
Overall			**109**	**10**
Short compounds				
Bertram & Hyönä (2003)	Finnish	7.5	52*	10*
Juhasz (2008)	English	6.6	69*	19*
Overall			**61**	**15**

* = significant.
† = either $.05 < p_{1,2} < .1$, or p_1 or $p_2 > .1$.

in compound word identification during reading. To account for the obtained pattern of results, Pollatsek et al. (2000) put forth a dual-route race model of compound word processing (see also Schreuder & Baayen, 1995). The model assumes that the morphological decomposition route and the holistic route operate in parallel during word identification, with the decomposition route typically running ahead of the holistic route—at least when reading long compound words. The study of Bertram and Hyönä (2003) suggests a clear difference in processing long versus short compounds, with the decomposition route dominating the initial processing of long compounds and the holistic route being dominant (i.e., winning the race) in reading short compound words. They posited that visual acuity is an important determinant of how compound words are processed and that the reader uses whatever visual information is available in foveal vision where the acuity is at its best. They coined this account the visual acuity principle. With long compound words, the first constituent is more readily available early in processing than the whole-word form, which gives a head start for the decomposition route. With shorter compound words, on the other hand, the whole-word form is within foveal reach; thus, the holistic route becomes active early during the identification process and is thus likely to win the race. As noted earlier, Juhasz (2008) provided evidence contradicting that reasoning. When discussing morphological segmentation in what follows, I will discuss additional evidence pertinent to this issue (for a further discussion of the issue, see Hyönä, 2012).

Inhoff et al. (2008) suggested that compound word processing is better described by an interactive use of the two routes (see also Taft, 1994) rather than the decomposition and holistic route being independent of each other. The reason for suggesting an interactive framework comes from their finding that an effect of second-constituent frequency was obtained for tailed compounds even during the first fixation. Recall that the second constituent of tailed compounds hosts the meaning-defining constituent. That is, the holistic route boosts the activation of the meaning-defining constituent more than the less meaning-dominant constituent; hence, an early effect of second-constituent frequency can be obtained. It is clear that more data are needed to assess to what extent the two routes interact with each other and to what extent they operate independently.

Morphological Segmentation During Compound Word Reading

As reviewed, reading compound words (at least long ones) involves morphological decomposition. Decomposition presupposes that the word can be segmented into its morphological components. Thus, factors that help segmentation should also facilitate compound word processing. Inhoff, Radach, and Heller (2000) conducted the first eye-tracking study to examine effects of signaling morphological boundaries in three-constituent compound words (e.g., *Datenschutzexpert*). A highly salient cue is to mark the boundaries by inserting a space between the constituents (e.g., *Daten schutz expert*). The downside of this is that it destroys the visual unity of the word; moreover, the form appears unfamiliar to German readers. Indeed, Inhoff et al. found both facilitation and inhibition in processing illegally spaced compound words. Facilitation was seen in shorter gaze durations on compounds where spaces demarcated the morphemes; inhibition was apparent in longer third and fourth fixations made on spaced than unspaced compounds. The pattern of data suggests that spacing facilitates morphological decomposition and component access but may interfere with meaning computation, as the (grammatical) relation of the constituents to each other is obscured by spacing. This is because it is unclear whether or not the lexemes separated by spaces belong to the same linguistic unit (noun phrase). Inhoff et al. also examined the effects of marking the morpheme boundaries (1) by an infrequent bigram spanning the morpheme boundary or (2) by capital letters (*DatenSchutzExpert*). Bigram frequency exerted no effects that could be discerned from the eye-movement records. Given the length of the target words (15–25 letters), it may not come as a surprise that single bigrams do not sufficiently stand out within such long letter strings. Capitalization had either no effect (Experiment 3) or smaller effects (Experiment 4) than spacing. Thus it appears that as a morphological boundary cue, capitalization is less effective than spacing.

Juhasz, Inhoff, and Rayner (2005) provided further evidence to support the view that spacing facilitates the decomposition of compound words. Their study was conducted in English. The target words were either unspaced (e.g., *bookcase*) or spaced (e.g., *rush hour*): two-constituent compound words appeared either in their "correct" form (as just shown) or an "incorrect" form (*book case* vs. *rush-hour*). Correctness is somewhat arbitrary in English, as there is no grammatical rule for deciding whether a compound is spaced or not. Juhasz et al. found that the first fixation on a compound was shorter if the compound was shown with a space, irrespective of whether the compound was supposed to be spaced—a finding consistent with the notion that providing a space between the constituents aids decomposition. On the other hand, inserting a space for normally unspaced compounds lengthened the gaze duration, whereas deleting the space for normally spaced compounds did not affect gaze duration. The gaze duration data indicate that later processing is facilitated by removal of the space, and may be taken to support the view that the holistic route benefits from concatenation.

Another visually salient cue is the hyphen (about the use of hyphens in English compound words, see Kuperman & Bertram, 2013) that is often inserted at morpheme boundaries (e.g., *word-play*). Bertram and Hyönä (2013) examined how hyphens inserted at constituent boundaries of long and short Finnish bimorphemic compound words influence the identification process. The Finnish spelling convention dictates a hyphen at the constituent boundary when the same vowel appears as the last letter of the first constituent and the first letter of the second constituent (e.g., *ulko-ovi* 'front door'; *musiikki-ilta* 'musical evening'). Reading of hyphenated compounds was compared with that of nonhyphenated compounds matched on frequency and length of the word and of the first and second constituent. Note that all words appeared in their legal format. Based on the study of Bertram and Hyönä (2003), Bertram and Hyönä (2013) predicted that a hyphen should facilitate processing of long compounds and inhibit processing of short compounds. That was exactly what they observed. Gaze durations were 64 ms shorter for long hyphenated compounds than for long concatenated compounds, whereas gaze durations for short hyphenated compounds were 43 ms longer than for short concatenated compounds (see also Häikiö, Bertram, & Hyönä, 2011). That is, a hyphen facilitates the processing of long compounds because the decomposition route gets a head start over the holistic route due to the initial constituent being more readily available in foveal vision than the whole-word form. On the other hand, a hyphen encourages sequential processing of short compounds when holistic processing is a viable option (i.e., the whole word is within foveal reach)—hence inhibition is observed in processing short compound words.

Similarly to Inhoff et al. (2000), Bertram, Kuperman, Baayen, and Hyönä (2011) examined the processing of three-constituent compounds. However, instead of using spaces to demarcate the morphemic boundaries, they illegally inserted hyphens at the boundaries in three-constituent Finnish (*lentokenttä-taksi* 'airport taxi') and Dutch (*voetbal-bond* 'football association') compounds. As mentioned earlier, hyphens are legally used in Finnish compound words when one constituent ends with the same vowel with which the following constituent begins. In Dutch, a hyphen may be inserted in a multiconstituent compound word if the writer thinks the word is difficult to parse otherwise. In the study of Bertram et al. (2011), illegal insertion of hyphens was done separately for so-called *left-branching* and *right-branching* compound words. Right-branching compounds are those where the second and third constituent constitute the compound head and the first constituent functions as its modifier, as in *zaalvoetbal* ('indoor football'), whereas left-branching compounds are those where the third constituent is the head modified by the first two constituents, as in *voetbalbond* ('football association'). Bertram et al. (2011) found an early processing benefit when the hyphen was inserted at the boundary between the modifier and the head (*voetbal-bond; zaal-voetbal*). Fixation time on the first constituent(s) prior to the hyphen was shorter than fixation time on the corresponding region without the presence of a hyphen. This is taken to suggest that a hyphen helps readers to assign a correct hierarchical structure to a three-constituent compound word. This early facilitation was offset by a later processing cost (not present for Finnish left-branching compounds) for reading the word region on the right side of the hyphen. The latter effect is likely to reflect readers' response to illegal spelling (i.e., the presence of the hyphen).

The pattern of results (early facilitation combined with later inhibition) obtained by Bertram et al. (2011) in Dutch is consistent with what Inhoff et al. (2000) observed in German for illegally inserting a space between constituents in three-constituent compound words. In word processing in Dutch and Finnish, as indexed by gaze duration, illegally inserting a hyphen at the minor boundary (i.e., not at the modifier-head boundary) resulted in slower word identification compared with the correct concatenated word form. The two languages differed in that the presence of a hyphen at the major boundary (i.e., at the modifier-head boundary) resulted in shorter gaze durations in Finnish but in no difference in Dutch, compared with the concatenated form. This language difference is assumed to reflect Finnish readers' greater exposure to hyphenated compound words in their native language. (Finnish contains many more compound words in general than Dutch and many more hyphenated compound words.)

Effects of more subtle segmentation cues were studied by Bertram, Pollatsek, and Hyönä (2004). They asked whether the type of letter clusters spanning the morpheme boundary is used in morphological segmentation. They made use of a specific feature of Finnish, so-called vowel harmony: back vowels (*a, o, u*; pronounced toward the back of the mouth) and front vowels (*ä, ö, y*; pronounced toward the front of the mouth) cannot appear in the same lexeme (i.e., within the same compound word constituent), whereas vowels of different quality can appear across constituents. Thus when two vowels of different quality appear next to each other as in *öljy**o**nnettomus* ('oil spill'), it is a logical cue for a constituent boundary: The front vowel *y* must be the end of the first constituent, and the back vowel *o* must be the beginning of the second constituent. Bertram et al. (2004) demonstrated that such vowel disharmony may indeed be used as a segmentation cue. Gaze durations were shorter when the two vowels spanning the morpheme boundary were disharmonious rather than harmonious. A reliable effect of vowel harmony was also obtained even when the (dis)harmonious vowels were not adjacent to each other (*kylpyl**ä**h**o**telli* 'spa hotel'). However, the vowel harmony effect was only apparent when the first constituent was long (7–9 letters) but not when the first constituent was short (3–5 letters); here the length of the whole word was controlled for. This pattern of results was taken to suggest that when the identification of the first constituent is fast, as is presumably the case with short initial constituents, and the initial fixation on the word is close to the boundary, morphological segmentation cues are not needed. On the other hand, such cues come in handy when the initial constituent is not readily parsed (i.e., it is longer and the initial fixation is positioned further away from it), as is supposedly the case with long first constituents.

Finally, Bertram et al. (2004) examined another more subtle segmentation cue, the probability of a consonant appearing as the initial or final letter of a lexeme in Finnish. Effects of consonant quality were investigated in reading long, bimorphemic compound words. They contrasted two conditions: one where the initial consonant of the second constituent cannot appear as the final letter in a lexeme, and another where the initial consonant of the second constituent can appear either as an initial or final letter in a lexeme. The former condition unambiguously signals a morphemic boundary, whereas the latter condition does not. Such cuing resulted in a reliable effect observed in gaze duration: the unambiguous consonant cue produced shorter gaze durations than the ambiguous consonant cue. The size of the effect was similar to that of vowel harmony. Moreover, effects of consonant type and vowel harmony were independent of each other. The effects of cuing morphological boundaries obtained by Bertram et al. (2004) were relatively late effects; they peaked at the third fixation made on the word (the effects were also significant in the second and fourth fixation). The late appearance of the effect is inconsistent with morphological processing models that assume morphological decomposition to occur early in the processing timeline, prior to lexical access (Rastle, Davis, & New, 2004; Taft, 1979, 1994). There are several possible reasons for the observed differences in the timing of the effects. One possibility is that the early prelexical effects are short-lived and can only be recovered using the masking paradigm (e.g., Rastle et al., 2004). It is also possible that sentence context obscures prelexical effects that would be present when isolated words are presented for identification.

In summary, there is consistent evidence demonstrating facilitation in reading long compound words due to the presence of morphological boundary cues. These cues are not only visually salient ones, such as spaces and hyphens, but also more subtle cues related to letter quality. However, in response to an illegal insertion of a space or hyphen at the morphological boundary, a later slowing in processing may be observed. Finally, legal hyphens at morphemic boundaries of short compound words result in a processing detriment. The summarized pattern of results for long and short compound words can be readily accounted for by the visual acuity principle (Bertram & Hyönä, 2003). According to this account, due to visual acuity constraints the decomposition route gets a head start in processing long compound words, presumably leading to facilitation in processing in the presence of cues aiding decomposition. On the other hand, as short compound words are within foveal reach, the holistic route is more likely to be the prevailing access route. Thus, morphological boundary cues encouraging decomposition will lead to slowing of processing for shorter compound words.

Identifying Derived Words in Reading

Eye movement studies investigating reading of derived words are scarcer than those on compound word reading. The first eye-tracking study of normal reading is that of Niswander, Pollatsek, and Rayner (2000; see also Holmes & O'Regan, 1992). They were interested in finding out whether suffixed words (e.g., *adoption*) are identified via the decomposition route or via the whole-word route. In order to do so, they employed the logic put forth by Taft and Forster (1976; see the preceding discussion for more details) by independently manipulating root (*adopt*) frequency and word (*adoption*) frequency for a set of 7- to 12-letter suffixed words in English (Experiment 1). They found an early effect of root frequency, as indexed by first-fixation duration, and a bit later effect of word frequency, as indexed by gaze duration. Both effects spilled over to the processing of the word following the target word. The spillover effect was assessed by the duration of the first fixation on the word to the right of the target word. The only puzzling result was that the root-frequency effect was not reliably observed in gaze duration on the target word. The pattern of data was interpreted to suggest that both the decomposition route, indexed by the root-frequency effect, and the holistic route, indexed by the word frequency effect, are in operation in reading derived suffixed words, with the decomposition route being active a bit earlier than the holistic route.

Niswander-Klement and Pollatsek (2006) studied reading English prefixed words such as *remove*. Similarly to Niswander et al. (2000), they manipulated the frequency of the word root (*move*) and that of the whole word (*remove*); in addition, they varied word length. Niswander-Klement and Pollatsek observed a significant root-frequency effect in gaze duration for longer but not for shorter, prefixed words. On the other hand, the word frequency effect in gaze duration was reliable for shorter, but not for longer, prefixed words. The pattern of results is similar to that observed by Bertram and Hyönä (2003) for short and long compound words. The results can be readily interpreted by the visual acuity principle. Due to visual acuity constraints, shorter prefixed words are more likely to be identified holistically, whereas compositional processing appears more important in reading longer prefixed words.

Pollatsek, Slattery, and Juhasz (2008) investigated the reading of relatively long novel (e.g., *overmelt*) or existing (e.g., *overload*) prefixed English words. Novel prefixed words were constructed such that their meaning could be readily derived from the component morphemes. By their nature, novel prefixed words have to be identified compositionally. Thus by comparing the processing of existing prefixed words to that of novel prefixed words, one can estimate the degree to which the decomposition and the holistic routes operate in reading existing prefixed words. In Experiment 1, reading of novel and existing prefixed words was studied by matching them on length (mean length of 9.6 letters). A sizable novelty effect (104 ms in gaze duration) was observed, which suggests that reading of existing prefixed words heavily relied on the holistic route. In Experiment 2, Pollatsek et al. (2008) varied root frequency factorially for novel and existing prefixed words to estimate the relative strength of the decomposition route in reading these two types of words. Root frequency exerted a reliable effect both in first-fixation duration (an effect size of 12 ms) and gaze duration (an effect size of 54 ms). Surprisingly, however, the root-frequency effect was no bigger for novel than for existing prefixed words. If anything, it was slightly bigger in gaze duration for existing prefixed words. The results were interpreted to support a dual-route race model where the two routes interact with each other. A dual-route race model with independent routes would predict a substantially larger root-frequency effect for novel than for existing prefixed words. This is because for novel prefixed words there is no competing holistic route that could win and thus terminate the race.

In sum, the current evidence on reading derived words consistently demonstrates that both the decomposition route and the holistic route are in operation in identifying them. As with compound words, it seems that the holistic route dominates the identification of short derived words, while the decomposition route is active early on in the processing of longer derived words.

Identifying Inflected Words in Reading

Reading inflected words in sentence contexts has been studied to a lesser extent than other morphologically complex words. Niswander et al. (2000; Experiment 2) examined reading of pluralized nouns (e.g., *uncles, beaches*), verbs with *-ing* endings (e.g., *watching*), and regular past-tense verbs (e.g., *directed*) in English (5–9 letters in length). Root frequency and word frequency were manipulated. There were reliable word frequency effects both in first-fixation and gaze duration; however, no significant root-frequency effects were obtained. In follow-up analyses of these three types of inflected words, a significant root-frequency effect emerged in first-fixation and gaze duration for plural nouns but not for verbs. Niswander et al. pointed out that many verb roots used in the study were also nouns (e.g., *handed*), which may explain the absence of a root-frequency effect when used as verbs. For many of the inflected verbs, another part of speech (typically noun) was the more common use of the root. This kind of root-word conflict was not apparent for the target nouns. Follow-up regression analyses provided suggestive evidence that the root-word conflict apparent for verbs may modulate the total fixation time, indexing a late effect. (Total fixation time sums the durations of fixations made during the first-pass reading and regressive fixations made on the target word after first-pass reading.)

Bertram, Hyönä, and Laine (2000) examined the reading of inflected nouns in Finnish by manipulating word and root frequency for words containing the *-ja* inflection denoting the partitive plural case (e.g., *hattuja* 'some hats'). Partitive is the third most common case ending of the 13 inflectional cases of Finnish. Average word length was about 7 letters. Bertram et al. obtained a 17-ms word-frequency effect in gaze duration on the target word as well as a 21-ms effect in gaze duration on the word following the target. However, no root-frequency effect was observed on the target word either in first-fixation duration or gaze duration. There was an 8-ms delayed effect (observed in gaze duration on the following word) of root frequency that was only significant in the participant analysis. However, the delayed effect of root frequency was fully significant (in both the participant and item analysis) in a word-by-word self-paced reading experiment. The pattern of results suggests that the holistic route is active in identifying inflected words, while the decomposition route becomes active after lexical access is achieved. One possibility is that inflections are processed by a syntactic module (Taft, 1994), which could explain why the word's decomposed form becomes active after lexical access. The tentative suggestion of Taft (1994) was proposed to account for the identification of inflected verbs. The idea is that the inflection is decomposed from its stem and fed into a syntactic module for processing.

Further evidence in support of the postlexical (i.e., effects appearing only after fixating away from the critical word to a subsequent word) view is reported by Hyönä, Vainio, and Laine (2002), who compared reading of morphologically complex, inflected words with that of morphologically simple, noninflected words. The study was conducted in Finnish. The inflected words appeared either in the partitive singular case, denoted by the inflection *-a* or *-ä* (e.g., *juustoa* '(some) cheese'), or in the genitive case denoted by the inflection *-n* (e.g., *peiton* 'blanket's'), while the noninflected words appeared in the nominative case; the three employed cases are among the most frequent ones in Finnish. The compared word forms were syntactically equally plausible in the sentence frames they appeared in. They were all clause objects, which can take any one of these three cases. No difference in gaze duration was found between the inflected and noninflected word forms. The only tendency for an effect (nonsignificant in the item analysis) was obtained in the second-pass fixation time (i.e., the duration of fixations returning to the target word after first-pass reading) of the target word, which was 19 ms longer for the inflected words.

The null effect of morphological complexity in first-pass reading may be interpreted to suggest that the word forms studied by Hyönä et al. (2002) are all identified via the whole-word route. Moreover, the trend observed in the second-pass fixation time may be taken to suggest that these kinds of contextual inflections (Booij, 1996) are processed by a syntactic module (Taft, 1994) after lexical access is achieved. This account, not necessarily the view Taft (1994) envisioned, would assume that the lagged effect of morphological complexity obtained for inflected words (assuming that the effect is real) reflects postaccess syntactic checking or integration. However, the fact that the three word forms studied by Hyönä et al. (2002) were syntactically equally plausible in the sentence context is challenging to the postlexical syntactic account that predicts no processing difference between them. More empirical evidence is definitely needed before any firmer conclusions can be drawn.

To sum up, the number of studies investigating reading inflected words is still rather limited; thus, the conclusions drawn are bound to be very tentative. Yet the available evidence paints a relatively consistent picture of reading inflected nouns. The reviewed studies suggest that the holistic route dominates the identification of inflected nouns and that the decomposition route becomes active after lexical access, possibly indicating that inflections are processed by a syntactic module. However, the only reading study examining the processing of verb inflections found no evidence for morphological decomposition. Thus more empirical evidence is definitely needed before it is possible to say anything more conclusive about reading inflected words. For example, it may be of interest to examine whether the distinction between contextual (e.g., structural case markers) and inherent (e.g., category number) inflections (Booij, 1996) would have processing consequences. By definition, contextual inflections are dictated by syntax, while inherent inflections are not required by the syntactic context. Thus the linguistic context where an inflected word appears may impose different constraints on reading words containing contextual versus inherent inflections. Finally, the studied word forms have been rather short; it is important to examine to what extent the obtained effects generalize to longer inflected words. It is possible that short inflected words can be recognized holistically, as they fit in the foveal area, whereas longer inflected words may encourage morphological parsing for their identification, as they cannot be identified with a single eye fixation.

Meaning Computation in Reading Polymorphemic Words

In the last section I will review studies examining the computation of meaning of polymorphemic words whose meaning is either ambiguous or nontransparent given the meaning of individual morphemes, or whose meaning has to be computed from the component meanings, as is the case with novel polymorphemic words. Effects of ambiguity in meaning (i.e., the studied words allow two alternative parses) have been studied with derived words, while meaning computation has been investigated using both semantically opaque compound words and novel derived and compound words as stimuli.

Ambiguity in Meaning

Pollatsek, Drieghe, Stockall, and de Almeida (2010) studied the reading of trimorphemic derived words (e.g., *unlockable*) that contained a prefix (*un*), a root (*lock*), and a suffix (*able*). The studied words were ambiguous in meaning, as their morphemic structure allowed two alternative parses. The word *unlockable* can be parsed as *un-lockable* to mean something that cannot be locked or as *unlock-able* to refer to something that can be unlocked. In the former case the structure is referred to as right-branching, in the latter case as left-branching. Pollatsek et al. (2010) manipulated the preceding sentence either to be neutral, to constrain the left-branching structure, or to constrain the right-branching structure, in order to find out what the preferred interpretation is for such ambiguous multimorphemic words. An example of a context biasing the left-branching structure for *unlockable* is as follows: *The zookeeper needed to get the new bird out of its cage. As the cage was unlockable, his key quickly opened it and he removed the bird*. An example of a context biasing the right-branching structure is as follows: *The zookeeper wanted to make sure the bird stayed in its cage. As the cage was unlockable, he needed to fasten the door with wire instead*. The neutral context sentence read as follows: *The zookeeper inspected the new bird and its cage*.

Pollatsek et al. (2010) observed no effects on the word itself. However, the so-called go-past fixation time reflecting further processing (i.e., initiated by a regression from the target word to a preceding sentence and terminated by a fixation going beyond the target word) suggested that the left-branching interpretation (i.e., *unlock-able*) is the favored one. More precisely, the right biasing context sentence facilitated later processing of right-branching structures relative to the neutral context sentence, but there was no difference for the left-branching structures between the neutral context sentence and left-branching biasing context sentence, which suggests that the left-branching alternative is considered to be the default structure.

Meaning Computation in Reading Existing and Novel Polymorphemic Words

Pollatsek et al. (2008) compared the reading of existing and novel prefixed words (e.g., *unmarried* vs. *unblamed*) and found a robust novelty effect (104 ms) in gaze duration, despite the fact that the meaning of both lexicalized and novel prefixed words can be easily computed from the component morphemes. This indicates that accessing the meaning of a polymorphemic word takes time when the word form cannot be accessed via the direct route, which uses an existing whole-word representation in the mental lexicon. Yet as reviewed earlier, their data showed that the decomposition route did not speed up processing of the novel prefixed words any more than the lexicalized prefixed words, which points to the view that the routes operate interactively.

Pollatsek, Bertram, and Hyönä (2011) came to a somewhat different conclusion in their study of novel and lexicalized Finnish two-constituent compound words (average length about 13 letters). Apart from novelty, they manipulated the frequency of the first constituent as a separate lexeme. They observed a robust novelty effect of 187 ms in gaze duration. It began to manifest itself during the second fixation, as the readers' need to access the second constituent for the novelty became apparent. Moreover, the two manipulated factors interacted with each other: The first-constituent frequency effect was greater for novel than lexicalized compound words. This result suggests that the two routes operate independently of each other—a conclusion divergent from that proposed by Pollatsek et al. (2008) for reading novel and lexicalized prefixed words. The different conclusions are reconciled by Pollatsek et al. (2011) by assuming that the identification process entails two levels of processing (see also Libben, 1998): identification of the orthographic and phonological forms and meaning computation. The identification stage is assumed to work similarly for compound and prefixed words. The early constituent- and root-frequency effects (at least for long words) presumably reflect this stage. However, the operation of the second stage differs between the two word types. For prefixed words, meaning computation is rule-governed (misX, e.g., *miscircle*, means that someone did X wrongly in some way), while meaning computation is not as straightforward for compound words due to the multiple ways the constituents may relate to each other (Gagné & Spalding, 2009). Being rule-governed, the second processing stage is not influenced by the first stage in reading prefixed words. On the other hand, with compound words the first stage, identifying the first constituent, impacts on the second stage. This is because the prototypical relationships the first constituent is involved in are less firmly established for infrequent than frequent lexemes.

In the case of lexicalized compound words, when the compound word meaning is opaque in the sense that it cannot readily be derived from the constituent meanings (e.g., *jailbird*), meaning computation appears to be absent (i.e., meaning is directly retrieved). Evidence consistent with this view was provided by Pollatsek and Hyönä (2005; Experiment 1), who manipulated first-constituent frequency for both semantically opaque and transparent two-constituent compound words. They observed a reliable first-constituent frequency effect in gaze duration for long (an average of about 13 letters) Finnish two-constituent compound words, which was similar in size for semantically opaque and transparent compounds. However, there was no overall transparency effect (a difference of 1 ms in gaze duration), which suggests that the meaning of existing opaque compound words is retrieved rather than computed from the component meanings. If the latter had been the case, a slowdown in processing should have been observed for semantically opaque in comparison to semantically transparent compounds.

Frisson, Niswander-Klement, and Pollatsek (2008) compared reading of semantically transparent compounds with that of partially opaque (either the first or second constituent is opaquely related to the meaning of the whole word, e.g., *trenchcoat*, *heirloom*) and fully opaque (both constituents are opaquely related to the word meaning, e.g., *cocktail*). The study was conducted in English using two-constituent compounds with an average length of about 8.8 letters. Similarly to Pollatsek and Hyönä (2005), they found no effect of semantic transparency in gaze duration (Experiment 1). Interestingly, when compositional processing was strengthened (see Experiment 2) by inserting a space between the constituents (e.g., *cock tail*), an effect of opacity was observed as a delayed effect (in gaze duration of the word subsequent to the target). This is likely to reflect meaning computation, which was presumably absent (i.e., meaning is retrieved for the compound word without separately accessing the meaning of its components and computing the meaning from them) when the words were presented legally, in concatenated form.

Finally, Juhasz (2007) conducted a study similar to those of Pollatsek and Hyönä (2005) and Frisson et al. (2008). She orthogonally manipulated first- and second-constituent frequency for semantically opaque and transparent English compound words. Unlike Pollatsek and Hyönä and Frisson et al., Juhasz obtained a reliable 24-ms transparency effect in gaze duration. On the other hand, similarly to Pollatsek and Hyönä, transparency did not interact with first-constituent frequency, nor did it interact with second-constituent frequency. Thus semantic transparency did not mediate the decomposition process, but morphological decomposition took place similarly with transparent and opaque compounds.

In sum, on the basis of the available evidence, the following conclusions may be derived concerning meaning computation for polymorphemic words during reading: The process is initiated with a lexical access of the morphemes (at least when the words are relatively long). This lexical phase operates similarly regardless of semantic transparency or novelty. The second stage involves meaning computation, which is not needed for lexicalized words, even when their meaning is opaque given the meaning of individual morphemes (but see Juhasz, 2007). The two lexical routes, the decomposition and holistic route, work in tandem to provide access to the word represented in the mental lexicon. The lexical access in turn leads to the retrieval of the word's meaning without the need to compute the compound word's meaning from the meanings of its components. With novel word forms, however, the second stage either entails an application of a rule to derive the word's meaning, as is the case with novel prefixed words, or the meaning has to be computed by figuring out the relationship between constituent meanings, as is the case with novel compound words.

Concluding Remarks

In this chapter I have reviewed eye-movement studies of reading polymorphemic words in sentence context. To date, the research has focused on the identification of compound words; relatively few studies have been conducted on reading derived and inflected words. Moreover, the existing evidence for alphabetic scripts is based on a small set of languages (Finnish, English, German, and Dutch). Obviously, future research should expand this limited set of languages to include a wider variety of typologically different languages.

The currently available evidence relatively consistently converges on three general conclusions: (1) Polymorphemic words are identified during reading using both whole-word representations available in the mental lexicon as well as accessing the word identity via the component meanings. (2) Word length plays a significant role in modulating the relative dominance of the two access routes, with the decomposition route dominating the early stages of the identification process of long polymorphemic words (compound and derived words; studies on long inflected words are still missing), while the whole-word route is more dominant in identifying short polymorphemic words. (3) Based on a limited amount of evidence, it is tentatively suggested that inflected word forms are identified via the holistic route with the inflection being processed postlexically by a syntactic module. More research on reading inflected words in sentence context is definitely needed before firmer conclusions can be made on their processing.

References

Amenta, S., & Crepaldi, D. (2012). Morphological processing as we know it: An analytical review of morphological effects in visual word recognition. *Frontiers in Psychology, 3,* 232.

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in the mousetrap. *European Journal of Cognitive Psychology, 16,* 285–311.

Bertram, R. (2011). Eye movements and morphological processing in reading. *The Mental Lexicon, 6,* 83–109.

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language, 48,* 615–634.

Bertram, R., & Hyönä, J. (2013). The role of hyphenation in compound word identification: Facilitative for long, detrimental for short words. *Experimental Psychology, 60,* 157–163.

Bertram, R., Hyönä, J., & Laine, M. (2000). The role of context in morphological processing: Evidence from Finnish. *Language and Cognitive Processes, 15,* 367–388.

Bertram, R., Kuperman, V., Baayen, R. H., & Hyönä, J. (2011). The hyphen as a segmentation cue in triconstituent compound processing: It's getting better all the time. *Scandinavian Journal of Psychology, 52,* 530–544.

Bertram, R., Pollatsek, A., & Hyönä, J. (2004). Morphological parsing and the use of segmentation cues in reading Finnish compounds. *Journal of Memory and Language, 51,* 325–345.

Booij, G. (1996). Inherent versus contextual inflection and the split morphology hypothesis. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1995* (pp. 1–16). Dordrecht, the Netherlands: Kluwer.

Frisson, S., Niswander-Klement, E., & Pollatsek, A. (2008). The role of semantic transparency in processing of English compound words. *British Journal of Psychology, 99,* 87–107.

Gagné, C. L., & Spalding, T. L. (2009). Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language, 60,* 20–35.

Giraudo, H., & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes, 15,* 421–444.

Häikiö, T., Bertram, R., & Hyönä, J. (2011). The development of whole-word representations in compound word processing: Evidence from eye fixation patters of elementary school children. *Applied Psycholinguistics, 32,* 533–551.

Holmes, V. M., & O'Regan, J. K. (1992). Reading derivationally affixed French words. *Language and Cognitive Processes, 7,* 163–192.

Hyönä, J. (2011). Foveal and parafoveal processing during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 819–838). Oxford, England: Oxford University Press.

Hyönä, J. (2012). The role of visual acuity and segmentation cues in compound word identification. *Frontiers in Psychology, 3,* 188.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1612–1627.

Hyönä, J., Vainio, S., & Laine, M. (2002). A morphological effect obtains for isolated words but not for words in sentence context. *European Journal of Cognitive Psychology, 14,* 417–433.

Inhoff, A. W., Radach, R. & Heller, D. (2000). Complex compounds in German: Interword spaces facilitate segmentation but hinder assignment of meaning. *Journal of Memory and Language, 42,* 23–50.

Inhoff, A. W., Starr, M. S., Solomon, M., & Placke, L. (2008). Eye movements during reading of compound words and the influence of lexeme meanings. *Memory & Cognition, 36,* 675–687.

Juhasz, B. J. (2007). The influence of semantic transparency on eye movements during English compound word recognition. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 373–389). Oxford, England: Elsevier.

Juhasz, B. J. (2008). The processing compound words in English: Effects of word length on eye movements during reading. *Language and Cognitive Processes, 23,* 1057–1088.

Juhasz, B. J., Inhoff, A. W. & Rayner, K. (2005). The role of interword spaces in the processing of English compound words. *Language and Cognitive Processes, 20,* 291–316.

Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decision, and eye fixations. *British Journal of Psychology, 94,* 223–244.

Kuperman, V., & Bertram, R. (2013). Moving spaces: Spelling alternation in English noun-noun compounds. *Language and Cognitive Processes, 28,* 939–966.

Libben, G. (1998). Semantic transparency in the processing of compounds: Consequences for representation, processing, and impairment. *Brain and Language, 61,* 30–44.

Niswander, E., Pollatsek, A., & Rayner, K. (2000). The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes, 15,* 389–420.

Niswander-Klement, E., & Pollatsek, A. (2006). The effects of root frequency, word frequency, and length on the processing of prefixed English words during reading. *Memory & Cognition, 34,* 685–702.

Pollatsek, A., Bertram, R., & Hyönä, J. (2011). Processing novel and lexicalized Finnish compound words. *Journal of Cognitive Psychology, 23,* 795–810.

Pollatsek, A., Drieghe, D., Stockall, L., & de Almeida, R. G. (2010). The interpretation of ambiguous trimorphemic words in sentence context. *Psychonomic Bulletin & Review, 17,* 88–94.

Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes, 20,* 261–290.

Pollatsek, A., & Hyönä, J. (2006). Processing of morphologically complex words in context: What can be learned from eye movements. In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 275–298). Hove, England: Psychology Press.

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 820–833.

Pollatsek, A., Slattery, T. J., & Juhasz, B. J. (2008). The processing of novel and lexicalised prefixed words in reading. *Language and Cognitive Processes, 23,* 1133–1158.

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review, 11,* 1090–1098.

Rayner, K., Well, A. D, Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception & Psychophysics, 31,* 537–550.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–154). Hillsdale, NJ: Erlbaum.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition, 7,* 263–272.

Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes, 9,* 271–294.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14,* 638–647.

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior, 15,* 607–620.

CHAPTER

10 Individual Differences Among Skilled Readers: The Role of Lexical Quality

Sally Andrews

Abstract

Theories of visual word recognition and reading have been based on averaged data from relatively small samples of skilled readers, reflecting an implicit assumption that all skilled readers read in the same way. This chapter reviews evidence of systematic individual differences in the early stages of lexical retrieval among samples of above-average readers that challenges this assumption. Individual differences in patterns of masked priming and parafoveal processing during sentence reading provide evidence of variability in the precision and coherence of lexical knowledge that are consistent with Perfetti's (2007) construct of lexical quality. This evidence is compatible with neuroimaging evidence that literacy drives the development of specialized neural systems for processing written words. Understanding these dimensions of individual differences among expert readers has important implications for future refinements of theories of visual word recognition and reading.

Key Words: orthographic processing, lexical quality, individual differences, masked priming, reading skill

Even a casual glance at the chapters in this volume reveals that we know a lot about how the average skilled reader reads words. Hundreds of papers reporting average data for samples of 20 to 30 university students in both single word and text reading tasks have confirmed a range of robust phenomena that generalize across samples, tasks, and languages. These findings provide the foundation for a rich landscape of theories of lexical retrieval and reading. Many of these models have been computationally implemented and systematically assessed against empirical benchmarks. Despite this degree of specificity and systematic evaluation, the validity of the different theoretical accounts remains a focus of intense debate. The central argument of this chapter is that investigations of individual differences among skilled readers can—and should—play an important role in future refinement of these competing models.

Most theories of skilled word recognition and reading have been validated against average data for samples of 20 to 30 university students. I am not questioning the value of this experimental research strategy, which underpins much of the empirical evidence and theoretical frameworks discussed in other chapters on skilled reading. Too early a focus on individual differences may have impeded extraction of these general principles of skilled reading behavior (Andrews, 2012). However, continuing to rely on averaged data to test the validity of different theories may lead the field astray.

Many computational models of visual word recognition can accurately simulate the average effects of major variables such as word frequency, length, and regularity on people's performance. But more fine-grained evaluation, such as comparisons of performance for different items, reveals that the models are very poor at predicting more detailed aspects of human performance. For example, the highly influential dual-route cascaded (DRC) (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) and parallel distributed processing (PDP) (Plaut, McClelland, Seidenberg, & Patterson, 1996) models account for less than 10% of variance in the average human naming latencies from a number of large-scale databases (Balota & Spieler, 1998). Computational modelers' solution to this problem has been to develop more refined models, often by combining the features of earlier models, to achieve a better fit with the human data (e.g., Perry, Ziegler, & Zorzi, 2007). However, the problem may not lie in the models but in the validity of the data used to evaluate them.

Relying on averaged data to evaluate theories reflects an implicit *uniformity assumption* (Andrews, 2012): that all skilled readers have developed the same cognitive architecture and read in the same way. The evidence reviewed in later sections shows that averaged data can obscure systematic individual differences and potentially lead to misleading conclusions about underlying cognitive processes. Rather than modifying models to better fit average patterns of performance, it is time for experimental psycholinguists to consider whether and how individual differences modulate skilled lexical retrieval.

The more specific goal of this chapter is to review and evaluate the contribution of *lexical quality* to individual differences among skilled readers. This construct was first introduced by Perfetti (1985) in the context of his *verbal efficiency theory* and subsequently defined as the "extent to which a mental representation of a word specifies its form and meaning components in a way that is both precise and flexible" (Perfetti, 2007, p. 359). High-quality representations are said to be orthographically fully specified, phonologically redundant, and semantically more generalized and less context-bound. The orthographic, phonological, and semantic constituents of a word are also assumed to be more tightly bound. These strong connections lead to more reliable, synchronous activation of the various constituent codes that define a word's identity, which in turn supports flexible use of this information to achieve the reader's goals (Perfetti, 2007, p. 360).

What critically distinguishes the lexical quality hypothesis (LQH) from many other accounts of individual differences in reading is that it assigns a causal role to lexical *knowledge*. Rather than attributing reading difficulties to dysfunctions in phonological, semantic, or working memory *mechanisms*, "the LQH is about knowledge that has not been acquired or practiced to a high-enough level" (Perfetti, 2007, p. 380). Ineffective or inefficient processes are a consequence of low quality knowledge representations.

The utility of the construct of lexical quality depends on how clearly it is defined. Perfetti's (2007) definition emphasizes the *precision* and *flexibility* of lexical knowledge. Precision is identified with the content of lexical knowledge: the specificity and completeness of the orthographic, phonological, grammatical, and semantic constituents of the lexical representation. As elaborated in what follows, orthographic precision is particularly critical because written forms are the gateway to lexical knowledge (Dehaene, 2009). Flexibility arises from the interconnectedness or binding between the different constituents of lexical representations. Strong connections allow printed word forms to trigger synchronous, coherent activation of all components of the word's identity required for comprehension (Perfetti, 2007).

According to Perfetti's (2007) definition, lexical quality is both word-specific and graded. The quality of lexical representations varies within individuals, as they gradually increase the size of their vocabulary and refine the specificity of the knowledge stored for existing words through reading experience. There are also differences between individuals in the extent to which they have established high-quality representations for most of the words in their written vocabulary. Such differences could arise from genetic differences in the capacity for orthographic learning (Byrne et al., 2008) or environmental factors, such as methods of instruction, as well as from differences in reading strategy. Readers who rely heavily on context to identify words may devote little attention to the details of words' internal structure and therefore be less likely to develop fully specified orthographic codes for all words (Frith, 1986).

Most investigations of lexical quality have focused on children, reflecting the common assumption that, although individual differences in word recognition are a major source of variance during early reading development, they make little contribution to variability in successful reading comprehension at later grades (e.g., Eason & Cutting, 2009). However, the university students who provide the averaged data in published studies vary considerably in their level and profile of performance in tests of written language proficiency (Andrews, 2012). Recognition of this variability, combined with difficulties in distinguishing between models, has increased interest in individual differences in word-level processing among skilled readers (e.g., Ashby, Rayner, & Clifton, 2005; Kuperman & Van Dyke, 2011; Yap, Balota, Sibley, & Ratcliff, 2012; Yap, Tse, & Balota, 2009). However, these studies have used a range of different measures to assess proficiency that vary in how effectively they capture the knowledge that Perfetti (2007) identified with lexical quality.

The chapter is structured around three broad questions that are central to evaluating the contribution of individual differences in lexical quality to skilled reading.

- Why is lexical quality important for skilled reading?
- What is lexical quality?
- How does lexical quality influence skilled reading?

Why Is Lexical Quality Important for Skilled Reading?

Reading is a culturally engineered skill (Carr, 1999) that was acquired too recently in the evolutionary history of our species to be part of our genetic blueprint (Dehaene, 2009). The first writing systems were invented about 5,000 years ago, and the alphabet is less than 4,000 years old. Humans are therefore not evolutionarily prepared to learn to read as they appear to be for comprehending and producing spoken language (Pinker, 1994). Reading is also acquired much later in a child's development than spoken language. Gough and Hillinger (1980) argued that this is because learning to read is "an unnatural act" that requires children to crack the code that connects written and spoken language so that they can harness their well-established spoken-language knowledge and skills for the new task of reading. Achieving these feats requires children to gain new metalinguistic insights, such as phonological awareness, that were not necessary for spoken-language acquisition (Ehri, this volume). These cognitive developments can be fostered by a supportive, literate environment. Learning to read is therefore subject to systematic influences of both the instructional and home environment (Sénéchal, this volume).

However, at least in cultures with universal access to education, genetic factors are the strongest predictor of early reading acquisition (Byrne et al., 2009). Behavioral genetic analyses of continuities and discontinuities in shared genetic and environmental variance, within a large-scale cross-national longitudinal study of twin pairs from before preschool until grade two, have demonstrated substantial shared genetic variance between decoding, spelling, and orthographic learning that is independent of the genetic variance associated with intelligence and significantly stronger than the variance due to environmental factors (Byrne et al., 2008). Phonological processing measures, such as phonological awareness and rapid naming speed, showed significant heritability at both preschool and kindergarten time points, demonstrating a genetic basis for the well-established contribution of phonological processes to early reading acquisition. However, analyses tracking changes to grade two showed that genetic effects of phonological awareness were shared with measures of orthographic knowledge, such as word decoding and spelling (Byrne et al., 2009). Byrne et al. (2008) argued that this common genetic variance is most parsimoniously attributed to a *learning rate* factor that influences the efficiency of learning associations between orthography and phonology.

Consistent with this evidence of genetic contributions to the development of orthographic knowledge, recent neuroimaging evidence from skilled readers has shown that the same highly localized brain regions respond selectively to written text across a wide range of languages. Meta-analyses of studies of readers of languages that vary in both script type and reading direction have identified a reproducible location of activation in the lateral occipito temporal sulcus. Dehaene (2009) argues that the uniformity of this *visual word form area* (VWFA) across individuals and languages presents a reading paradox: despite its evolutionary recency, reading involves "fixed cortical mechanisms that are exquisitely attuned to the recognition of written words . . . as though there was a cerebral organ for reading" (p. 4). Understanding how humans have developed a uniform solution to such a recent cultural invention is important not only for understanding reading itself but also for understanding how the brain adapts to a new cognitive skill.

The resolution to the reading paradox provided by Dehaene's (2009) *neuronal recycling hypothesis* is that the specialized visual word form system develops by partially recycling "a cortical territory evolved for object and face recognition" (Dehaene & Cohen, 2011, p. 254). This region is specialized for extracting features such as line orientation, junctions, and contours that are important for identifying objects. Such features are also nonaccidental properties of all writing systems because these cultural inventions have been shaped by the "learnability constraints" of the human brain (Dehaene, 2009). The anatomical connectivity of this region is also ideally suited to serve as the "brain's letterbox" in reading (Dehaene, 2009): filtering visual word forms from the perceptual stream and funneling them to temporal and frontal areas involved in phonological and semantic processing. Direct evidence for the emergence of the visual word form system with reading experience is provided by Dehaene et al.'s (2010) comparison of neural activity in populations varying in levels of literacy. This study found that the acquisition of literacy is associated with increased activation of the VWFA area by words and reduced activation of this region by faces and objects.

Convergent evidence for the important role of early visual processing in skilled reading is provided by neuroimaging comparisons of successful and unsuccessful readers. Pugh et al.'s (2001) review of such evidence concluded that successful reading was associated with higher levels of activation in the ventral visual processing areas corresponding to the VWFA during the early stages of word identification, as well as with stronger activation in temporoparietal areas of the dorsal visual system during later processing. However, successful readers showed less activation than disabled readers in frontal regions, centered around Broca's area, which are associated with phonological output processes such as fine-grained articulatory coding. Pugh et al. interpreted these data as indicating that the dorsal system plays an important role during the initial stages of reading as children learn to integrate the orthographic, phonological, and semantic features of words. Such integration enables development of the specialized "linguistically structured, memory-based word identification system" of the VWFA (Pugh et al., 2001, p. 245). Disabled readers rely on a different reading circuit because deficits in the temporoparietal processes required to integrate orthographic, phonological, and semantic features impede development of a well-structured VWFA system. They are therefore forced to rely on compensatory strategies indexed by increased activation of frontal areas associated with phonological processing. Pugh et al. suggest that this may reflect increased reliance on covert pronunciation to compensate for limitations in the automatic retrieval of the phonological codes required for comprehension.

The Cognitive Architecture of Skilled Reading

This selective overview of evidence about the neural circuitry underlying reading has important implications for the current discussion of individual differences in skilled reading because it demonstrates how the cognitive system of the developing reader is shaped by the process of learning to read. Dehaene et al. found changes in populations who acquired literacy in adulthood, emphasizing that "both childhood and adult education can profoundly refine cortical organization" (2010, p. 1359). This evidence of late plasticity provides a mechanism for the gradual refinements to the quality of lexical knowledge that is implied by the assumption that lexical quality is a word-specific property which "accumulates with age and experience" (Perfetti, 2007, p. 380). The VWFA's selective response to the visual features that define letters and words also demonstrates how orthography-specific features come to be embedded in the "mosaic of category-specific regions in ventral cortex" (Dehaene et al., 2010, p. 362) to take advantage of areas with appropriate receptive fields and anatomical connectivity to the brain regions that code semantic information. The VWFA therefore provides a neural mechanism that is compatible with the construct of orthographic precision emphasized by the LQH.

Pugh et al. (2001)'s systems-level analysis demonstrates how strengths and weaknesses in the processes required to integrate orthography with existing representations of spoken language can shape the functional architecture of the reading system. Successful readers develop a specialized automatic word identification system that supports fast, efficient access to all components of a word's identity. This reduces the need for compensatory, phonologically mediated word identification strategies, like those employed by less proficient readers.

This interpretation is consistent with the well-established interactive compensatory framework of individual differences in reading (Stanovich, 1992). This framework assumes that bottom-up decoding processes and top-down comprehension processes both contribute to word identification. The balance between them depends on a range of factors, including reading skill. In early applications of this framework, the central causal factor determining the balance between bottom-up and top-down processes was working memory. Automatic word decoding was assumed to benefit comprehension by reducing the drain on central processing resources so that they could be directed to comprehension. As reviewed earlier, Perfetti (2007) shifted the causal focus to the quality of lexical knowledge. High-quality representations support effective comprehension not simply because they support automatic retrieval but also because coherent, synchronous activation of the word's constituents provides the semantic and grammatical information required for comprehension. Readers who fail to develop such knowledge need to co-opt other processes, such as articulatory recoding and contextual prediction, to support both word identification and comprehension. This leads to differences in the *balance of labor* between automatic word identification and top-down processes.

The neuroimaging evidence therefore provides an answer to the question of why the precision and coherence of lexical knowledge is important to skilled reading. Precise knowledge about orthographic word forms provides rapid, specialized access to the tightly interconnected lexical information necessary for effective comprehension.

Lexical Quality and Models of Skilled Reading

This characterization of skilled reading is also consistent with modular views of lexical retrieval in reading (Forster, 1989), which assume that perceptual features are automatically mapped to lexical forms with minimal, if any, influence of context. This is the implicit or explicit assumption of many current theories of visual word recognition, which focus on retrieval of orthographic word forms without specifying how they interact with semantic knowledge (e.g., Coltheart et al., 2001; Davis, 2010; Perry et al., 2007).

Many models are based on the hierarchical interactive-activation (IA) architecture (McClelland & Rumelhart, 1981), which assumes that activation of orthographic word forms is a prerequisite for accessing semantic knowledge (see Taft, this volume). There are two senses in which these models implicitly assume that all skilled readers have developed high-quality representations for all words in their vocabulary. At the micro level, all words have fully specified connections between letter and word units, paralleling the precision criterion for lexical quality. At a systems level, few models allow the possibility that readers vary in the time course or weighting of different sources of knowledge, reflecting differences in the strength of connections between lexical constituents that might influence the coherence of their activation.

The limitation of these models from the perspective of the LQH is that they fail to accommodate the graded nature of lexical quality both within and between individuals or to recognize that the refinement of lexical representations may be a gradual process that continues at least into adolescence and shapes the cognitive architecture of the reading system. I will consider how the models might be modified to accommodate individual differences in lexical quality among skilled readers in the concluding section. But it is first necessary to consider how lexical quality should be defined and measured.

What Is Lexical Quality?

As reviewed earlier, Perfetti (2007) identified lexical quality as precise, stable, word-specific knowledge that supports coherent activation of all components of a word's identity. Although this definition incorporates both form and meaning, the neuroimaging evidence of a specialized system for processing the features of written words highlights the special status of orthographic knowledge in the development of literacy. This conclusion converges with theories of reading development that also emphasize the importance of precise orthographic representations in successful reading acquisition (Ehri, 2005, this volume).

Orthographic Precision and Lexical Quality

Phase theories of reading development (see Ehri, this volume) propose a sequence of transitions in children's representations of written words. As children acquire phonological awareness, their initial *prealphabetic* codes, based on idiosyncratic visual features and context (e.g., identifying *McDonalds* by the golden arches of its initial letter), are gradually connected to developing knowledge of letter–sound relationships to form *partial alphabetic* codes. To progress to the next stage of *full alphabetic* codes, in which all of a word's letters are connected to their pronunciations, requires children to be able to segment words into phonemes and map them to graphemes. Achieving this transition benefits from systematic phonics instruction (Ehri, Nunes, Stahl, & Willows, 2001) and training that emphasizes articulatory features (Boyer & Ehri, 2011). Full alphabetic codes are assumed to support accurate identification and spelling of most familiar words and relatively effective decoding of novel words.

However, fully automating word identification depends on a further transition to a *consolidated alphabetic phase* that is defined by *unitization* of larger orthographic chunks, such as syllables, rimes (e.g., ‹amp› in *camp, damp, stamp*), and morphological units such as roots (e.g., ‹mit› from *admit, submit*) and affixes (e.g., ‹pre›, ‹ed›). Word codes that are fully unitized can be read as whole units, reflected in reading speeds as fast as for single digits (Ehri & Wilce, 1983). Consolidated representations also enhance acquisition of new words by facilitating the use of analogy strategies. Representing multiple levels of orthographic-phonological (O-P) correspondence facilitates automatic word identification by securely bonding words' spellings to their pronunciations so that they are directly activated by print. It also reduces the number of links required to connect them—for example, *interesting* can be linked as four syllabic chunks rather than multiple grapheme-phoneme units (Ehri, 2005).

Ehri's description of the phases of lexical development aligns well with the LQH's assumption that high-quality lexical knowledge is acquired gradually at the level of the individual word. Transition between phases is gradual rather than discrete, so that a reader's vocabulary may include words at different phases of development. Ehri also emphasizes the importance of establishing strong connections between the different components of lexical identity. Like Perfetti (2007), she argues that this *amalgamation* of consolidated alphabetic codes with their semantic and syntactic features supports synchronous coactivation of these constituents, which facilitates automatic word identification,

The distinction between full and consolidated alphabetic representations fleshes out the critical construct of orthographic precision. According to Ehri, full alphabetic representations specify all the letters of a word and link them to phonology, the properties that Perfetti (2007) identified with precision. However, they do not necessarily support direct, automatic identification, particularly for words that have many orthographically similar word neighbors that contain many of the same letters. Consolidated alphabetic codes that represent multiple *grain size*s of O-P correspondence (Ziegler & Goswami, 2005) facilitate discrimination between similar words by reducing ambiguities in spelling-sound correspondence. In English, for example, many irregularities in O-P mapping are reduced when larger chunks are considered (Kessler & Treiman, 2001). For example, the grapheme ‹ea› takes many different pronunciations (e.g., *bead, dead, great*), but it is more likely to be pronounced /ɛ/ when followed by ‹d› than by any other letter (e.g., *dead* vs *dean; head* vs *heal*). Context also reduces ambiguity in mapping from phonology to orthography. For example, the phoneme /ʊ/ can be spelled in different ways, but ‹oo› is more likely before /k/ (e.g., *book, took, look*) while ‹u› is more common before /l/ or /ʃ/ (e.g., *full, bull, push, bush*).

The averaged performance of skilled readers is sensitive to these statistical patterns in both spelling-sound mapping, reflected in the pronunciations assigned to nonword items (e.g., Andrews & Scarratt, 1998; Treiman, Kessler, & Bick, 2003) and sound-spelling mapping, assessed in spelling of both nonwords and words (e.g., Treiman, Kessler, & Bick, 2002). Investigations of children have shown that this sensitivity develops gradually and appears to take longer to manifest in spelling than in reading: Vowel pronunciations of nonwords were significantly modulated by the following consonants as early as grade three (Treiman, Kessler, Zevin, Bick, & Davis, 2006), but corresponding effects on spelling performance only became evident between grades 5 and 7 (Treiman & Kessler, 2006). These findings are consistent with evidence that larger orthographic units play an increasing role in more skilled readers (Ehri, this volume). Treiman et al. (2002) showed that spelling ability was correlated with contextual influences on vowel spelling, but there is little systematic evidence about individual differences in sensitivity to these factors among skilled readers.

Structured orthographic codes that represent larger orthographic chunks such as rimes would facilitate discrimination between words that are similar at the level of individual letters and graphemes (e.g., *good* and *book* would be less similar to *goon* and *boot* than if they were coded purely as grapheme-phoneme units). They also contribute to securing the representation of the order of letters within the word. A representation that only codes individual grapheme-phoneme units is vulnerable to being confused with another word that contains the same letters in a different order (e.g., *calm/clam; eager/agree*). This vulnerability is reduced if the representation also codes rime (e.g., ‹am› is not contained in *calm*) or syllable units (e.g., *agree* does not contain ‹er›). Unitized morphemic units provide another dimension that can differentiate between similar words (e.g., *hunter* and *hunger* share five of their six letters, but contain different root morphemes). Morphologically structured orthographic representations also support the early morphological parsing and lemma extraction proposed in Taft's (this volume) AUSTRAL model.

The refinement of orthographic representations that results from integrating multiple grain sizes of mapping to phonology has been referred to as *lexical restructuring* (Ziegler & Goswami, 2005) and *lexical tuning* (Castles et al., 2007) to highlight its impact on the similarity relationships between orthographic forms. More finely tuned representations overlap less with the representations of orthographically similar words, making them less vulnerable to being activated by letter strings that contain similar components.

Thus, converging evidence from developing and skilled readers highlights the critical role of precise orthographic representations of words. However, there are major gaps between these bodies of evidence. Critically, there is little evidence about the later phases of reading development. Few developmental studies go past the early primary years, as if mastery of the alphabetic principle is sufficient to ensure success at reading. This view ignores the further refinements in orthographic learning required to achieve consolidated representations.

There are also differences in conceptual approach. Studies of reading development have attempted to identify the factors that predict orthographic learning but have paid little attention to underlying mechanisms (Castles & Nation, 2008). By contrast, the skilled-reading literature has used averaged data to build very detailed models of orthographic processing (see Grainger, 2008) but virtually ignored individual differences.

However, because orthographic learning operates at the item level (Share, 2004) and the acquisition of written vocabulary takes more than a decade (Nation, 2005), it must continue at least into adolescence (Andrews, 2012). Moreover, a substantial proportion of readers' written language vocabulary is acquired through reading. Of the 60,000 to 90,000 unique words known by the average high school graduate, as few as 200 are explicitly taught at school and less than 20,000 are encountered outside the school context (Landauer, Kireyev, & Panaccione, 2011). If rich, consolidated representations enhance effective application of analogy strategies to learn new words, as Ehri (this volume) argues, individual differences in the quality of lexical knowledge across readers will magnify over adolescence because readers who fail to develop consolidated representations for many words will be able to identify fewer words automatically and be less effective at developing stable representations of new words. This will encourage greater reliance on contextual prediction to identify words, which will, in turn, reduce the attention to orthography necessary to extract the larger orthographic chunks that contribute to consolidated representations.

Frith (1986) argued that this *partial cue reading strategy* of using coarse orthographic cues, supplemented by contextual prediction, is characteristic of good readers/poor spellers who demonstrate above average reading comprehension and knowledge of grapheme-phoneme correspondences but who perform well below their reading level on tests of spelling. A contextually driven strategy supports fast, efficient comprehension in many reading contexts but impedes extraction and unitization of the larger grain size units required for consolidated representations. Adoption of such a strategy is therefore likely to perpetuate and exaggerate individual differences in the precision of readers' orthographic representations. Differences in orthographic precision are likely to remain a source of variability in adult populations of competent readers (Ziegler & Goswami, 2005). The next section reviews evidence evaluating the validity of this hypothesis.

How Does Lexical Quality Influence Skilled Reading?

Most of the relatively sparse literature on individual differences among skilled readers has used measures of reading comprehension or vocabulary as predictor variables (e.g., Ashby et al., 2005; Perfetti & Hart, 2001; Yap et al., 2009, 2012). While such studies have provided useful evidence of systematic variability within this population, these relatively gross measures cannot provide direct evidence that these differences are due to lexical quality.

The most systematic body of evidence about individual differences in word identification among skilled readers has relied on measures of vocabulary, because it is an easily administered and calibrated index of lexical knowledge. It has therefore been included in a number of recent megastudies of skilled word recognition, including the English Lexicon Project (Balota et al., 2007), which collected measures of performance in a range of word identification tasks for large samples of words from participants from different universities (see Yap & Balota, this volume). To complement their analyses of the averaged data collected in this project, Yap, Balota, and colleagues have recently investigated how vocabulary modulates lexical decision and naming performance (Yap et al., 2012) and semantic priming effects (Yap et al., 2009). Vocabulary level is interpreted as an index of *lexical integrity*, which is identified with the coherence and stability that define high-quality lexical representations (Perfetti, 2007).

Using sophisticated analysis of response time data, Yap et al. (2009) showed that a sample of participants with high average vocabulary showed an additive relationship between word frequency and semantic priming, while a sample of lower vocabulary participants showed larger priming effects for low-frequency words, particularly for slower responses. Similarly, Yap et al.'s (2012) analysis of naming and lexical decision data showed that vocabulary was most strongly associated with variability in slow responses and that higher vocabulary was associated with reduced effects of frequency and O-P consistency. These findings were interpreted as showing that higher lexical integrity is associated with more automatic, modular lexical retrieval processes.

Similar conclusions have been drawn from studies using the combination of vocabulary and reading comprehension to index individual differences among skilled readers. Using the more naturalistic method of assessing eye movements during sentence reading, Ashby et al. (2005) compared the effects of predictability and frequency on the eye movements of highly skilled and average readers classified according to a median split on a composite measure of reading comprehension and vocabulary. The highly skilled group showed similar patterns of reading behavior for predictable, neutral, and unpredictable sentences, suggesting that they relied on automatic, context-independent word identification processes. By contrast, the average readers were strongly influenced by sentence context. Ashby et al. concluded that even within this sample of generally skilled readers, lower levels of proficiency were associated with increased reliance on context to compensate for less efficient word identification. They suggested that this contextually based strategy might "interfere with feedback processes that help form the crisp, stable … representations needed to access that word in the future" (p. 1083). This interpretation attributes the automatic, context-independent word identification strategy adopted by highly skilled readers to the precision of their orthographic representations. However, the measures of reading comprehension and vocabulary used to assess proficiency do not provide a direct measure of this construct.

Reading comprehension is a multifaceted skill that depends on domain-general processes such as working memory and inferential processes, as well as knowledge and processes that are specific to reading. Tests of reading comprehension vary in how effectively they assess these different components (Keenan, Olson, & Betjeman, 2008). Vocabulary size, breadth, and depth are also multiply determined and not specific to reading. Comprehension and vocabulary also tend to be at least moderately correlated and are therefore often combined to provide a more reliable index (e.g., Ashby et al., 2005). However, given the range of other cognitive processes that contribute to both measures, there is no guarantee that the variance they share is related to the quality or precision of lexical knowledge. Most notably, these measures fail to capture the critical construct of orthographic precision highlighted by both the neuroimaging and developmental evidence.

Isolating the Components of Lexical Quality

The goal of recent research in my laboratory has been to go beyond global measures of lexical quality to investigate specific contributions of the precision and coherence of lexical knowledge to skilled word identification and reading. To assess the contribution of word-specific orthographic knowledge, I have developed measures of spelling ability. Separate tests of spelling dictation and spelling recognition are combined to provide a broad index of spelling ability. These highly correlated measures (r between .7 and .8) both have high test-retest reliability ($r > .9$) and yield a wide range of scores in our university student samples (Andrews & Hersch, 2010).

A central feature of our approach to assessing orthographic precision has been to combine measures of spelling ability with tests of reading comprehension and vocabulary in order to isolate variance specific to orthographic precision. As would be expected, these three measures are intercorrelated: In a recent, typical sample of 200 students from my laboratory, vocabulary was moderately correlated (r = .54 and .55) with both reading comprehension and spelling ability, which were more weakly related (r =. 37). Nevertheless, by using reasonably large samples of participants and a small focused test battery, we have successfully applied multivariate regression procedures (Baayen, 2008) to distinguish variance common to all tests from variance unique to spelling. Using this approach, we have made substantial progress in demonstrating specific contributions of individual differences in orthographic precision to both word identification (Andrews & Hersch, 2010; Andrews & Lo, 2012) and sentence processing (Andrews & Bond, 2009; Hersch & Andrews, 2012; Veldre & Andrews, 2014, in press) and are now extending this work to distinguish individual differences in precision and lexical coherence (Andrews & Lo, 2013; Andrews, Xia, & Lo, 2014).

It is important to emphasize that our participants are screened to exclude non-native English speakers to ensure that variability is not due to bilingualism. Our participants also tend to be of superior reading ability. According to the US four-year college norms for the Nelson-Denny Reading Test (Brown, Fishco, & Hanna, 1993), our participants generally fall above the 50th percentile (e.g., medians of 80 and 86 for the sample of 200 referred to previously). We focus on this population of above-average readers because, as in other domains of expertise, understanding how *lexical experts* perform the task of reading provides insight into the optimal cognitive architecture for this complex skill (Andrews, 2008).

Another important feature of our approach has been the choice of experimental paradigms. Detecting individual differences among above average readers requires methods that test the limits of the reading system to reveal constraints and limitations that are masked or compensated for under less demanding conditions. At the same time, it is important to avoid tasks that are difficult because they place heavy demands on central resources. For example, a common method of increasing the demands of word identification is to present stimuli in a degraded form. Comparisons of the effects of degradation have also been used to investigate individual differences (Ziegler, Rey, & Jacobs, 1998). But high-quality representations facilitate both more efficient extraction of perceptual information and more effective use of educated guessing strategies based on partial information, which depend on central resources. Any interactions with individual difference observed in such tasks may reflect general differences in cognitive capacity rather than reading-specific effects.

The more general methodological problem in isolating processes associated with lexical quality is that more proficient readers tend to be better at all the processes required for effective reading. The LQH attributes causality to precise lexical representations but assumes that such representations also support effective higher-level processes. Many tasks conflate these sources of variability, making it difficult to disentangle the effects of precise lexical knowledge from the flexibility of integration and decision strategies that they afford. For example, meaning judgment tasks that require participants to judge the synonymy of word pairs (Perfetti & Hart, 2001) or decide whether a word is related to the meaning of a sentence (Gernsbacher & Faust, 1991) require a combination of lexical retrieval and decision processes. This makes it difficult to isolate the cause of any individual differences that are observed.

To avoid these problems, our recent research has relied on two major methodological approaches. To investigate the contributions of orthographic precision and lexical coherence to isolated word identification, we have used variants of the masked priming task. To assess how lexical quality contributes to sentence processing, we are investigating individual differences in eye movements during sentence reading. I focus primarily on masked priming data but also briefly describe some of our recent eye movement data.

Individual Differences in Masked Orthographic Priming

The masked priming paradigm developed by Forster and Davis (1984) is ideal for separating individual differences in early lexical retrieval from those due to decision processes. In this task participants respond to a clearly presented uppercase target stimulus, which is preceded by a brief (~50 ms) lowercase prime that is forward-masked by a symbol string (#####) and backward-masked by the target. The fact that responses are made to a clear, visible target eliminates the problems associated with using stimulus degradation to tax the system. Instead, the ability to extract information from the degraded, masked prime is assessed by investigating the impact of different types of primes on responses to the target. Because participants are usually unaware of the presence or identity of the prime, such influences are unlikely to reflect prediction or decision strategies. These higher order processes should affect overall speed and sensitivity to the impact of target attributes, while individual differences that affect early lexical retrieval should appear as interactions with priming manipulations.

Our initial investigations of the contribution of orthographic precision to lexical retrieval focused on *masked neighbor priming* (Andrews & Hersch, 2010) because data from this paradigm had been used in both the developmental and skilled reading literature to track the tuning of lexical representations with reading experience. Using the typical definition of orthographic neighbors as stimuli that share all but one letter with the target word, Castles, Davis, Cavalot, and Forster (2007) found that third-grade children showed facilitation from masked nonword neighbor primes (e.g., *pley-PLAY*), which disappeared by the time they were in fifth grade. Castles and colleagues interpreted this finding as suggesting that as children's written word vocabulary increases, their orthographic representations are fine-tuned to reduce the overlap between similar words. Similar conclusions derive from differences in skilled readers' masked priming for different items. Averaged data for samples of skilled readers typically show facilitatory masked neighbor priming for words with few neighbors (e.g., *eble-ABLE*), but not for targets (e.g., *tand-SAND*) that are similar to many words (Forster, Davis, Schoknecht, & Carter, 1987). The lack of priming for words with many neighbors suggests that the representations of these words may have been restructured by incorporating larger grain size units (Ziegler & Goswami, 2005) to enhance discrimination between similar words (Forster & Taft, 1994).

These data are consistent with the LQH's assumption that the acquisition of precise orthographic representations is a gradual process that varies between the words in an individual's vocabulary. We hypothesized that if the extent and success of the tuning process varies among skilled readers, as the LQH implies, individual differences in masked neighbor priming should still be evident within samples of skilled readers (Andrews & Hersch, 2010). Further, if spelling ability provides an index of the orthographic precision of word representations, it should be the best predictor of individual differences in neighbor priming. Our masked neighbor priming lexical decision data supported these predictions. The averaged data replicated Forster et al.'s (1987) finding that neighbor priming only occurred for targets with few neighbors. However, analyses including individual difference measures showed that the null priming for many neighbor targets in the averaged data obscured significant effects of spelling: good spellers showed inhibitory priming from neighbor primes, while poor spellers showed facilitation. We interpreted these findings as showing that spelling ability is selectively associated with inhibitory effects of lexical competition. However, the inhibitory priming shown by better spellers was specific to targets with many word neighbors. For targets with few neighbors, better spellers showed facilitation effects equivalent to those shown by poorer spellers.

This pattern is consistent with interactive-activation models, in which orthographic priming reflects the combination of facilitation due to sublexical overlap and inhibition due to lexical competition (Perry, Lupker, & Davis, 2008). Neighbor primes benefit target processing at the sublexical level because they preactivate most of the target's letters. However, if the prime becomes sufficiently strongly activated at the word level it will inhibit the representations of other similar words, potentially including the target word. Spelling ability appears to selectively predict the strength of the competition between words with many neighbors.

To provide further insight into the basis of these individual differences in orthographic precision, we subsequently compared the effects of neighbor primes with *transposed-letter (TL) primes*, which contain all the letters of the target word in a different order (e.g., *clam-CALM*). Transposed-letter items have played an important role in theories of visual word recognition because they provide insight into the relative contribution of letter identity and letter order to word recognition (Kinoshita, this volume). Averaged data for skilled readers has revealed that they are relatively poor at resolving letter order: TL nonwords (e.g., *jugde*) are often misclassified as words, and masked TL nonword primes (e.g., *jugde-JUDGE*) tend to facilitate performance more than neighbor primes (e.g., *jurge-JUDGE*) (e.g., Grainger, 2008). In contrast, TL *word* primes (e.g., *salt-SLAT*) often yield interference (e.g., Andrews, 1996). Such evidence has stimulated a substantial recent body of empirical research and computational modeling focused on "cracking the orthographic code" of skilled reading (see Grainger, 2008). However, this research has relied on averaged data. Our evidence that spelling ability modulates neighbor priming suggests that the averaged effects of TL priming may also obscure systematic individual differences.

To investigate these issues, we directly compared TL and neighbor primes for the same target words (Andrews & Lo, 2012). To disentangle facilitatory effects of sublexical overlap from lexical competition, we compared word and nonword primes (e.g., *plot, colt, CLOT; clib, cilp, CLIP*). The sample of 90 students in this study showed higher correlations between reading ability, spelling ability, and vocabulary (r = .64–.75) than those tested by Andrews and Hersch (2010). To separate the effects of general proficiency from effects specific to spelling, we therefore used principal components analysis to define two orthogonal individual difference measures: an index of overall proficiency that was essentially the average of the three composite scores and a second factor corresponding to the difference between spelling and reading/vocabulary that I will refer to as the *spelling-meaning* factor.

Analyses including these two independent individual difference measures showed that overall proficiency was associated with faster overall RT, confirming the expected association between higher proficiency and more efficient lexical classification, as well as stronger effects of prime lexicality. As summarized in Figure 10.1, which presents average priming effects for TL and neighbor primes relative to unrelated primes, lower proficiency participants showed

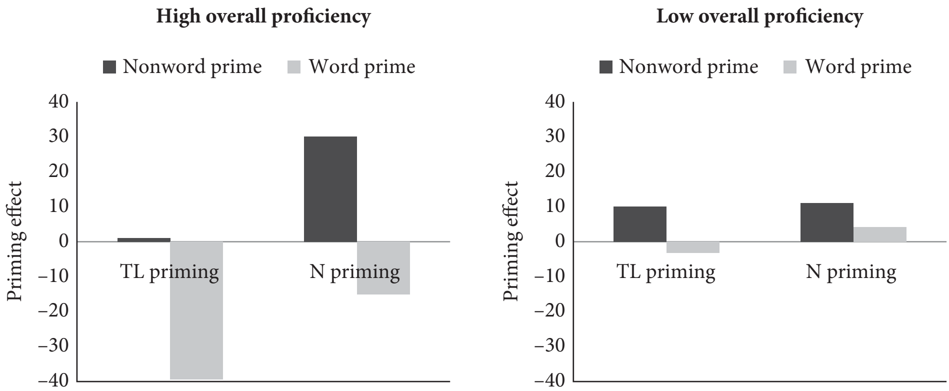


Fig. 10.1 Average transposed-letter (TL) and neighbor (N) priming effects (in ms) separately for participants above the median (high overall proficiency) and below the median (low overall proficiency) on the overall proficiency factor. Positive priming scores indicate facilitation relative to the unrelated prime condition, while negative scores indicate inhibition. (Adapted from data reported in Andrews and Lo, 2012.)

minimal, and equivalent, priming for both types of prime.[1] In contrast, high proficiency participants showed facilitatory priming from nonword neighbor primes (e.g., *clib-CLIP*) but inhibitory priming from TL word primes (e.g., *calm-CLAM*). Over and above these effects of overall proficiency, the second, spelling-meaning factor also significantly modulated priming. As depicted in Figure 10.2, individuals with unexpectedly high spelling ability for their level of reading/vocabulary showed inhibitory priming, particularly for TL primes, that was not evident in those with poorer spelling ability than expected from their reading/vocabulary scores. The latter subgroup showed facilitatory priming, particularly from neighbor primes.

These results complement and extend Andrews and Hersch's (2010) findings. By using principal components analysis, we identified two independent dimensions of individual difference that both significantly interacted with priming. The composite measure of overall written language proficiency was associated with faster performance and strong sensitivity to the lexical status of the prime and its orthographic relationship to the target. These data suggest that the dimension tapped by this index of overall proficiency is associated with rapid, efficient processing of the briefly presented prime. When the prime is a nonword neighbor that orthographically overlaps with the target (e.g., *clib-CLIP*), the shared sublexical constituents preactivate the target representation so that it is more quickly retrieved and classified when the target is presented. However, a TL nonword prime that contains all the letters of the target in a different order (e.g., *cilp-CLIP*)

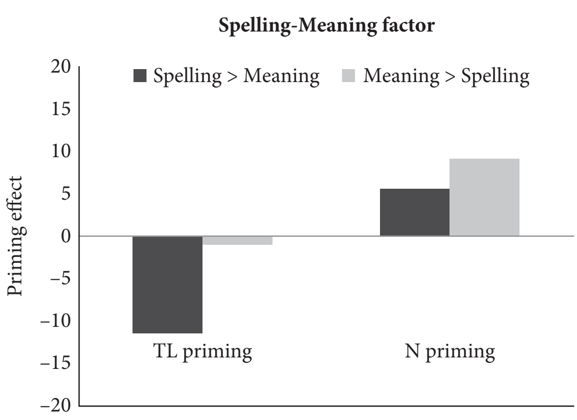


Fig. 10.2 Average transposed-letter (TL) and neighbor (N) priming effects (in ms) separately for participants above the median (spelling>meaning) and below the median (meaning>spelling) on the spelling-meaning factor. Positive priming scores indicate facilitation relative to the unrelated prime condition, while negative scores indicate inhibition. (Adapted from data reported in Andrews and Lo, 2012.)

creates perceptual ambiguity about letter position/order that reduces the extent of target preactivation. When the prime is a similar word, rapid prime processing allows the prime word's representation to become sufficiently strongly activated to inhibit the representation of the target word, counteracting the facilitation derived from sublexical overlap. If the prime is a word neighbor of the target (e.g., *plot-CLOT*), the net effect is no priming. But TL word pairs (e.g., *colt-CLOT*) yield strong inhibitory priming that systematically increases with reading proficiency. This result suggests that lexical competition may be stronger between more orthographically similar words, perhaps particularly those that differ only in letter order (Davis, 2010).

Andrews and Hersch (2010) found that lexical competition was selectively associated with spelling ability, but they did not directly evaluate the contribution of overall proficiency. Andrews and Lo's (2012) evidence that this dimension was also associated with stronger sublexical facilitation suggests that it taps broader aspects of lexical quality than orthographic precision. By combining spelling, reading comprehension, and vocabulary, this measure captures both orthographic and semantic aspects of lexical knowledge. It may therefore index the strength of connections between different lexical constituents that both Perfetti (2007) and Ehri (this volume) identify as critical to the coherent, automatic activation of the multiple constituents of a word's identity necessary for flexible, effective use of lexical knowledge.

Over and above the effects of overall proficiency, spelling also made an additional contribution to predicting priming reflected in stronger inhibition—particularly from TL primes—in individuals with unexpectedly high spelling ability for their level of reading and vocabulary. This independent source of variance in inhibitory priming was insensitive to prime lexicality, suggesting this factor may tap perceptual processes involved in resolving letter order that are specifically associated with orthographic precision. Many current models of orthographic coding attribute TL priming effects to the greater perceptual ambiguity of information about letter order than letter identity (e.g., Gomez, Ratcliff, & Perea, 2008; Norris & Kinoshita, 2012). The unique variance tapped by the spelling-meaning factor may reflect differences in the efficiency of resolving ambiguous information about letter order, which is a prerequisite for the rapid activation of prime word representations required to trigger inhibition of similar words. However, the strength of that lexical competition depends on broader properties of high-quality representations, such as tight constituent binding and coherence. Nonword primes yield diffuse activation across multiple letter and word representations that does not converge on a single lexical identity, as occurs for word primes. The more tightly interconnected the lexical constituents of the prime word, the more strongly and coherently they will be activated by the brief prime presentation, and the more they will inhibit representations of similar words. It is therefore the coherence and synchrony of the early activation associated with a stimulus, tapped by the overall proficiency factor, that distinguishes word from nonword primes and predicts effects of prime lexicality.

These results provided insights into the sources of variance differentiating skilled readers that go beyond the subgroup (Andrews & Bond, 2009) and regression (Hersch & Andrews, 2012) approaches that we have used previously. They provide evidence for independent contributions of orthographic precision, indexed by the spelling-meaning factor, and the broader aspects of lexical quality indexed by overall proficiency, which I will refer to as *lexical coherence*. In masked orthographic priming tasks, lexical coherence predicts both greater benefit from sublexical overlap and greater sensitivity to lexical competition. The second dimension of orthographic precision adds an additional, perhaps more perceptually based, component that is associated with stronger competition between similar words in individuals with the *orthographic profile* of higher spelling than reading/vocabulary ability.

The apparent paradox of identifying high lexical quality with competition processes that are detrimental to target identification is specific to the task demands of the masked priming paradigm. Under these conditions, rapid activation of the prime hurts performance because faster access to the prime leads to stronger inhibition of similar words, including the target. In normal reading, where readers do not receive contradictory perceptual input, rapid, precise lexical retrieval benefits reading (Andrews, 2012). As discussed later, orthographic precision also facilitates parafoveal processing during text reading (Veldre & Andrews, 2014, in press).

Individual Differences in Masked Morphological Priming

To provide further insight into the role of individual differences in orthographic precision and lexical coherence, we have begun to assess semantic processing of masked primes. The modular models of skilled lexical retrieval described earlier typically assume that processing follows a hierarchical, *form-first* sequence in which activation of orthographic forms precedes, and operates independently of, activation of meaning units at higher levels of the hierarchy (Forster, 2004). These assumptions have been most explicitly elaborated in models of morphological priming, which assume that complex words are segmented into morphemes that serve as access units to semantic and conceptual information about whole words. These *morphographic* models vary in their assumptions about the precise representational levels involved and how they interact (see Taft, this volume), but they share a commitment to an early orthographic decomposition process that is insensitive to semantic similarity (e.g., Rastle & Davis, 2008). Other researchers are equally committed to the alternative view that semantic information is available very early in processing and influences morphological segmentation and decision processes (e.g., Feldman, O'Connor, & Moscoso del Prado Martin, 2009).

Efforts to distinguish these competing theoretical claims have relied on comparisons of masked morphological priming for semantically *transparent* morphologically related pairs (e.g., *dreamer-DREAM*); *pseudomorphemic* pairs (e.g., *corner-CORN*), which have the same apparent morphological structure as transparent pairs but are not semantically related; and *orthographic control* pairs (e.g., *brothel-BROTH*) in which the prime word is not morphologically complex. An informal meta-analysis of masked priming studies comparing these conditions led Rastle and Davis (2008) to conclude that transparent and pseudomorphemic pairs showed equivalent priming relative to control pairs, as predicted by the orthographic decomposition account. However, a reanalysis of essentially the same data by proponents of the early semantic view concluded that priming was stronger for transparent than pseudomorphemic pairs (Feldman et al., 2009). The studies evaluated by both groups all relied on averaged skilled reader data.

We reasoned that these contradictory conclusions may reflect individual differences in orthographic precision and lexical coherence paralleling those observed in masked orthographic priming tasks (Andrews & Lo, 2013). Given Yap et al.'s (2009) evidence, reviewed earlier, that individual differences in vocabulary predict semantic priming, we used measures of spelling and vocabulary as predictors in order to tease apart individual differences in reliance on orthographic and semantic knowledge. Following the procedures in Andrews and Lo (2012), principal components analysis was used to separate overall proficiency from a second, independent factor reflecting the difference between spelling and vocabulary.

The results of Andrews and Lo (2013) showed that high overall proficiency predicted significantly faster responses to transparent stems but did not interact with any measures of priming. Significant interactions with morphological priming did, however, emerge for the second, spelling-meaning factor. As summarized in Figure 10.3, individuals with higher spelling than vocabulary showed equivalent priming for transparent and pseudomorphemic pairs, which both produced stronger priming than orthographic controls. However, those with the opposite profile of higher vocabulary than spelling showed priming for transparent but not pseudomorphemic or control pairs. Thus the orthographic profile was associated with the pattern predicted by form-first models, while the reverse *semantic profile* predicted the pattern identified with early semantic influences.

Andrews and Lo (2013) emphasized that the orthographic and semantic profiles should not be seen as distinct subgroups but as evidence of an independent source of individual differences over and above differences due to overall proficiency. In the morphological priming task overall proficiency did

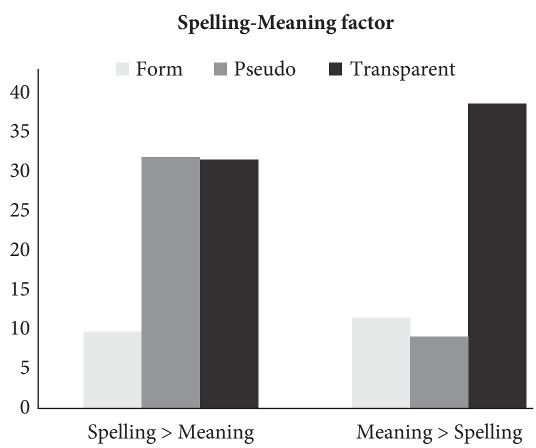


Fig. 10.3 Average morphological priming effects for the form, pseudomorphemic, and transparent conditions separately for participants above the median (spelling > meaning) and below the median (meaning > spelling) on the spelling-meaning factor. The positive priming scores indicate facilitation relative to the matched unrelated prime for each condition. (Adapted from data reported in Andrews and Lo, 2013.)

not interact with priming, perhaps because the critical stimuli were equated on orthographic overlap. Rather, differences in the pattern of priming were predicted by the independent spelling-meaning dimension that appeared to tap the availability and use of orthographic relative to semantic knowledge in lexical classification.

Individual Differences in Masked Semantic Priming

To provide further insight into this dimension of individual difference among skilled readers and assess its generality across tasks, we are currently investigating masked semantic priming in the semantic categorization task (Andrews et al., 2014). In contrast to the lexical decision task, which is relatively insensitive to semantic priming (e.g., Forster, 2004), semantic categorization judgments for broad categories like "animal" yield *category congruence effects* from masked primes, which implicate activation of semantic attributes of the prime (Quinn & Kinoshita, 2008). These effects are indexed by faster categorization of target words preceded by primes from the same category as the target: both faster "yes" responses to exemplar targets preceded by animal than nonanimal primes (e.g., *mole-EAGLE* vs. *boots-EAGLE*) and faster "no" responses to nonexemplars preceded by nonanimal than animal primes (e.g., *boots-RIFLE* vs. *mole-RIFLE*). Forster (2004) found that such effects only occurred for narrow, finite categories like single-digit numbers or months, for which decisions can be based on searching the entire category. However, Quinn and Kinoshita (2008) reported significant category congruence priming on animal categorization responses, but only for category congruent primes high in semantic feature overlap with the target (e.g., *hawk-EAGLE*; *pistol-RIFLE*).

We have recently replicated this evidence that high overlap semantic primes significantly facilitate average semantic categorization judgments, but we found that the extent and breadth of priming was significantly modulated by both the overall proficiency and spelling-meaning factors. Higher proficiency participants showed equivalent congruence priming for both high and low overlap primes (e.g., *pistol-RIFLE* and *boots-RIFLE*), while low proficiency participants only showed priming for high overlap primes. Significant interactions with the independent spelling-meaning factor showed that semantic priming was stronger in those with the semantic profile of higher vocabulary than spelling ability and entirely absent in individuals with above average spelling ability but below average vocabulary knowledge. These findings converge with those of Andrews and Lo (2013) in revealing individual differences in sensitivity to semantic attributes of masked primes. However, in contrast to our morphological priming data, in which semantic influences were predicted by the spelling-meaning dimension of variability, overall proficiency was a stronger predictor of priming in the semantic categorization task. The different contributions of the two dimensions of variability to predicting morphological and semantic priming may be due to task requirements.

The lexical decision task that we used to assess morphological priming appears to rely more on orthographic than semantic information (Balota, Cortese, Sergent-Marshall, & Spieler, 2004). This may account for why variability in morphological priming was predicted by the more specific spelling-meaning factor rather than overall proficiency. By contrast, congruence priming indexes facilitation from primes that share semantic but not orthographic features with the target. Such priming depends on the tight connections between different lexical constituents that enable rapid activation of the prime's semantic features, which facilitates rapid categorization of targets that share these features. This coherence dimension of lexical quality is indexed by overall proficiency, which was associated with stronger semantic priming, particularly from more distantly related category congruent primes. Nevertheless, paralleling Andrews and Lo (2013), there was an additional, independent contribution from the second spelling-meaning dimension reflecting an association between orthographic precision and reduced semantic priming.

There are at least two ways of interpreting the contributions of the spelling-meaning dimension of individual differences to masked semantic priming. One possibility is that the absence of semantic priming associated with the orthographic profile is a direct consequence of orthographic precision: Greater sensitivity to the orthographic features of the prime prevents preactivation of words that are orthographically different from the prime, even when they share semantic features. From this perspective, the reduced semantic priming may be specific to the masked priming task, which presents contradictory orthographic information without any cues to allow the system to distinguish between prime and target events (Norris & Kinoshita, 2012).

The second possibility is that the spelling-meaning factor indexes variability in the speed and efficiency of accessing semantic activation about words. This may be a lingering relic of the poor comprehender profile identified in developmental populations (Nation & Snowling, 1998) who have age-appropriate decoding skills but weak vocabulary and working memory. When the latter deficits are not too severe, cognitively able individuals, perhaps particularly those who are genetically endowed for efficient learning of orthographic-phonological correspondences (Byrne et al., 2008), achieve above average levels of reading skill but remain poorer at reading comprehension than expected from their word-level skills. From the perspective of the LQH, this profile reflects the precision required for rapid lexical retrieval without the synchronous, coherent activation of other components of word identity required for effective, flexible use of this knowledge.

According to this view, orthographic precision is an independent dimension of lexical quality that can coexist with deficits in word-level semantic knowledge that impede lexical coherence. Precision without coherence is not major disadvantage in the lexical decision task, where decisions depend on activation of the word's orthographic form, but is reflected in stronger sensitivity to discrepancies in letter order and sublexical morpheme units and reduced benefit from semantic overlap between morphologically related words. However, in the semantic categorization task, the reduced speed of activation of semantic information is reflected in reduced benefit from semantic feature overlap.

Individual Differences in Parafoveal Processing

The masked priming studies that I have reviewed provide clear evidence of individual differences in word-level processes. To evaluate whether they also influence reading of connected text, we are currently using eye-movement methods to investigate the contribution of lexical quality to sentence reading. The most exciting and informative strand of that research, led by Aaron Veldre, is focused on individual differences in parafoveal processing (see Schotter & Rayner, this volume). We first demonstrated that reading and spelling ability both predict skilled readers' *perceptual span* (Veldre & Andrews, 2014): the size of window of text that readers use during reading (Rayner, 2009). The combination of high reading and spelling ability was associated with a greater benefit from being provided with a larger window of text to the right, but not the left, of fixation, and a greater disruption from being given a very small window (Veldre & Andrews, 2014). This result indicates that higher overall proficiency is associated with greater reliance on parafoveal information.

We are now using gaze-contingent boundary paradigms (Cutter, Liversedge, & Drieghe, this volume) to investigate how individual differences modulate the nature of the information extracted from the parafovea. Investigations of individual differences in sensitivity to parafoveal previews of word and nonword neighbors converge with our masked priming data in showing that high proficiency readers who are better at spelling than reading show inhibitory parafoveal preview effects of word neighbors on early fixation measures (Veldre & Andrews, in press). This result suggests that, in the same way that orthographic precision facilitates rapid processing of masked primes leading to inhibition of the target word representation, precision also supports more rapid extraction of lexical information from the parafovea. When the parafoveal preview is a word neighbor of the target its identification inhibits orthographically similar words, leading to longer fixation durations on targets.

The eye movement measures also provide insight into the costs associated with lower quality lexical knowledge. Low scores on both factors were associated with facilitation from word previews on early fixation measures, paralleling the facilitatory masked priming effects shown by low proficiency readers. This result implies that these readers only extract sublexical features of parafoveal words. Low proficiency readers also showed inhibitory effects of word neighbor previews on later regressions to the target, suggesting that they failed to resolve the conflict between the target and the word neighbor preview, leading to disruptions in later integration processes.

Conclusions

The data reviewed in the preceding section provide converging evidence of systematic differences among skilled readers in the early stages of lexical retrieval tapped by the masked priming task. Varying the nature of the similarity relationships between the prime and the target and the decision requirements of the task revealed two independent dimensions of variance among skilled readers: an overall proficiency dimension, indexed by the shared variance between measures of reading comprehension, spelling and vocabulary, and a separate spelling-meaning factor indexed by the discrepancy between spelling ability and reading comprehension/vocabulary (Andrews & Lo, 2012, 2013; Andrews et al., 2014).

Both factors also significantly modulate skilled readers' extraction of parafoveal information during sentence reading (Veldre & Andrews, 2014, in press). The eye movement data provide insight into one of the mechanisms underlying the benefits of lexical quality for the reading process: Precise orthographic representations support rapid lexical retrieval, which in turn enables deeper processing of upcoming words (see "Future Directions").

Implications for Models of Visual Word Recognition

This evidence of systematic individual differences among skilled readers has important implications for computational models of visual word recognition that have simulated the average skilled reader, implying that all skilled readers read in the same way. The models that most obviously lend themselves to accommodating individual differences are the various *multiple-route* models of both reading aloud (e.g., Coltheart et al., 2001) and morphological priming (e.g., Diependaele, Sandra, & Grainger, 2005), which allow for differences between items as a function of lexical attributes such as frequency, regularity, and morphological structure. Within such frameworks, higher lexical quality can be conceptualized as an extension of these item differences: Readers with high-quality lexical representations have more functionally high-frequency words that are identified as whole units rather than through computational, sublexical processes. But such accounts describe rather than explain the source of the individual differences (Andrews & Lo, 2013).

An alternative, more perceptual approach to incorporating individual variability in lexical quality is to assume that readers vary in the efficiency with which they extract the visual features of written words and map them to representations in lexical memory. Many models of orthographic coding attribute TL priming effects to difficulties in resolving information about letter position (Gomez et al., 2008; Norris & Kinoshita, 2012), and this is clearly a parameter that could be varied to accommodate individual differences. Systematic modeling is required to determine whether individual differences in perceptual efficiency are sufficient to explain the differential orthographic priming from word and nonword primes and the modulation of semantic influences on morphological and category congruence priming. These effects appear to implicate individual differences in the online activation of lexical and semantic knowledge that depend on more than perceptual efficiency. However, these could be consequences of faster perceptual analysis of the prime. Within interactive activation models, word detectors must reach a threshold level of activation before they begin to inhibit other words. If this is achieved faster because of more efficient perceptual processing, there is a greater probability that the prime word will both inhibit other similar words to yield inhibitory orthographic priming and activate its semantic features to support category congruence effects. This analysis suggests that, within a hierarchical interactive-activation architecture, the cascading effects of individual differences in the efficiency of low level perceptual processes could potentially contribute to stronger inhibitory effects of masked orthographic primes and stronger facilitation from masked semantic primes. However, it remains to be seen whether such patterns can be simulated by modulating parameters governing low level perceptual processes. If not, other modifications, such as the addition of noise or weaker connectivity in higher layers of the network, may be necessary to effectively simulate the consequences of inefficient perceptual processes for the precision of lexical knowledge.

Another critical issue raised by the present data concerns the role of lexical competition in word identification. Such competition is central to interactive activation frameworks' account of lexical retrieval. Modulating the strength of lexical competition may simulate the association between overall proficiency and inhibitory neighbor priming (Perry et al., 2008). The additional assumption that competition is stronger between more orthographically similar words, embodied in Davis's (2010) spatial coding model, offers another potential source of individual differences in sensitivity to inhibition.

Fully understanding how individual differences in lexical quality emerge during reading development and shape the organization of lexical knowledge and the architecture of the reading system requires the development of computational accounts of how lexical knowledge is learned and refined, rather than further tweaking of models that stipulate the form and organization of lexical knowledge. Individual differences among skilled readers provide empirical constraints that should play an important role in future empirical and theoretical progress toward this goal.

Future Directions

Symbolic Versus Distributed Lexical Representations

From one perspective, the gradual refinement of lexical representations required to develop high-quality lexical knowledge seems compatible with *emergentist* theories of cognition such as PDP accounts, which attribute the rich, complex structure of human language and thought to simple processes that dynamically modify the weightings of connections between neuron-like processing units. McClelland et al. (2010) contrast these approaches with theories that assume abstract units corresponding to psycholinguistic entities such as phonemes, words, and concepts, and identify cognitive processing with the manipulation of these symbols. Proponents of emergentist approaches argue that these symbolic approaches are misleading approximations that fail to capture the dynamic, flexible similarity relationships that characterize real-world knowledge. However, acknowledging the dynamic interactions underlying the development of symbolic knowledge does not undermine the possibility that the "advantage of imposing a specific structural form on knowledge [may] outweigh its disadvantages" (McClelland et al., 2010, p. 354) and contribute to achieving expertise in highly practiced skills like identifying written words.

How Does Lexical Quality Contribute to Sentence Comprehension?

Skilled reading clearly involves more than word identification. The LQH assumes that high-quality representations directly contribute to reading by supporting automatic retrieval of the lexical knowledge required for effective comprehension. Support for this view is provided by our recent eye-movement evidence, reviewed briefly earlier, showing that skilled readers make more use of parafoveal information than less skilled readers (Veldre & Andrews, 2014; Veldre & Andrews, in press). Higher quality lexical knowledge supports rapid foveal processing, which in turn allows more efficient extraction of parafoveal information that facilitates processing of subsequently fixated words and enhances the efficiency and effectiveness of oculomotor planning. We are conducting further investigations of individual differences in skilled readers' eye movements to better understand the interplay between early lexical retrieval and the integrative processes required for comprehension.

Does Orthographic Precision Matter for All Languages?

Orthographic precision may not be equally important for all languages. According to Ziegler and Goswami's (2005) *psycholinguistic grain size theory,* differences in the consistency of the relationship between orthography and phonology determine the grain size of the O-P units that readers need to extract. The drive to develop consolidated representations may therefore be stronger in *deep orthographies* like English, in which inconsistent O-P correspondences "force the reading system into developing multiple grain size mappings" (Ziegler & Goswami, 2005, p. 18) to achieve the precision and redundancy required for automatic word identification. Consistent with the LQH, this pressure toward lexical restructuring is assumed to be word-specific: Words that have many similar neighbors will experience more pressure to reflect multiple O-P correspondences than those from less dense orthographic neighborhoods. The grain size account also assumes that lexical restructuring is an ongoing process that shapes the long-term organization and dynamics of the adult reading system.

Arguing that languages differ in the extent to which they drive the extraction and representation of larger orthographic units does not contradict claims for a universal theory of reading (Frost, 2012). Effective processing of all writing systems depends on learning to efficiently map orthography to phonology. However, the specific strategies and representational structures required to achieve this depend on the covariations between the phonology, orthography, and semantics of the language. Our data show that individual differences among skilled readers provide insight into the lexical structures that emerge for readers of English. Comparisons of individual differences in skilled readers of different writing systems will further contribute to understanding how universal cognitive constraints interact with the characteristics of different languages to shape the cognitive architecture of the skilled reading system.

Note

1 The analyses reported by Andrews and Lo (2012, 2013) used continuous measures of both the overall proficiency factor and the spelling-meaning factor. To summarize the significant interactions of priming with these individual difference measures, Figures 10.1, 10.2, and 10.3 present average priming effects for participants above and below the median on the relevant factor.

References

Andrews, S. (1996). Lexical retrieval and selection processes: Effects of transposed-letter confusability. *Journal of Memory and Language, 35,* 775–800.

Andrews, S. (2008). Lexical expertise and reading skill. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 49, pp. 247–281). San Diego, CA: Elsevier.

Andrews, S. (2012). Individual differences in skilled visual word recognition and reading: The role of lexical quality. In J. Adelman (Ed.), *Visual word recognition: Volume. 2. Meaning and context, individuals and development* (pp. 151–172). Hove, England: Psychology Press.

Andrews, S., & Bond, R. (2009). Lexical expertise and reading skill: Bottom-up and top-down processing of lexical ambiguity. *Reading and Writing: An Interdisciplinary Journal, 22,* 687–711.

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General, 139,* 299–318.

Andrews, S., & Lo, S. (2012). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 152–163.

Andrews, S., & Lo, S. (2013). Is morphological priming stronger for transparent than opaque words? It depends on individual differences in spelling and vocabulary. *Journal of Memory and Language, 68,* 279–296.

Andrews, S., & Scarratt, D. R. (1998). Rule and analogy mechanisms in reading nonwords: Hough dou peapel rede gnew wirds? *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1052–1086.

Andrews, S., Xia, V., & Lo, S. (2014, November). *Individual differences in masked semantic priming: The role of lexical quality.* Paper presented at the meeting of the Psychonomic Society, Long Beach, CA.

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *Quarterly Journal of Experimental Psychology, 58A,* 1065–1086.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, England: Cambridge University Press.

Balota, D. A., Cortese, M., Sergent-Marshall, S., & Spieler, D. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283–316.

Balota, D. A., & Spieler, D. (1998). The utility of item-level analyses in model evaluation. *Psychological Science, 9*, 238–240.

Balota, D. A., Yap, M., Cortese, M., Hutchinson, K., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459.

Boyer, N., & Ehri, L. (2011). Contribution of phonemic segmentation instruction with letters and articulation pictures to word reading and spelling. *Scientific Studies of Reading, 15,* 440–470.

Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *Nelson-Denny reading test*. Rolling Meadows, IL: Riverside.

Byrne, B., Coventry, W. L., Olson, R. K., Samuelsson, S., Corley, R., Wilcutt, E. G., . . . DeFries, J. (2009). Genetic and environmental influences on aspects of literacy and language in early childhood: Continuity and change from preschool to grade 2. *Journal of Neurolinguistics, 22,* 219–236.

Byrne, B., Olson, R. K., Hulslander, J., Wadsworth, S., Defries, J. C., & Samuelsson, S. (2008). A behavior-genetic analysis of orthographic learning, spelling and decoding. *Journal of Research in Reading, 31,* 8–21.

Carr, T. H. (1999). Trying to understand dyslexia: Mental chronometry, individual differences, cognitive neuroscience and the impact of instruction as converging sources of evidence. In R. M. Klein & P. A. McMullen (Eds.), *Converging methods for understanding reading and dyslexia* (pp. 459–491). Cambridge, MA: MIT Press.

Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers. *Journal of Experimental Child Psychology, 97,* 165–182.

Castles, A., & Nation, K. (2008). Learning to be a good orthographic reader. *Journal of Research in Reading, 31,* 1–7.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review, 117,* 713–758.

Dehaene, S. (2009). *Reading in the brain*. New York, NY: Penguin Viking.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Science, 15,* 254–262.

Dehaene, S., Pegado, F., Braga, L., Ventura, P., Filho, G., Jobert, A., … Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science, 3,* 1359–1364.

Diependaele, K., Sandra, D., & Grainger, J. (2005). Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes, 20,* 75–114.

Eason, S. H., & Cutting, L. E. (2009). Examining sources of poor comprehension in older poor readers. In R. K. Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.), *Beyond decoding: The behavioral and biological foundations of reading comprehension* (pp. 263–283). New York, NY: Guilford.

Ehri, L. (2005). Development of sight word reading: Phases and findings. In M. J. Snowling & C. Hulme (Eds.), *The science of reading* (pp. 135–154). Cornwall, England: Blackwell.

Ehri, L., Nunes, S., Stahl, S., & Willows, D. (2001). Systematic phonics instruction helps students learn to read. *Review of Educational Research, 71,* 393–447.

Ehri, L., & Wilce, L. (1983). Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology, 75,* 3–18.

Feldman, L. B., O'Connor, P. A., & Moscoso del Prado Martín, F. (2009). Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review, 16,* 684–691.

Forster, K. I. (1989). Basic issues in lexical processing. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 75–107). Cambridge, MA: MIT Press.

Forster, K. I. (2004). Category size effects revisited: Frequency and masked priming effects in semantic categorization. *Brain and Language, 90,* 276–286.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 680–698.

Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology, 39,* 211–251.

Forster, K. I., & Taft, M. (1994). Bodies, antibodies and neighborhood-density effects in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 844–863.

Frith, U. (1986). A developmental framework for developmental dyslexia. *Annals of Dyslexia, 36,* 69–81.

Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences, 35,* 263–279.

Gernsbacher, M., & Faust, M. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 245–262.

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review, 115,* 577–600.

Gough, P., & Hillinger, M. (1980). Learning to read: An unnatural act. *Bulletin of the Orton Society, 30,* 179–196.

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes, 23,* 1–35.

Hersch, J., & Andrews, S. (2012) Lexical expertise and reading skill: Bottom-up and top-down contributions to sentence processing. *Scientific Studies of Reading, 16,* 240–262.

Landauer, T., Kireyev, K., & Panaccione, C. (2011). Word maturity: A new metric for word knowledge. *Scientific Studies of Reading, 15,* 92–108.

Keenan, J. M., Olson, R. K., & Betjemann, D. (2008). Assessment and etiology of individual differences in reading comprehension. In R. K. Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.), *Beyond decoding: The behavioral and biological foundations of reading comprehension* (pp. 227–245). New York, NY: Guilford.

Kessler, B., & Treiman, R. (2001). Relationships between sounds and letters in English monosyllables. *Journal of Memory and Language, 44,* 592–617.

Kuperman, V., & Van Dyke, J. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language, 65,* 42–73.

McClelland, J., Botvinick, M., Noelle, D., Plaut, D. C., Rogers, T., Seidenberg, M., & Smith, L. (2010). Letting structure emerge: Connectionist and dynamical systems approaches to cognition. *Trends in Cognitive Sciences, 14,* 348–356.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

Nation, K. (2005). Children's reading comprehension difficulties. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 248–265). Oxford, England: Blackwell.

Nation, K., & Snowling, M. (1998). Semantic processing and the development of word recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39,* 85–101.

Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: Why there's nothing special about the perception of orthography. *Psychological Review, 119,* 517–545.

Perfetti, C. A. (1985). *Reading ability*. New York, NY: Oxford University Press.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11,* 357–383.

Perfetti, C., & Hart, L. (2001). The lexical basis of comprehension skill. In D. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). Washington, DC: American Psychological Association.

Perry, C., Ziegler, J., & Zorzi, M. (2007). Nested modeling and strong inference testing in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 27,* 301–333.

Perry, J. R., Lupker, S. J., & Davis, C. J. (2008). An evaluation of the interactive-activation model using masked partial-word priming. *Language and Cognitive Processes, 23,* 36–68.

Pinker, S. (1994). *The language instinct*. Cambridge, MA: Harvard University Press.

Plaut, D., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Pugh, K. R, Mencl, W. E., Jenner, A. R., Lee, J. R., Katz, L., Frost, S. J., . . . Shaywitz, B. A. (2001). Neuroimaging studies of reading development and reading disability. *Learning Disabilities Research and Practice, 16,* 240–249.

Quinn, W., & Kinoshita, S. (2008). Congruence effects in semantic categorization with masked primes with narrow and broad categories. *Journal of Memory and Language, 58,* 286–306.

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes, 23,* 942–971.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62,* 1457–1506.

Share, D. L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology, 87,* 267–298.

Stanovich, K. E. (1992). Speculations on the causes and consequences of individual differences in early reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 307–342). Hillsdale, NJ: Erlbaum.

Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology, 98,* 642–652.

Treiman, R., Kessler, B., & Bick, S. (2002). Context sensitivity in the spelling of English vowels. *Journal of Memory and Language, 47,* 448–468.

Treiman, R., Kessler, B., & Bick, S. (2003). Influence of consonantal context on the pronunciation of vowels. *Cognition, 88,* 49–78.

Treiman, R., Kessler, B., Zevin, J., Bick, S., & Davis, M. (2006). Influence of consonantal context on the reading of vowels: Evidence from children. *Journal of Experimental Child Psychology, 93,* 1–24.

Veldre, A., & Andrews, S. (2014). Lexical quality and eye movements: Individual differences in the perceptual span of skilled adult readers. *Quarterly Journal of Experimental Psychology, 67,* 703–727.

Veldre, A., & Andrews, S. (in press). Parafoveal activation depends on skilled reading proficiency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 53–79.

Yap, M. J., Tse, C.-S. & Balota, D. A. (2009). Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language, 61,* 303–325.

Ziegler, J., & Goswami, U. (2005). Reading acquisition, developmental dyslexia and skilled reading across language: A psycholinguistic grain size theory. *Psychological Review, 131,* 3–29.

Ziegler, J., Rey, A., & Jacobs, A. (1998) Simulating individual word identification thresholds and errors in the fragmentation task. *Memory & Cognition, 26,* 490–501.

CHAPTER

11

What Does Acquired Dyslexia Tell Us About Reading in the Mind and Brain?

Anna M. Woollams

Abstract

Reading is a fundamental cognitive skill that is often disrupted as a consequence of brain damage. The study of neurological patients with acquired reading disorders has proven pivotal in development of theoretical accounts of normal reading. This work initially involved a focus on cases of dissociation between reading and other cognitive functions using single-case methodology. This evidence was influential in the formation of dual-route models of reading aloud, which employ localist representations. More recent work has used simultaneous consideration of multiple cases to reveal associations between reading and other cognitive functions. This evidence has been captured by connectionist triangle models of reading aloud, which rely on learned distributed representations. Neuroimaging of patients with acquired dyslexia has provided insights into the mechanisms of dysfunction and the neural basis of normal reading. Consideration of neuropsychological patient data has highlighted the role of more basic perceptual and cognitive processes in skilled reading.

Key Words: reading, computational modeling, cognitive neuropsychology, pure alexia, phonological dyslexia, surface dyslexia, deep dyslexia

Fluent reading is a complex process that begins by accessing the meaning and sounds of single words (Stanovich, 1982). Brain damage often undermines reading ability, in many cases even at the most basic single-word level. The creation of cognitive theories of reading was fundamentally guided not only by normal behavior (Baron & Strawson, 1976; Forster & Chambers, 1973) but also by the performance of patients with *acquired dyslexia* (Marshall & Newcombe, 1973, 1981), which is defined as a disorder of reading that emerges after brain damage in formerly literate individuals. The study of such patients is part of the broader field of cognitive neuropsychology, which is based on the rationale that any adequate theory of normal information processing should also be able to account for its disorders. Normal reading performance can be explained by a variety of different theories, and data from patients with acquired dyslexia have provided crucial insights into the structure and function of the reading system. Both behavioral and lesion data from neurological patients have also drawn our attention to how reading is implemented in the brain, as neuropsychological evidence has a unique role to play in indicating which brain areas are necessary for a particular function. This chapter aims to illustrate the influence of neuropsychological data from patients with acquired dyslexia on the development of computational models of reading aloud in English.

Types of Acquired Dyslexia

Reading involves a variety of component processes that can be disrupted by brain damage. If these disorders are to be informative concerning the nature of the normal reading system, it must be supposed that there is sufficient localization of function within the brain that damage can selectively undermine individual components. It is useful to employ a distinction that was drawn early in the neuropsychological literature between *peripheral dyslexias*, which refer to a failure to derive a satisfactory internal visual representation of the letter string (see Kinoshita, this volume, for discussion of orthographic coding), and *central dyslexias*, which involve damage to phonological processing or semantic access/representation (Shallice & Warrington, 1980). This distinction reflects the extent to which the cause of the reading difficulty involves cognitive components that are core to the system that supports spoken language processing. Across both classes of dyslexia, deficits are most often defined in terms of the accuracy of reading single words aloud, although of course these impairments will undermine silent reading of connected text (see Pollatsek, this volume).

Peripheral Dyslexias

The peripheral dyslexias involve impairment in a person's ability to accurately encode the letter strings that form the basis for the reading process. The idea is that the root cause of the reading problem for these dyslexics is in more general cognitive processes that fall outside the language system and that affect orthographic processing. Across the peripheral dyslexias, we see variation in the extent to which the primary deficit affects reading of connected text versus isolated words.

In *attentional dyslexia*, associated with lesions to the left parietal lobe (Hall, Humphreys, & Cooper, 2001; Warrington, Cipolotti, & McNeil, 1993), there is a problem with simultaneous perception of multiple elements in a stimulus. At the text level, this means that such patients have difficulty accurately reading aloud adjacent words, while at the word level these patients have difficulty accurately encoding adjacent component letters. A distinctive feature of this disorder is the presence of strong effects of surrounding elements on perception of not just letters, but also numbers and symbols (Shallice & Warrington, 1977; Warrington et al., 1993), such that elements that can be identified accurately in isolation (e.g., the letter ‹R›) cannot be identified when surrounded by comparable stimuli (e.g., ‹R› in ‹CLRNS›).

In *neglect dyslexia*, typically associated with lesions to the right parietal lobe (Ptak, Di Pietro, & Schnider, 2012), patients show a tendency to overlook information in the left visual field. At the text level, this means that such patients have trouble finding the beginnings of lines and will often omit the initial words of lines, while at the word level these patients will often omit or misreport the initial letters of words. These patients show the same kind of deficits in nonreading and nonlinguistic tasks like picture description and marking the middle position of lines, and these additional deficits go hand in hand with their reading problems (Anderson, 1999; McGlinchey-Berroth et al., 1996).

Hemianopic alexia is somewhat unusual in that it primarily affects text processing with little disruption at the single-word level. This is in contrast to the disruption to both levels seen in attentional and neglect dyslexia. *Hemianopia* refers to the loss of some portion of the visual field and is usually seen after a posterior cerebral artery stroke in the left hemisphere that damages primary visual cortex (Leff, Spitsyna, Plant, & Wise, 2006; Pflugshaupt et al., 2009). At the text level, a right hemianopia means that fixation in text is made more difficult as it disrupts the visual input that contributes to parafoveal preview and oculomotor planning that determines the position of next fixation (Pflugshaupt et al., 2009). However, at the single-word level, at least with relatively short words, a patient with hemianopia can fixate toward the end of the string, which appears to permit relatively fluent oral reading (Leff et al., 2006).

Pure alexia contrasts with hemianopic alexia in that perception of single words is the primary deficit; in this sense, it clusters with the central dyslexias that are described later. It is distinguished from these, however, by a focus on speed rather than accuracy of single-word processing, such that it is defined as very slow single-word reading with an abnormally large effect of letter length on reaction time, in contrast with the minimal length effects seen in normal skilled reading of words (Weekes, 1997). Patients with pure alexia often adopt a laborious letter-by-letter reading strategy, in which they overtly name each letter before reading the word aloud. While it has been suggested that such patients have difficulties specifically in word identification, letter identification in these patients is not always accurate (Cumming, Patterson, Verfaellie, & Graham, 2006; Woollams, Hoffman, Roberts, Lambon Ralph, & Patterson, 2014).

Pure alexia is a striking disorder that appears to be extremely reading specific. These patients are unable to rapidly and accurately identify the component letters in written strings but are seemingly able to identify familiar objects, process spoken language, and write fluently. It is associated with damage to the left ventral-occipitotemporal cortex (Roberts et al., 2013), which responds strongly to written letter strings in neuroimaging studies of normal reading (Dehaene & Cohen, 2011). Yet despite its apparent reading specificity, deficits in processing visual stimuli that have similar processing demands to letter strings do emerge in pure alexia when speed of response is considered. Written letter strings represent a class of complex visual stimuli that are composed of a small set of highly confusable elements, which require high acuity to accurately perceive. Pure alexic patients often have reduced sensitivity to the medium to high spatial frequencies that support foveal vision (Roberts et al., 2013), and indeed their lesion encompasses the neural region shown to be particularly responsive to high relative to low spatial frequencies (Woodhead, Wise, Sereno, & Leech, 2011). When patients with pure alexia are required to identify pictures of objects high in visual complexity, they are in fact impaired (Behrmann, Nelson, & Sekuler, 1998). The same pattern is seen when pure alexic patients are required to match complex abstract visual patterns such as checkerboards or logographic characters (Mycroft, Behrmann, & Kay, 2009; Roberts et al., 2013). Moreover, the extent of their impairment on this task correlates with the severity of their reading impairment (Roberts et al., 2013).

Central Dyslexias

The central dyslexias are each identified by a different profile of reading accuracy across word classes and distinctive error types. They share the feature that the underlying deficit is damage to phonological or semantic processes that also support spoken language processing. Although the original definition emphasizes the overlap between the processes involved in reading and speaking, some research has suggested that the reading disorder can occur independently of speech processing deficits. As there are differences between current models in terms of the degree of reading specificity of their components, these claims have been vigorously debated, as will be considered in more detail later.

Surface dyslexia is defined by a relatively selective deficit in the reading aloud of exception words (Marshall & Newcombe, 1973), which are those that contain components with atypical mappings between spelling and sound at the level of the grapheme and phoneme (e.g., *crook* vs. *broom*) and the body and rime (e.g., *blood* vs. *food*). Although these are separable dimensions of regularity and consistency, respectively (see Yap & Balota, this volume), it appears that both influence performance in surface dyslexia (Jefferies, Ralph, Jones, Bateman, & Patterson, 2004). This stands in contrast to much better or even normal reading of regular words with typical mappings between spelling and sound (e.g., *black, broom)* and also good reading of novel letter strings or nonwords (e.g., *preak, splof)*. This deficit for exception words is particularly pronounced when they are lower in frequency. The hallmark error in surface dyslexia is regularization, where an exception word is pronounced as if it had typical mappings between spellings and sounds, such as *sew* read as "sue." Surface dyslexia has most often been reported among patients suffering from *semantic dementia*, which is associated with atrophy of the anterior temporal lobes (Wilson et al., 2012, 2009; Woollams, Lambon Ralph, Plaut, & Patterson, 2007).

Phonological dyslexia, in contrast to surface dyslexia, is defined by a relatively selective deficit in reading nonwords aloud (e.g., *preak, splof*) (Beauvois & Derouesne, 1979). Reading of familiar words is considerably better, and may even fall within normal limits, although this is unusual. Word performance is more accurate for those words that are high in frequency than those low in frequency. Performance is also somewhat better for words with typical (e.g., *black, broom*) as opposed to atypical (e.g., *crook, blood*) mappings between spelling and sound (Crisp & Lambon Ralph, 2006). Accuracy is strongly influenced by the meaning of a word, such that words with high imageability referents (e.g., *table, chair*) are read more accurately than those with low imageable referents (e.g., *poem, hope*) (Crisp & Lambon Ralph, 2006). The hallmark error in phonological dyslexia is lexicalization, whereby a novel letter string like *stoip* is read as a familiar item like *stoop.* Phonological dyslexia is most often observed in patients with chronic spoken language problems that arise due to damage to some portion of the left perisylvian phonological processing network (Rapcsak et al., 2009).

Deep dyslexia is characterized by an extremely severe nonword reading deficit, with relatively preserved reading performance seen for words, particularly those high in imageability. This syndrome is distinguished from very severe phonological dyslexia by the presence of the highly distinctive semantic error, whereby a familiar word like *blood* is misread as a semantically related but formally distinct word such as *heart* (Marshall & Newcombe, 1973). Deep dyslexia is a relatively rare syndrome that is associated with very large lesions that encompass multiple aspects of the left hemisphere language network (Price et al., 1998).

Dissociations

Within cognitive neuropsychology, the observation of a dissociation between a patient's performance across different tasks or item types is taken as evidence of the functional independence of the underlying cognitive processes. Hence the observation of impaired word reading with intact object naming, as reported in some cases of pure alexia, is taken as evidence for the independence of written word and object recognition (e.g., Miozzo & Caramazza, 1998). Of course, it is possible that word reading is simply a more difficult task. Much stronger evidence of independence comes from contrasting the performance of a pure alexic patient with that of a *visual agnosic* patient. *Visual agnosia* is defined as an inability to recognize visually presented objects, but some cases have been found to show intact reading (e.g., Gomori & Hawryluk, 1984). Taken together, the two patients provide a double dissociation between written word recognition and object recognition, which implies that different cognitive and neural mechanisms support each process.

Turning to dissociations over items, in the case of surface dyslexia we see impaired exception word reading in the face of intact nonword reading (McCarthy & Warrington, 1986). This single dissociation suggests that there is a component of the reading system that deals specifically with mappings for familiar items that supports reading of exception words, distinct from a component able to process mappings for unfamiliar items that supports reading of nonwords. Of course, it could be that exception word reading is simply harder than nonword reading. However, we can rule out this possibility if we contrast surface dyslexia with phonological dyslexia, where we see impaired nonword reading in the face of intact exception word reading (Funnell, 1983). The double dissociation that this contrast provides suggests that independent cognitive and neural mechanisms are required to process exception words and nonwords. However, the interpretation of double dissociations in terms of functional independence has been challenged, because it has been shown that these can arise from differential damage within a common system (Plaut, 1995, 2003; Van Orden, Pennington, & Stone, 2001). Nevertheless, the double dissociation between surface and phonological dyslexia has proven fundamental to the formulation of current computational models of reading aloud.

Dual-Route Cascaded Model

Early case study reports of patients with surface and phonological dyslexia (Beauvois & Derouesne, 1979; Marshall & Newcombe, 1973; Shallice & Warrington, 1975; Warrington, 1975) formed the basis of the proposal of the dual-route model of reading aloud (Morton, 1980). This theory has as its central assumption that there are two functionally independent direct pathways between spelling and sound (Morton, 1980), one at the whole-word level and one at the subword level. Initially expressed in box and arrow notation (Patterson & Shewell, 1987), one version of this view has been implemented in a large-scale computational model, known as the dual-route cascaded (DRC) model, shown in the top portion of Figure 11.1 (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). Within this model, reading begins with the identification of the letters of the printed input string via the visual feature analysis and then the letter identification modules. The direct lexical route maps between whole-word orthographic and phonological representations, stored in the orthographic input lexicon and phonological output lexicon respectively, and it is this pathway which allows correct reading of exception words. The orthographic input lexicon can activate the semantic system, which can then activate the phonological output lexicon. It is this lexical-semantic pathway that allows access to word meaning. Critically, because word meaning is not considered to be necessary for reading aloud within this theory, the lexical-semantic route remains unimplemented within the model. Because the orthographic input lexicon contains localist representations that involve a specific unit for each known word, neither the direct lexical nor lexical-semantic pathways can, by definition, pronounce nonwords. The processing of novel letter strings therefore requires an additional pathway, termed the nonlexical route, which consists of a series of grapheme-to-phoneme conversion rules. These rules are applied serially from left to right across the input string. The outputs from the pathways are then combined in the phoneme system to permit a spoken response.

The DRC model is capable of capturing a variety of key aspects of normal reading aloud (see Coltheart et al., 2001). Word frequency influences performance as the resting levels of lexical entries are frequency indexed (Forster & Chambers, 1973). Exception words are processed more slowly than regular words because the correct lexical pronunciation conflicts with the regularized nonlexical

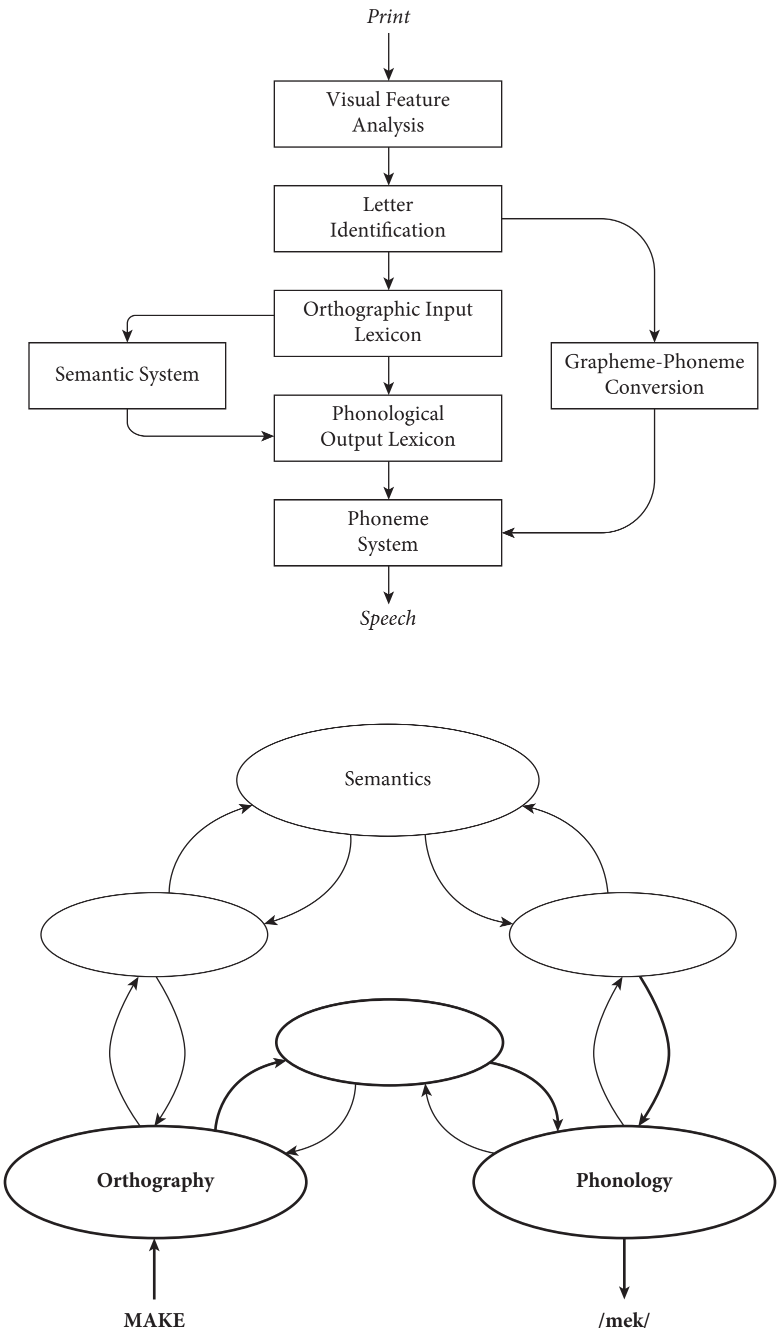


Fig. 11.1 The architecture of the dual-route cascaded model (top), redrawn from Coltheart (2006), and of the connectionist triangle model (bottom), from Woollams et al. (2007).

pronunciation at the phoneme level (Baron & Strawson, 1976). This conflict takes time to resolve, and if it fails, a regularization error will result. This conflict is more pronounced the lower the frequency of a word because lexical output for these words arrives later at the phoneme level and overlaps more with that of the nonlexical pathway (Seidenberg, Waters, Barnes, & Tanenhaus, 1984). Nonwords are processed more slowly than words because the correct nonlexical pronunciation conflicts with the incorrect word pronunciation offered by the direct lexical route (Forster & Chambers, 1973). In contrast to both exception words and nonwords, there is not any conflict between the outputs of the two pathways for regular words; hence they are processed relatively rapidly and accurately.

When it comes to simulation of the highly selective deficits in exception and nonword reading seen in pure cases of acquired surface and phonological dyslexia, respectively, this "is almost *too* simple" (Coltheart, 2006, p. 102), as the model's structure is founded on this double dissociation. Abolition of any component of the direct lexical pathway will produce severe and pure surface dyslexia, whereas damage to any aspect of the lexical pathways will produce severe and pure phonological dyslexia. What is more challenging for the DRC model is to capture the more common graded impairments seen in most cases of acquired dyslexia. The deficits that are seen for exception words or nonwords stand in comparison to better performance for regular words in all cases. However, in most cases, performance for regular words still remains below normal levels. While previous simulations within this framework have managed to capture this more graded pattern of impairments, to do so has required damaging multiple components of the model (Coltheart, Tree, & Saunders, 2010; Nickels, Biedermann, Coltheart, Saunders, & Tree, 2008).

Connectionist Triangle Model

A very different approach to English spelling-sound mappings was adopted by Seidenberg and McClelland (1989), who introduced a model of reading aloud that was composed of distributed representations of orthography and phonology in which processing was determined by the weights on connections between units. Distributed representations are those in which each word corresponds to a particular pattern of activation over a fixed number of units. This scheme offers a parsimonious and efficient approach because each unit can participate in the representation of multiple words. This contrasts with the localist approach of the DRC model, in which the number of lexical units required for each modality corresponds to the number of known words. In connectionist models, the weights on the connections between units start out as random and are gradually refined through exposure to a training corpus. Such learned processing parameters again contrast with the DRC approach, in which these are hardwired.

The connectionist triangle framework is shown in the bottom portion of Figure 11.1. The initial implementation (Seidenberg & McClelland, 1989) contained only a direct subword pathway between spelling and sound, and the goal was to show that a single process could support correct reading of exception words and nonwords. Refinements to the form of the distributed orthographic and phonological units (Plaut, McClelland, Seidenberg, & Patterson, 1996) rendered the model as accurate as skilled adults on both exception word and nonword reading, thereby challenging the long-held belief that separate processes were required to deal with each word type. Frequency affects the performance of this model because exposure to the training corpus was frequency-indexed. Exception words are processed less efficiently as the weights on the mappings for their components are weaker, particularly for the mappings in low-frequency words. Occasionally the stronger mappings for a component may win out, and the result would be a regularization error. Known words enjoy a processing advantage because, over training, the familiar patterns of activation across the phonological units of the model come to settle more rapidly. Nonwords are processed more slowly because their pattern of activation across the phonological units is unfamiliar. Indeed, if the initial pattern of activation for a nonword is too similar to a known word, it will be captured by the familiar pattern, and a lexicalization error would result.

Although this version of the connectionist triangle model could capture many of the key facets of normal reading behavior without recourse to separate procedures for exception words and nonwords, could it simulate the key cases of surface and phonological dyslexia? Even though lesioning various different types of connections within the model was able to provide a good fit for the milder exception word impairment seen in one surface dyslexic patient, MP (Bub, Cancelliere, & Kertesz, 1985), no pattern of lesioning was able to reproduce the severe exception word impairment seen in another patient, KT (McCarthy & Warrington, 1986), without also undermining performance for regular words and nonwords.

Noting that the vast majority of reported cases of surface dyslexia also had semantic impairments, Plaut et al. (1996) suggested that exception word reading also received some support from the whole-word information carried by word meaning. If so, then semantic damage would produce selective deficits of the sort seen in severe cases of surface dyslexia. In support of this view, Plaut et al. provided a simulation in which the phonological representations of the model were trained with extra frequency-weighted activation to approximate the influence of the semantic pathway. This caused a graded division of labor to emerge within the model such that the external activation was particularly important for uncommon words with atypical spelling-to-sound mappings (i.e., low-frequency exception words). When this input was decreased, mild damage captured the pattern of MP and severe damage captured the pattern of KT. A fuller implementation of the semantic pathway has since been provided by Harm and Seidenberg (2004). This extended model is capable of reproducing the finding that imageability, a dimension of word meaning, has a larger effect on exception word reading than on regular word reading in normal participants (Strain, Patterson, & Seidenberg, 1995). It was, therefore, neuropsychological evidence concerning surface dyslexia that led to the full instantiation of the connectionist triangle model.

Given the co-occurrence of phonological dyslexia with spoken language processing deficits, Plaut et al. (1996) assumed that lesioning of phonological representations would result in a lexicality effect. This is because words would be protected against the impact of a phonological lesion as their phonological representations are familiar and would be activated via semantics, but nonwords would suffer as they are unfamiliar and do not have meaning. Due to the divergence between current models in terms of their structure and function, quantitative simulation of data from patients with acquired dyslexia has proven important in model evaluation and continues to teach us a great deal about the structure and function of the normal reading system.

Associations

The focus in the triangle model on functional, as opposed to structural, specialization has meant that it is relatively simple in terms of its component modules. In accounting for surface and phonological dyslexia, the triangle model therefore nominates either general semantic or phonological processing deficits as the underlying cause. This perspective sits well with the original conception of the central acquired dyslexias as arising from deficits in spoken language processing (Shallice & Warrington, 1980) and with the more recent primary systems perspective on acquired reading disorders, which proposes that these are caused by damage to more basic cognitive and neural systems (vision, semantics, phonology; Patterson & Lambon Ralph, 1999).

The connectionist triangle model approach to acquired dyslexia predicts associations between spoken language impairments and reading deficits. Within traditional cognitive neuropsychology, associations in performance across tasks have generally been distrusted. This is because such associations may arise as a consequence of the anatomical contiguity of functionally independent neural regions. Hence the mere co-occurrence of semantic deficits and surface dyslexia or phonological deficits and phonological dyslexia may simply indicate the proximity of these regions within the brain, rather than a causal link between language and reading difficulties.

If we consider individual performance on language tasks and reading tasks over multiple cases (an approach known as a *case series*) and find a significant quantitative relationship between the two deficits, then this provides stronger evidence that problems in the two domains are meaningfully linked. Although it is always possible that both tasks map a common underlying dimension of severity mediated by lesion extent such that the patients with more damage are likely to show multiple deficits, advanced structural and functional neuroimaging of neuropsychological patients can be used to rule out this concern in many cases.

Impairments of the Whole-Word Pathway

A key prediction of the triangle model account of surface dyslexia is that we should see exception word reading deficits among those patients with damage to semantic representations or access. By this view, it is no coincidence that, from the first reports (Marshall & Newcombe, 1973; Warrington, 1975), the vast majority of surface dyslexic cases have some form of semantic deficit. The most common disorder accompanying surface dyslexia is semantic dementia (Patterson & Hodges, 1992), in which progressive atrophy to the anterior temporal lobes undermines semantic processing (Acosta-Cabronero et al., 2011; Mion et al., 2010). While this impairment may at first be most apparent in the linguistic domain, deficits emerge over multiple modalities on tasks such as picture association, sound-picture matching, and odor identification as the disease progresses (Adlam et al., 2006; Bozeat, Lambon Ralph, Patterson, Garrard, & Hodges, 2000; Luzzi et al., 2007).

In the largest assessment of semantic dementia patients' reading to date, Woollams et al. (2007) demonstrated ultimately universal prevalence of surface dyslexia in 51 cases of semantic dementia. Moreover, they demonstrated a quantitative relationship between the degree of the semantic deficit (as measured by performance on nonreading tests involving matching of spoken words to pictures and naming of pictures) and reading performance that was particularly pronounced for low-frequency exception words, with over 50% of the variance in reading accounted for by semantic score. While three cases demonstrated initially intact reading of exception words, all three progressed into a surface dyslexic reading pattern. These results were accurately simulated using a modified version of the Plaut et al. (1996) connectionist triangle model in which the contribution of semantic activation to phonology was reduced and disrupted in order to simulate the effect of semantic dementia on reading.

Although there is no denying the very strong association between semantic deficits and exception word reading seen in semantic dementia, there have been a few cases with normal exception word reading accuracy. WLP, early in her illness, showed no evidence of deficits in reading aloud high-frequency exception words, although these did emerge as the disease progressed (Schwartz, Marin, & Saffran, 1979; Schwartz, Saffran, & Marin, 1980). Two other cases of dissociation have been reported in English (Blazely, Coltheart, & Casey, 2005; Cipolotti & Warrington, 1995), although these have never been followed up. To date, the converse pattern of impaired exception word reading in the face of apparently intact semantic processing has been reported only once (Weekes & Coltheart, 1996).

The occurrence of such dissociations has led some to dismiss the observed large-scale association as arising from two independent deficits due to left temporal lobe atrophy encompassing functionally distinct regions: a semantic deficit associated with the anterior temporal lobe damage and orthographic deficits associated with damage to the occipitotemporal region, which can be damaged in later stage semantic dementia (Rohrer et al., 2009). This region has been dubbed the visual word form area, and although activation in this region does not differ according to whether a letter string forms a word or a nonword (Dehaene, Le Clec'H, Poline, Le Bihan, & Cohen, 2002), it has been nominated by some as the neural location of the orthographic input lexicon (Noble, Glosser, & Grossman, 2000). Coltheart et al. (2010) have simulated the reading data reported by Woollams et al. (2007) by removing various proportions of the lowest frequency words from the orthographic input lexicon and various proportions of grapheme-phoneme rules from the nonlexical route, using an automated parameter search to derive the level of damage that most accurately reproduces the reading deficit for each observation. By this account then, it is damage to reading-specific elements that is responsible for surface dyslexia.

However, a number of objections may be raised to the DRC approach to the simulation of surface dyslexia. First, it leaves a key feature of the patients' performance, the semantic deficit, unexplained. Nor can the DRC view account for the reports from a number of elegant studies (Funnell, 1996; Graham, Hodges, & Patterson, 1994; McKay, Castles, Davis, & Savage, 2007) that the association between semantic knowledge and exception word reading holds at the item level (e.g., if a patient cannot identify a picture of a brooch, then he or she is also likely to mispronounce the written word, while if the same person can identify a picture of a hearth, then that written word will be read correctly). The DRC view does not address the observation that semantic dementia patients do not just fail at low-frequency atypical items in reading aloud, but show the same pattern on a diverse range of tasks including spelling, past-tense inflection, lexical decision, object decision, and delayed copy drawing (Patterson et al., 2006). These observations suggest a pivotal role for semantic knowledge not only in supporting correct reading of words with atypical spelling-sound mappings, but also more generally in processing items that are atypical in their respective domains, such as past-tense irregular verbs like *run,* orthographically unusual words like *yacht,* and visually distinctive animals like a swan (Patterson, 2007).

Evidence From Imaging

Although the common locus for the semantic and reading deficits as proposed by the connectionist triangle model has the benefit of parsimony, there is more direct evidence available from neuroimaging to argue against a reading-specific account implicating damage to the left ventral occipitotemporal cortex as the critical locus for surface dyslexic reading. Seven of the nine semantic dementia cases considered by Nestor, Fryer, and Hodges (2006) had their reading assessed before being scanned, and all were surface dyslexic. The scans for this group showed the expected reduction in blood flow to the bilateral anterior temporal lobes, and while this reduction was more extensive in the left hemisphere than in the right hemisphere, it did not encompass the left ventral occipitotemporal cortex. Hence disruption of the left ventral occipitotemporal cortex could not be the cause of the reading deficits in these patients.

The localization of exception word processing to the left anterior temporal lobe rather than the left ventral occipitotemporal cortex is also supported by functional imaging of semantic patients and healthy participants. A surface dyslexic semantic dementia patient reported by Wilson et al. (2012) showed left anterior temporal atrophy that did not encompass left ventral occipitotemporal cortex, in line with the findings reported by Nestor et al. (2006). Functional imaging of sixteen healthy normal subjects revealed an area of the left anterior temporal lobe that was more active during reading of exception words than nonwords and that overlapped with the patient's atrophy. Functional imaging of a group of six semantic dementia patients (Wilson et al., 2009) and nine healthy participants also revealed activation of the intraparietal sulcus during reading of nonwords in both groups, but activation of the same area for exception words only in the patients, and particularly so when the reading response was a regularization error. This neural region has been implicated as the direct subword pathway in meta-analytic studies of neuroimaging of reading (Cattinelli, Borghese, Gallucci, & Paulesu, 2013; Taylor, Rastle, & Davis, 2013). Hence, this result provides a neural correlate of the increased reliance on the subword pathway that is presumed to occur in surface dyslexia in cognitive models.

Individual Differences

While the connectionist triangle model's account of surface dyslexia in terms of underlying semantic deficits fits with the available neuroimaging evidence, it still needs to account for the occurrence of dissociations between reading and meaning, and more generally, the range of reading performance seen across individuals with a given level of semantic impairment. While the computational simulations of Coltheart et al. (2010) were effective in reproducing this variation, this was on the basis of selecting the particular lesion parameters that most accurately fit each data point. In this sense, the focus was on reproduction of the variance in the data rather than explanation of underlying causes.

The connectionist triangle model simulations of surface dyslexia (Woollams, Lambon Ralph, Plaut, & Patterson, 2010; Woollams et al., 2007) were based on the hypothesis that there are individual differences in the extent to which healthy adult readers rely on semantic information for correct reading of exception words, which was based on evidence of variations in the size of the imageability effect among normal readers (Strain & Herdman, 1999; see also Andrews, this volume). Woollams et al. implemented this notion by creating a population of instantiations of the model that varied in the degree of semantic support provided while learning to read, which was approximated with frequency-indexed whole-word activation of phonology. The models were then lesioned through the reduction and disruption of this semantic support. This resulted in considerable variance in the degree of reading impairment according to the premorbid state of the model (i.e., before the lesion), such that those models trained with weak semantic support are less impaired by a given level of lesion than those trained with strong semantic support. This approach to simulation resulted in a close fit to the patient data, with the model data accounting for more than 90% of the variance in reading performance. Within this view, the cases of single dissociation consisting of intact reading with impaired semantics come from individuals with weak premorbid semantic reliance for exception word reading. The cases of single dissociation consisting of impaired reading and intact semantics come from individuals with strong semantic reliance (e.g., Mendez, 2002; Weekes & Coltheart, 1996). While plausible, this proposal concerning individual differences in the premorbid state of the reading system requires further investigation via behavioral and neuroimaging investigations of normal participants.

Impairments of the Subword Pathway

Turning to the deficits in nonword reading that characterize phonological dyslexia, the dual-route account has focused on the existence of dissociations between impaired nonword reading and intact processing of spoken nonwords, while the connectionist triangle model view has focused on associations between performance in tasks involving written and spoken nonwords. In one of the first reports of phonological dyslexia (Beauvois & Derouesne, 1979), a patient with good spoken language function showed a dissociation between impaired reading of nonwords and intact immediate nonword repetition. However, a later case series of six patients with the poor spoken language that characterizes nonfluent aphasia (Patterson & Marcel, 1992) showed an association between impaired nonword reading and poor performance on phonological tests. Performance was particularly poor on phonological tests that involved nonwords and that were more challenging (e.g., involving phonemic manipulation as opposed to simple repetition).

A larger study of twelve cases presenting with enhanced lexicality effects in reading (Crisp & Lambon Ralph, 2006) involved patients with varying degrees of nonfluent aphasia. This study showed that the strong lexicality and imageability effects seen in reading for these patients also emerged in spoken repetition when this task was made more demanding by the insertion of a 5-second delay in which patients repeated their own names. Moreover, there was specifically a significant correlation between nonword reading ability and performance on the more challenging phonological tasks of phoneme segmentation and blending, with simple repetition being relatively insensitive to variables that affected reading. A similar quantitative association between spoken language processing and reading aloud performance has been reported by Rapcsak et al. (2009) among patients selected for damage to a region of the left perisylvian region that is associated with phonological processing in normal participants. In this study, both nonword reading and nonword spelling were associated with a phonological composite measure consisting of the average of performance across word and nonword repetition, rhyme production and judgment, and phoneme segmentation, blending, and substitution (although it should be noted that the same measure also correlated with word reading and spelling).

It was assumed in the triangle model formulation of Plaut et al. (1996) that lesioning of phonological representations would result in a lexicality effect, as this damage would be offset for words by the additional support provided by activation over the whole-word semantic pathway. Simulations with the version of the model that successfully simulated surface dyslexia in semantic dementia reproduced the profile of mild phonological dyslexia (Welbourne & Lambon Ralph, 2007). However, the structure of that version of the model meant that the lesions were to aspects of the direct pathway (in the form of the input links to hidden units and adding noise to hidden unit outputs) rather than to phonological representations themselves. In a fuller implementation of the triangle model framework, in which it was possible to disrupt phonological processing (by lesioning of all incoming connections via both the whole-word semantic and subword pathways), the lexicality (and also imageability) effects that characterize phonological dyslexia did emerge after a period of retraining in the model, akin to recovery after brain damage (Welbourne, Woollams, Crisp, & Ralph, 2011). Phonological dyslexia has therefore been simulated within the connectionist triangle framework both as a consequence of damage to the reading-specific aspects of the direct subword pathways, and also to the primary language system of phonological processing.

While the case-series evidence clearly suggests that there is a relationship between deficits in nonword reading and phonological processing (see also Tree, 2008, for a review of single cases), this does not necessarily imply causation, as the association may arise as a consequence of the proximity of functionally independent brain regions that support nonword reading and phonological processing. Evidence for this view consists of a number of cases of impaired nonword reading with intact phonological processing, but the issue concerning the relative sensitivity of different kinds of phonological tests becomes very important in the interpretation of such dissociations. Reports of impaired nonword reading with intact phonological processing in a patient with Alzheimer's disease (Caccappolo-van Vliet, Miozzo, & Stern, 2004a) are difficult to interpret. Although nonword repetition was intact and performance was good on rhyme-level tasks that used words as stimuli, the more difficult and sensitive tests like phoneme segmentation and blending could not be performed by the patient. A further case of mild to moderate phonological dyslexia reported by Tree and Kay (2006) also showed mild impairment on the more demanding phoneme level tests (Castles, Holmes, Neath, & Kinoshita, 2003; Patterson & Marcel, 1992). There are, however, other cases where the dissociation appears stronger: A further two Alzheimer's cases showed impaired nonword reading but good phonological processing for words and nonwords across a variety of tests, although it is worth noting that phoneme deletion was not included and each patient failed at least one of the phonological tests that were given (Caccappolo-van Vliet, Miozzo, & Stern, 2004b).

Phonological dyslexia both with and without associated phonological impairment has been simulated within the DRC model by Nickels, Biedermann, Coltheart, Saunders, and Tree (2008). Increasing how long it took the model to start to process each new letter of the string via the nonlexical route captured performance well in one patient with unimpaired word reading and intact repetition performance, but additional noise to the output of the phonological lexicon was needed to capture the slightly impaired word reading seen in another patient with intact repetition performance. The third case of nonword reading deficits with impaired repetition was captured using the addition of noise to the phonemic output level shared by both pathways (similar to the idea of a phonological lesion in the triangle model), but performance for words in the model remained too high. A different approach was adopted by Coltheart et al. (2010), who used combinations of proportional deletions of orthographic input entries and nonlexical rules to simulate the case series data of Crisp and Lambon Ralph (2006). While the fit of these simulations in terms of the size of the lexicality effect is essentially perfect, the values for each case were selected from 40,200 possible combinations. Hence, the emphasis is once again on precise reproduction of the data rather than parsimonious explanation of associated deficits.

Across both the connectionist triangle model and the DRC model, it seems that there is potential for simulation of the rare cases of phonological dyslexia that are relatively reading specific, and also for simulation of the more common pattern of association between impaired nonword reading and phonological processing. Test sensitivity is always an issue in assessing the adequacy of reports of cross-task dissociation. Also, the speed of phonological processing ability has not been formally measured in patients with phonological dyslexia. Both of these factors can affect the strength of the relationship between performance on measures of phonological processing and nonword reading. Nevertheless, there does appear to be variation in the degree of phonological processing deficit observed in phonological dyslexia, suggesting some underlying heterogeneity in terms of cause.

Evidence From Imaging

The heterogeneity seen across cases of phonological dyslexia in terms of the degree of correspondence between the reading and language deficit is consistent with the results of meta-analyses of normal neuroimaging studies that implicate a large left perisylvian language network in phonological processing (Vigneau et al., 2006). In the case series of thirty-one participants with left perisylvian lesions reported by Rapcsak et al. (2009), approximately 80% showed an abnormally large disadvantage for nonword relative to word reading. The presence of an increased lexicality effect was not associated specifically with the presence of damage to any of five perisylvian regions of interest (posterior inferior frontal gyrus/Broca's area [BA44/45], precentral gyrus [BA4/6], insula, superior temporal gyrus/Wernicke's area [BA22], and supramarginal gyrus [BA40]). Rather, damage to any of these regions was sufficient to produce a profile of phonological dyslexia. The more of these regions that were damaged, the stronger the impairments were to both word and nonword reading and spelling.

The regions proposed to support direct subword mapping during reading (Cattinelli et al., 2013; Taylor et al., 2013) overlap with only some parts of the larger left perisylvian network implicated in phonological processing. This means that variation in the relative strength of nonword reading and phonological processing deficits according to lesion location is to be expected. Indeed, Rapcsak and colleagues (2009) found that the size of the lexicality effect in reading was correlated with the presence of lesions in the inferior frontal gyrus and precentral gyrus, suggesting a particular role for this region in processing both written and spoken nonwords. In contrast, posterior superior temporal regions would seem more involved in processing input in phonological tasks, as shown by their consistent activation during speech perception, which is less relevant for reading (Hickok & Poeppel, 2007).

This proposal that there is a graded task specialization across the left perisylvian phonological processing network is highly speculative, but it does have the potential to account for the variance in the strength of association between written and spoken language deficits across cases of phonological dyslexia. What will be required to validate this proposal are studies where the structural and functional abnormalities of a large number of patients with nonword reading deficits are mapped onto a range of language processing tasks. This would need to be complemented by more detailed considerations of parallel tasks like reading and repetition in functional imaging studies of normal participants (e.g., Hope et al., 2014) that would allow a more complete understanding of functional subdivisions within the left perisylvian phonological processing network.

Recovery and Relearning

One of the most interesting insights to emerge from the computational modeling of phonological dyslexia is the possibility that reading deficits may change over the course of time after brain damage. In the simulations of Welbourne et al. (2011), lesions to the phonological system alone were unable to produce a pattern of phonological dyslexia, as the impairments to nonword reading were accompanied by impairments to word reading. This meant that the lexicality effects produced by the model were weaker than those reported in the literature.

Crisp and Lambon Ralph (2006) noted that the vast majority of cases of phonological dyslexia have been observed in patients with chronic stroke aphasia. Hence it is possible that the advantage for word processing only emerges after a period of recovery because patients will continue to encounter words in their day-to-day environment.

When Welbourne et al. (2011) continued to train their damaged model on words to approximate the recovery process, the lexicality effect that defines phonological dyslexia emerged. Some evidence for this notion of recovery as key in phonological dyslexia is provided by consideration of reading performance in progressive nonfluent aphasia, where there is a continued deterioration of phonological processing over time associated with left perisylvian damage (Nestor et al., 2003). Woollams and Patterson (2012) reported that progressive nonfluent aphasic patients, who presumably experience limited opportunity for relearning in the face of continuing neurodegeneration, showed deficits for both nonwords and low-frequency exception words, a pattern seen in the connectionist triangle model after phonological lesions without relearning (Plaut et al., 1996).

Interestingly, as the lexicality effect emerged in Welbourne et al.'s (2011) model with retraining on words, so too did the characteristic imageability effect, indicating that the recovery of word reading was supported by an increased reliance on the semantic pathway. This is consistent with longitudinal evidence (Read, Welbourne, Sage, & Lambon Ralph, 2010) that semantic processing capacity three months after a stroke is a significant predictor of whether phonological dyslexia will emerge when the patient is tested again nine months after a stroke. Essentially, it is those patients with a more intact semantic pathway who will show the improvement in word reading that produces the hallmark lexicality effect. More longitudinal behavioral and neuroimaging work on patients with left perisylvian lesions is needed to understand the role of recovery in nonword reading and phonological processing, but this simulation work demonstrates the exciting potential of connectionist modeling to capture recovery via its emphasis on learning.

Interaction of the Whole-Word and Subword Pathways

In considering the simulation of surface and phonological dyslexia within the dual-route and connectionist triangle models, we see that the two approaches vary in the extent to which they posit interaction between the whole-word and subword pathways. Although the output of both the direct lexical and nonlexical routes is pooled at the phoneme level in the DRC model, the interplay between the semantic and direct pathways of the connectionist triangle model is far greater, as this occurs during throughout learning (Harm & Seidenberg, 2004). In this way, we see complimentary partial functional specialization over the two pathways according to the demands of the particular word being processed. For example, semantic activation compensates for inefficient processing via the direct pathway for words with exceptional spelling-sound mappings, particularly those that are low in frequency. Support for this view is provided by the stronger imageability effects seen for these items in normal reading (Strain et al., 1995; Strain, Patterson, & Seidenberg, 2002) and the association between surface dyslexia and semantic deficits. Similarly, phonological activation compensates for inefficient processing via the semantic pathway for words with low imageability meaning. Support for this view is provided by the stronger effects of phonological variables for low imageability items (Tyler, Voice, & Moss, 2000; Westbury & Moroschan, 2009) and the marked impairment for low imageability words in many cases of phonological dyslexia.

Within the connectionist triangle model, the differential contribution of the semantic and direct pathways to processing different kinds of words corresponds to item-based variation in the division of labor. As noted in accounting for variation in surface dyslexia, the division of labor can vary over different individuals, with some relying more on semantic information for correct exception word reading than others. And, as covered in consideration of lexicality effects in the simulation of phonological dyslexia, the division of labor can be adjusted over time in response to brain damage. The division of labor within the connectionist triangle model is therefore dynamic, varying over different items, different individuals, and over time.

Deep Dyslexia

This final type of central dyslexia shares with phonological dyslexia a profile of impaired nonword reading (albeit more severely so) and prominent imageability effects in word reading. The cardinal feature of deep dyslexia is the occurrence of semantic paralexias in reading aloud, such that a word is read as another word that is related semantically but unrelated in form, such as "blood" for *heart* (Marshall & Newcombe, 1973). The combination of reading deficits both for nonwords and for real words implies that there is disruption to both the subword and whole-word semantic pathways in deep dyslexia. This is consistent with the observation that deep dyslexia results from large left perisylvian lesions. It has therefore been suggested that the reading performance of such patients reflects right hemisphere reading mechanisms (Weekes, Coltheart, & Gordon, 1997), although at least some cases show no more right-hemisphere activation than healthy readers (Price et al., 1998).

The cause of deep dyslexia remains unclear, with some researchers proposing a single phonological deficit and others suggesting an additional semantic deficit. It is certainly the case that these patients have marked phonological processing deficits that accompany their severe nonword reading deficits (Jefferies, Sage, & Lambon Ralph, 2007). By one view (Wilshire & Fisher, 2004), the semantic errors seen in deep dyslexic reading reflect the operation of the normal semantic pathway without any constraint from subword phonological processing. Although deep dyslexic patients do not always show semantic errors in spoken word repetition, these do emerge when the task is made more demanding by the insertion of a filled delay (Jefferies et al., 2007). Further, Welbourne et al. (2011) found that a severe phonological lesion within an implementation of the connectionist triangle model led to above-chance rates of semantic errors in reading after recovery. Taken together, these findings suggest that phonological impairment alone could be sufficient to produce semantic paralexias.

Nevertheless, additional semantic deficits may be present in deep dyslexia. Riley and Thompson (2010) considered performance of nine phonological dyslexics and five deep dyslexics in a category verification task, presented in both written and spoken form, for words that were either typical or atypical within their semantic category. Like healthy controls, the phonological dyslexics showed a significant advantage for typical over atypical items. However, this was abolished in the deep dyslexic patients in both modalities, suggesting a semantic impairment. There is also evidence that these semantic problems are linked to reading deficits. Crisp and Lambon Ralph (2006) found that the incidence of semantic errors in their phonologically impaired patients' reading was correlated with semantic processing ability as measured by a demanding synonym judgment test. Conversely, in a large case series of Italian aphasic patients who made semantic errors in picture naming, Ciaghi, Pancheri, and Miceli (2010) found that those who produced semantic paralexias in reading were distinguished by their more severe nonword reading problems. These reports of both phonological and semantic deficits in deep dyslexia align with the connectionist model proposed by Plaut and Shallice (1993), in which the subword pathway was assumed to be completely abolished, and semantic paralexias (e.g., *flan* read as "tart") resulted from lesions to the connections and representations of the semantic whole-word pathway.

Further case-series patient research with detailed neuroimaging is needed to understand the causes of deep dyslexia. It is possible that there is heterogeneity in the underlying causes that corresponds to variation in lesion site and extent within the left perisylvian region. It may be that some cases do reflect the impact of severely impaired subword reading mechanisms interacting with an intact whole-word semantic pathway, while in other cases with more extensive damage, behavior is a reflection of damage to both mechanisms. In either case, in computational modeling terms, the interaction of the two pathways will prove crucial in formulating accounts of deep dyslexic performance.

Conclusions and Future Directions

This chapter has provided a brief overview of the peripheral and central acquired dyslexias. Emphasis has been placed on surface and phonological dyslexia, as this double dissociation between exception and nonword reading has been at the heart of formulation of theories and more recently computational models of reading aloud. Both dual-route and connectionist models have converged on a system that involves whole-word and subword processing mechanisms. Behavioral and imaging evidence concerning exception word reading deficits has provided strong evidence for a semantic whole-word pathway. Consideration of nonword reading deficits has emphasized the role of general phonological processing. What acquired dyslexia has told us about reading in the mind and brain is that this ability is underpinned by more basic perceptual and cognitive processes (Patterson & Lambon Ralph, 1999; Woollams, 2014). This framework could motivate future investigations concerning how variation in these more basic perceptual and cognitive processes relates to performance among normal readers and those with developmental reading disorders (see Pennington & Peterson, this volume).

References

Acosta-Cabronero, J., Patterson, K., Fryer, T. D., Hodges, J. R., Pengas, G., Williams, G. B., & Nestor, P. J. (2011). Atrophy, hypometabolism and white matter abnormalities in semantic dementia tell a coherent story. *Brain, 134,* 2025–2035.

Adlam, A. L. R., Patterson, K., Rogers, T. T., Nestor, P. J., Salmond, C. H., Acosta-Cabronero, J., & Hodges, J. R. (2006). Semantic dementia and fluent primary progressive aphasia: Two sides of the same coin? *Brain, 129,* 3066–3080.

Anderson, B. (1999). A computational model of neglect dyslexia. *Cortex, 35,* 201–218.

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 386–393.

Beauvois, M. F., & Derouesne, J. (1979). Phonological alexia: Three dissociations. *Journal of Neurology Neurosurgery and Psychiatry, 42,* 1115–1124.

Behrmann, M., Nelson, J., & Sekuler, E. B. (1998). Visual complexity in letter-by-letter reading: "Pure" alexia is not pure. *Neuropsychologia, 36,* 1115–1132.

Blazely, A. M., Coltheart, M., & Casey, B. J. (2005). Semantic impairment with and without surface dyslexia: Implications for models of reading. *Cognitive Neuropsychology, 22,* 695–717.

Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia, 38,* 1207–1215.

Bub, D. N., Cancelliere, A., & Kertesz, A. (1985). Whole-word and analytic translation of spelling to sound in a nonsemantic reader. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 15–34). Hillsdale, NJ: Erlbaum.

Caccappolo-van Vliet, E., Miozzo, M., & Stern, Y. (2004a). Phonological dyslexia without phonological impairment? *Cognitive Neuropsychology, 21,* 820–839.

Caccappolo-van Vliet, E., Miozzo, M., & Stern, Y. (2004b). Phonological dyslexia: A test case for reading models. *Psychological Science, 15,* 583–590.

Castles, A., Holmes, V. M., Neath, J., & Kinoshita, S. (2003). How does orthographic knowledge influence performance on phonological awareness tasks? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 56A,* 445–467.

Cattinelli, I., Borghese, N. A., Gallucci, M., & Paulesu, E. (2013). Reading the reading brain: A new meta-analysis of functional imaging data on reading. *Journal of Neurolinguistics, 26,* 214–238.

Ciaghi, M., Pancheri, E., & Miceli, G. (2010). Semantic paralexias: A group-case study on the underlying functional mechanisms, incidence and clinical features in a consecutive series of 340 Italian aphasics. *Brain and Language, 115,* 121–132.

Cipolotti, L., & Warrington, E. K. (1995). Semantic memory and reading abilities: A case report. *Journal of the International Neuropsychological Society, 1,* 104–110.

Coltheart, M. (2006). Acquired dyslexias and the computational modelling of reading. *Cognitive Neuropsychology, 23,* 96–109.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Coltheart, M., Tree, J. J., & Saunders, S. J. (2010). Computational modeling of reading in semantic dementia: Comment on Woollams, Lambon Ralph, Plaut, and Patterson (2007). *Psychological Review, 117,* 256–271.

Crisp, J., & Lambon Ralph, M. A. (2006). Unlocking the nature of the phonological-deep dyslexia continuum: The keys to reading aloud are in phonology and semantics. *Journal of Cognitive Neuroscience, 18,* 348–362.

Cumming, T. B., Patterson, K., Verfaellie, M. M., & Graham, K. S. (2006). One bird with two stones: Abnormal word length effects in pure alexia and semantic dementia. *Cognitive Neuropsychology, 23,* 1130–1161.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences, 15,* 254–262.

Dehaene, S., Le Clec'H, G., Poline, J. B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: A prelexical representation of visual words in the fusiform gyrus. *NeuroReport, 13,* 321–325.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12,* 627–635.

Funnell, E. (1983). Phonological processes in reading: New evidence from acquired dyslexia. *British Journal of Psychology, 74,* 159–180.

Funnell, E. (1996). Response biases in oral reading: An account of the co-occurrence of surface dyslexia and semantic dementia. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49A,* 417–446.

Gomori, A. J., & Hawryluk, G. A. (1984). Visual agnosia without alexia. *Neurology, 34,* 947–950.

Graham, K. S., Hodges, J. R., & Patterson, K. (1994). The relationship between comprehension and oral reading in progressive fluent aphasia. *Neuropsychologia, 32,* 299–316.

Hall, D. A., Humphreys, G. W., & Cooper, A. C. G. (2001). Neuropsychological evidence for case-specific reading: Multi-letter units in visual word recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 54A,* 439–467.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111,* 662–720.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8,* 393–402.

Hope, T. M. H., Prejawa, S., Parker Jones, O., Oberhuber, M., Seghier, M. L., Green, D. W., & Price, C. J. (2014). Dissecting the functional anatomy of auditory word repetition. *Frontiers in Human Neuroscience, 8,* 246.

Jefferies, E., Lambon Ralph, M. A., Jones, R., Bateman, D., & Patterson, K. (2004). Surface dyslexia in semantic dementia: A comparison of the influence of consistency and regularity. *Neurocase, 10,* 290–299.

Jefferies, E., Sage, K., & Lambon Ralph, M. A. (2007). Do deep dyslexia, dysphasia and dysgraphia share a common phonological impairment? *Neuropsychologia, 45,* 1553–1570.

Leff, A. P., Spitsyna, G., Plant, G. T., & Wise, R. J. S. (2006). Structural anatomy of pure and hemianopic alexia. *Journal of Neurology, Neurosurgery, and Psychiatry, 77,* 1004–1007.

Luzzi, S., Snowden, J. S., Neary, D., Coccia, M., Provinciali, L., & Lambon Ralph, M. A. (2007). Distinct patterns of olfactory impairment in Alzheimer's disease, semantic dementia, frontotemporal dementia, and corticobasal degeneration. *Neuropsychologia, 45,* 1823–1831.

Marshall, J. C., & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research, 2,* 175–199.

Marshall, J. C., & Newcombe, F. (1981). Lexical access: A perspective from pathology. *Cognition, 10,* 209–214.

McCarthy, R. A., & Warrington, E. K. (1986). Phonological reading: Phenomena and paradoxes. *Cortex, 22,* 359–380.

McGlinchey-Berroth, R., Bullis, D. P., Milberg, W. P., Verfaellie, M., Alexander, M., & D'Esposito, M. (1996). Assessment of neglect reveals dissociable behavioral but not neuroanatomical subtypes. *Journal of the International Neuropsychological Society, 2,* 441–451.

McKay, A., Castles, A., Davis, C., & Savage, G. (2007). The impact of progressive semantic loss on reading aloud. *Cognitive Neuropsychology, 24,* 162–186.

Mendez, M. F. (2002). Slowly progressive alexia. *Journal of Neuropsychiatry and Clinical Neurosciences, 14,* 84.

Mion, M., Patterson, K., Acosta-Cabronero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y. T., … Nestor, P. J. (2010). What the left and right anterior fusiform gyri tell us about semantic memory. *Brain, 133,* 3256–3268.

Miozzo, M., & Caramazza, A. (1998). Varieties of pure alexia: The case of failure to access graphemic representations. *Cognitive Neuropsychology, 15,* 203–238.

Morton, J. (1980). The logogen model and orthographic structure. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 117–133). London, England: Academic Press.

Mycroft, R. H., Behrmann, M., & Kay, J. (2009). Visuoperceptual deficits in letter-by-letter reading? *Neuropsychologia, 47,* 1733–1744.

Nestor, P. J., Fryer, T. D., & Hodges, J. R. (2006). Declarative memory impairments in Alzheimer's disease and semantic dementia. *NeuroImage, 30,* 1010–1020.

Nestor, P. J., Graham, N. L., Fryer, T. D., Williams, G. B., Patterson, K., & Hodges, J. R. (2003). Progressive non-fluent aphasia is associated with hypometabolism centred on the left anterior insula. *Brain, 126,* 2406–2418.

Nickels, L., Biedermann, B., Coltheart, M., Saunders, S., & Tree, J. J. (2008). Computational modelling of phonological dyslexia: How does the DRC model fare? *Cognitive Neuropsychology, 25,* 165–193.

Noble, K., Glosser, G., & Grossman, M. (2000). Oral reading in dementia. *Brain and Language, 74,* 48–69.

Patterson, K. (2007). The reign of typicality in semantic memory. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362,* 813–821.

Patterson, K., & Hodges, J. R. (1992). Deterioration of word meaning: Implications for reading. *Neuropsychologia, 30,* 1025–1040.

Patterson, K., & Lambon Ralph, M. A. (1999). Selective disorders of reading? *Current Opinion in Neurobiology, 9,* 235–239.

Patterson, K., Lambon Ralph, M. A., Jefferies, E., Woollams, A., Jones, R., Hodges, J. R., & Rogers, T. T. (2006). "Presemantic" cognition in semantic dementia: Six deficits in search of an explanation. *Journal of Cognitive Neuroscience, 18*, 169–183.

Patterson, K., & Marcel, A. J. (1992). Phonological ALEXIA or PHONOLOGICAL alexia? In J. Alegria, D. Holende, J. Junça de Morais, & M. Radeau (Eds.), *Analytic approaches to human cognition* (pp. 259–274). New York, NY: Elsevier.

Patterson, K., & Shewell, C. (1987). Speak and spell: Dissociations and word-class effects. In M. Coltheart, G. Sartori, & R. Job (Eds.), *The cognitive neuropsychology of language* (pp. 273–295). Hillsdale, NJ: Erlbaum.

Pflugshaupt, T., Gutbrod, K., Wurtz, P., Von Wartburg, R., Nyffeler, T., De Haan, B., … Mueri, R. M. (2009). About the role of visual field defects in pure alexia. *Brain, 132,* 1907–1917.

Plaut, D. C. (1995). Double dissociation without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology, 17,* 291–321.

Plaut, D. C. (2003). Interpreting double dissociations in connectionist networks. *Cortex, 39,* 138–141.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology, 10,* 377–500.

Price, C. J., Howard, D., Patterson, K., Warburton, E. A., Friston, K. J., & Frackowiak, R. S. J. (1998). A functional neuroimaging description of two deep dyslexic patients. *Journal of Cognitive Neuroscience, 10,* 303–315.

Ptak, R., Di Pietro, M., & Schnider, A. (2012). The neural correlates of object-centered processing in reading: A lesion study of neglect dyslexia. *Neuropsychologia, 50,* 1142–1150.

Rapcsak, S. Z., Beeson, P. M., Henry, M. L., Leyden, A., Kim, E., Rising, K., … Cho, H. (2009). Phonological dyslexia and dysgraphia: Cognitive mechanisms and neural substrates. *Cortex, 45,* 575–591.

Read, J., Welbourne, S., Sage, K., & Lambon Ralph, M. (2010). Recovery of language and reading in post-CVA aphasia: A longitudinal study. Procedia—Social and Behavioral Sciences, *6,* 158–159.

Riley, E. A., & Thompson, C. K. (2010). Semantic typicality effects in acquired dyslexia: Evidence for semantic impairment in deep dyslexia. *Aphasiology, 24,* 802–813.

Roberts, D. J., Woollams, A. M., Kim, E., Beeson, P. M., Rapcsak, S. Z., & Lambon Ralph, M. A. (2013). Efficient visual object and word recognition relies on high spatial frequency coding in the left posterior fusiform gyrus: Evidence from a case-series of patients with ventral occipito-temporal cortex damage. *Cerebral Cortex, 23,* 2568–2580.

Rohrer, J. D., Warren, J. D., Modat, M., Ridgway, G. R., Douiri, A., Rossor, M. N., … Fox, N. C. (2009). Patterns of cortical thinning in the language variants of frontotemporal lobar degeneration. *Neurology, 72,* 1562–1569.

Schwartz, M. F., Marin, O. S. M., & Saffran, E. M. (1979). Dissociations of language function in dementia: A case study. *Brain and Language, 7,* 277–306.

Schwartz, M. F., Saffran, E. M., & Marin, O. S. M. (1980). Fractionating the reading process in dementia: Evidence for word-specific print-to-sound associations. In M. Coltheart K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 259–269). London, England: Routledge.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23,* 383–404.

Shallice, T., & Warrington, E. K. (1975). Word recognition in a phonemic dyslexic patient. *Quarterly Journal of Experimental Psychology, 27,* 187–199.

Shallice, T., & Warrington, E. K. (1977). The possible role of selective attention in acquired dyslexia. *Neuropsychologia, 15,* 31–41.

Shallice, T., & Warrington, E. K. (1980). Single and multiple component dyslexic syndromes. In K. Patterson, M. Coltheart, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 91–118). London, England: Routledge.

Stanovich, K. E. (1982). Individual differences in the cognitive processes of reading: I. Word decoding. *Journal of Learning Disabilities, 15,* 485–493.

Strain, E., & Herdman, C. M. (1999). Imageability effects in word naming: An individual differences analysis. *Canadian Journal of Experimental Psychology, 53,* 347–359.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 114–154.

Strain, E., Patterson, K., & Seidenberg, M. S. (2002). Theories of word naming interact with spelling-sound consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 207–214.

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin, 139,* 766–791.

Tree, J. J. (2008). Two types of phonological dyslexia—A contemporary review. *Cortex, 44,* 698–706.

Tree, J. J., & Kay, J. (2006). Phonological dyslexia and phonological impairment: An exception to the rule? *Neuropsychologia, 44,* 2861–2873.

Tyler, L. K., Voice, J. K., & Moss, H. E. (2000). The interaction of meaning and sound in spoken word recognition. *Psychonomic Bulletin & Review, 7,* 320–326.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (2001). What do double dissociations prove? *Cognitive Science, 25,* 111–172.

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., … Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage, 30,* 1414–1432.

Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology, 27,* 635–657.

Warrington, E. K., Cipolotti, L., & McNeil, J. (1993). Attentional dyslexia: A single case study. *Neuropsychologia, 31,* 871–885.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50A,* 439–456.

Weekes, B., & Coltheart, M. (1996). Surface dyslexia and surface dysgraphia: Treatment studies and their theoretical implications. *Cognitive Neuropsychology, 13,* 277–315.

Weekes, B. S., Coltheart, M., & Gordon, E. (1997). Deep dyslexia and right hemisphere reading—A regional cerebral blood flow study. *Aphasiology, 11,* 1139–1158.

Welbourne, S. R., & Lambon Ralph, M. A. (2007). Using parallel distributed processing models to simulate phonological dyslexia: The key role of plasticity-related recovery. *Journal of Cognitive Neuroscience, 19,* 1125–1139.

Welbourne, S. R., Woollams, A. M., Crisp, J., & Lambon Ralph, M. A. (2011). The role of plasticity-related functional reorganization in the explanation of central dyslexias. *Cognitive Neuropsychology, 28,* 65–108.

Westbury, C., & Moroschan, G. (2009). Imageability x phonology interactions during lexical access: Effects of modality, phonological neighbourhood, and phonological processing efficiency. *Mental Lexicon, 4,* 115–145.

Wilshire, C. E., & Fisher, C. A. (2004). “Phonological” dysphasia: A cross-modal phonological impairment affecting repetition, production, and comprehension. *Cognitive Neuropsychology, 21,* 187–210.

Wilson, M. A., Joubert, S., Ferré, P., Belleville, S., Ansaldo, A. I., Joanette, Y., . . . Brambati, S. M. (2012). The role of the left anterior temporal lobe in exception word reading: Reconciling patient and neuroimaging findings. *NeuroImage, 60,* 2000–2007.

Wilson, S. M., Brambati, S. M., Henry, R. G., Handwerker, D. A., Agosta, F., Miller, B. L., . . . Gorno-Tempini, M. L. (2009). The neural basis of surface dyslexia in semantic dementia. *Brain, 132,* 71–86.

Woodhead, Z. V. J., Wise, R. J. S., Sereno, M., & Leech, R. (2011). Dissociation of sensitivity to spatial frequency in word and face preferential areas of the fusiform gyrus. *Cerebral Cortex, 21,* 2307–2312.

Woollams, A. M. (2014). Connectionist neuropsychology: Uncovering ultimate causes of acquired dyslexia. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*.

Woollams, A. M., Hoffman, P., Roberts, D. J., Lambon Ralph, M. A., & Patterson, K. E. (2014). What lies beneath: A comparison of reading aloud in pure alexia and semantic dementia. *Cognitive Neuropsychology, 31,* 461–481.

Woollams, A. M., Lambon Ralph, M. A., Plaut, D. C., & Patterson, K. (2007). SD-squared: On the association between semantic dementia and surface dyslexia. *Psychological Review, 114,* 316–339.

Woollams, A. M., Lambon Ralph, M. A., Plaut, D. C., & Patterson, K. (2010). SD-squared revisited: Reply to Coltheart, Tree, and Saunders (2010). *Psychological Review, 117,* 273–281.

Woollams, A. M., & Patterson, K. (2012). The consequences of progressive phonological impairment for reading aloud. *Neuropsychologia, 50,* 3469–3477.

CHAPTER

12 Literacy and Literacy Development in Bilinguals

Debra Jared

Abstract

The first part of the chapter reviews the literature on reading in adult bilinguals, and the second part reviews the literature on reading development. The focus of the review is on how knowledge of one language influences reading in the other language. In the literature on adult bilinguals, most research on this issue has been on word recognition, with fewer studies on sentence processing and fewer still on text reading. A model of bilingual word recognition is discussed. In the literature on child bilinguals, the focus has been on predictors of reading development in the second language. Only a few studies have investigated how bilingual children represent and process their two languages. Several theoretical proposals regarding bilingual language development are discussed.

Key Words: bilingual, word recognition, sentence processing, reading development, biliteracy

Many individuals around the world can speak and read in more than one language. Until the later 1990s, however, most psycholinguistic research focused on how individuals represent and process a single language, typically English. Since that time there has been a very rapid increase in the number of studies on bilingual language processing. It has become clear from this research that a bilingual is not "two monolinguals in one head" (Grosjean, 1998). That is, the two languages of a bilingual do not operate in isolation from one another. What is emerging is an intriguing picture of a bilingual individual as one who manages two languages simultaneously. The focus of this chapter is on research that investigates how a bilingual's knowledge of one language influences reading in the other language.

This chapter contains two main sections, the first on adult bilingual reading and the second on biliteracy development in children. For the most part, these have been separate bodies of literature. However, the adult literature could be informed by a consideration of how skilled reading in bilinguals comes about, and the literature on children could benefit from what has been learned in the adult literature about how the bilinguals' two languages are represented and processed. Including these two bodies of literature in a single chapter should encourage cross-fertilization of ideas.

Biliteracy in Adults

Much of the literature on the interaction between an adult bilingual's languages when reading has focused on word recognition, with a smaller body of work on sentence processing and very little on text processing. Complicating the investigation of this issue is that the extent of the interaction between a bilingual's two languages when reading may depend on a variety of factors, including the similarity of the languages, the age at which the second language was acquired, relative proficiency in each language, and the extent to which each language is used in daily life.

Word Recognition

The issue of whether word recognition in one language is influenced by knowledge of another language has typically been addressed by investigating performance on words that have some overlap between the two languages. The logic is that if performance on such words differs from performance on matched control words that exist only in the language of the task (the target language), then representations from the other (nontarget) language were probably activated. There are three major kinds of studies on this issue: those that present words one at a time, priming studies in which a word is briefly presented and then followed by a target to which the participant responds, and studies in which words are embedded in sentences. A related issue is whether a bilingual uses the same brain structures when reading words in each language. Following a review of this research (see also Dijkstra, 2005; Kroll, Gullifer, & Rossi, 2013), a model of bilingual word recognition that attempts to account for these findings is presented.

STUDIES WITH SINGLE WORDS

The most commonly studied cross-language words are cognates. Cognates share orthographic forms, phonological forms, or both in the two languages and also share meanings (e.g., *animal* means the same thing in English and French). Numerous studies have found that cognates are responded to more quickly than matched single-language control words in lexical decision and naming tasks (e.g., Dijkstra, Grainger, & Van Heuven, 1999; Peeters, Dijkstra, & Grainger, 2013; Schwartz, Kroll, & Diaz, 2007; Van Hell & Dijkstra, 2002; see Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010, for a review), providing evidence that bilinguals activate representations in both of their languages when performing a task in one of them. One explanation of the cognate facilitation effect is that the orthographic representation activates a phonological representation in each language, which in turn activates a common semantic representation. The semantic representation sends feedback to the orthographic representation and the phonological representation in each language. The consequence is greater activation of representations for cognates than for single-language control words. Furthermore, Peeters et al. (2013) found that the magnitude of the cognate effect depends on the word's frequency both in the target language and the nontarget language, with cognate effects being especially large for Dutch-English bilinguals in an English lexical decision task when words were low in frequency in English and high in frequency in Dutch. This finding indicates that nontarget language representations have the most impact when they are highly familiar, and therefore activated quickly and strongly, and when the corresponding target word is much less familiar, and therefore activated more slowly and weakly. In addition, Schwartz et al. (2007) showed that Spanish-English cognates were named faster when they had similar phonological representations in the two languages (e.g., *piano*) than when they had more distinct pronunciations (e.g., *base*). Similarly, Dijkstra et al. (2010) observed that Dutch-English identical cognates had faster lexical decision latencies when there was greater phonological overlap between the two languages. These findings suggest more specifically that phonological representations from both of a bilingual's languages were activated.

Interlingual homographs are words that have the same spelling in two languages but different meanings (e.g., *main* means 'hand' in French) and often different pronunciations. Experiments using interlingual homographs in the lexical decision task have produced a complicated pattern of findings (e.g., De Groot, Delmaar, & Lupker, 2000; Dijkstra et al., 1999; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Lemhöfer & Dijkstra, 2004; Von Studnitz & Green, 2002), depending on list composition and task demands. However Dijkstra (2005) concluded that most results provide evidence that the nontarget language influenced lexical decision performance in the target language. The activation of multiple meanings that feed activation back to corresponding orthographic representations in each language can slow decisions when participants have to determine whether the homograph is a word in a particular language. However, activation of multiple meanings can speed decisions when the task is to indicate whether it is a word in either language. Studies using homographs in other behavioral tasks, such as naming (Jared & Szucs, 2002; Smits, Martensen, Dijkstra, & Sandra, 2006) and cross-language semantic priming (e.g., Paulmann, Elston-Güttler, Gunter, & Kotz, 2006), also provide evidence that representations from both languages are activated. Additional support comes from event-related potential (ERP) and functional magnetic resonance imaging (fMRI) studies. Kerkhofs, Dijkstra, Chwilla, and De Bruijn (2006) showed that the frequency of interlingual homographs in both the language of the task and the task-irrelevant language affect participants' behavioral responses and electrophysiological activity. In an fMRI study, Van Heuven, Schriefers, Dijkstra, and Hagoort (2008) showed that the between-language phonological and semantic conflict for interlingual homographs produced greater activation of the left inferior prefrontal cortex than for control words.

Cross-language effects are not limited to words that have the same orthographic form in two languages. A third kind of cross-language word is interlingual homophones, which share a pronunciation across languages (e.g., *shoe* in English and *choux* in French). Dijkstra et al. (1999) observed that interlingual homophones had slower lexical decisions than matched control words, but subsequent research has found that interlingual homophones produce facilitatory effects in lexical decision (Haigh & Jared, 2007; Lemhöfer & Dijkstra, 2004) and in an ERP experiment involving a semantic decision (Carrasco-Ortiz, Midgley, & Frenck-Mestre, 2012). These studies provide evidence for parallel activation of phonological representations in both languages. Other research has shown that these activated phonological representations in turn activate their corresponding meanings. Friesen and Jared (2012) gave bilinguals who knew English and French a category decision task in which the critical items (e.g., *shoe*) were not members of a category (e.g., *vegetable*), but they sounded like a word in the other language that was (e.g., *choux* 'cabbage'). Participants were more likely to make errors on these critical items than on matched single-language control words. Evidence of even more interactivity between a bilingual's languages comes from Wu and Thierry (2010), who found that decisions by Chinese-English bilinguals about the semantic similarity of two English words were influenced by the phonological similarity of their Chinese translations. Semantic representations activated by the presented English words must have fed activation back to phonological representations in the nontarget language.

In many word recognition models (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989), words in the lexicon are activated if they share any letters or sounds with the presented word, and the amount of activation is proportional to the degree of overlap. Words that are similar to a target word have been called its neighbors (see Perea, this volume). Several studies have suggested that the neighborhood of words that is activated upon seeing a word in one language includes similar words in the other language. Grainger and Dijkstra (1992) and Van Heuven, Dijkstra, and Grainger (1998) tested Dutch-English bilinguals on an English lexical decision task and found that decision latencies were longer for English words that had many Dutch orthographic neighbors than for English words with few Dutch neighbors. Similarly, Midgley, Holcomb, Van Heuven, and Grainger (2008) showed an effect of number of cross-language neighbors in an ERP study with French-English bilinguals. Also with French-English bilinguals, Jared and Kroll (2001) observed longer naming latencies for English words such as *bait* that have French neighbors with a different pronunciation (e.g., *fait* and *lait*) than for English words that did do not have French neighbors (e.g., *bump*). These inhibitory effects of French neighbors were not, however, as strong as inhibitory effects of English neighbors on words such as *bead* (*head, dead*), even though participants were more fluent in French. This latter result suggests that between-language connections can be weaker than within-language connections. The studies by Grainger and colleagues provide evidence that orthographic representations of words from both languages are activated, and the Jared and Kroll study provides evidence that phonological representations of words from both languages are activated.

Priming Studies

A priming paradigm has been used to investigate whether phonological representations activated by a prime in one language facilitate reading similarly pronounced targets in the other language. Brysbaert and colleagues (Brysbaert, Van Dyck, & Van de Poel, 1999; Van Wijnendaele & Brysbaert, 2002) demonstrated that briefly presented French target words were more accurately identified by Dutch-French and French-Dutch bilinguals if they were preceded by a pseudoword that sounded similar when pronounced using Dutch spelling-sound correspondences than when preceded by an unrelated pseudoword. Because the prime was presented very briefly and masked, this finding suggests that participants automatically applied their knowledge of Dutch spelling-sound correspondences to the prime even though the task appeared to be entirely in French. In a study by Duyck (2005), critical primes were pseudohomophones of a Dutch word if Dutch spelling-sound correspondences were applied (e.g., *tauw* is a not a word in Dutch but it is pronounced like the Dutch *touw* 'rope') and targets were the English translation equivalent (e.g., *rope*). Dutch-English bilinguals made faster lexical decisions on target words when they were preceded by these pseudohomophone primes than when they were preceded by control primes, indicating that the phonological representations generated from the prime activated their corresponding meanings. In other masked priming studies that have used word primes, phonological priming effects have been observed when a bilingual's two languages have different scripts, such as with Hebrew-English (Gollan, Forster, & Frost, 1997), Korean-English (Kim & Davis, 2003), Greek-French (Voga & Grainger, 2007), Chinese-English (Zhou, Chen, Yang, & Dunlap, 2010), Japanese-English (Ando, Matsuki, Sheridan, & Jared, 2015; Nakayama, Sears, Hino, & Lupker, 2012), Russian-English (Jouravlev, Lupker, & Jared, 2014), and Greek-Spanish bilinguals (Dimitropoulou, Duñabeitia, & Carreiras, 2011). Dimitropoulou et al. provided evidence that dissimilarity in the prime and target scripts was necessary to get a facilitatory phonological priming effect with word primes. Otherwise, phonological facilitation effects are cancelled out by competition between the orthographic lexical representations of the prime and target. Facilitatory phonological priming effects probably arise as a consequence of feedback from sublexical phonological representations to orthographic representations, because they are unaffected by target word frequency or proficiency of participants (Ando et al., in press; Nakayama et al., 2012).

Cross-language effects have also been found in priming studies where primes and targets share meaning (for reviews, see Altarriba & Basnight-Brown, 2007, 2009; Basnight-Brown & Altarriba, 2007). The most robust cross-language priming effects have been observed when primes are translations of targets (e.g., Duñabeitia, Perea, & Carreiras, 2010; Hoshino, Midgley, Holcomb, & Grainger, 2010; Schoonbaert, Duyck, Brysbaert, & Hartsuiker, 2009). In the Schoonbaert et al. study, the size of the translation priming effect did not depend on whether the words were abstract or concrete, suggesting that the effect could arise from lexical connections between translation equivalents.

Convincing evidence that both languages of a bilingual activate representations in a shared semantic store comes from experiments finding cross-language semantic priming (e.g., Perea, Duñabeitia, & Carreiras, 2008; Schoonbaert et al., 2009; Williams, 1994). Semantic priming effects arise when semantic features that are activated by the prime are also activated by the target word. Finkbeiner, Forster, Nicol, and Nakamura (2004) pointed out that many words share only some of their senses with their translation equivalent in another language and have other senses that are language-specific (e.g., *blue* can mean 'sad' in English but not in French). They hypothesized that the size of the facilitation effect in cross-language priming experiments depends on the number of shared senses that are known to the bilingual and on the ratio of primed to unprimed senses.

Words in Sentence Contexts

In the research discussed so far, evidence that both languages are activated when reading in one language has come from studies in which words were presented one at a time or in prime-target pairs. In contrast, in most natural reading, words are embedded in sentences. The sentence contains plenty of clues about the language of the text, and researchers have wondered whether these clues can be used to limit activation to a single language (for a review, see Schwartz & Van Hell, 2012).

Studies addressing this question have included cognates and interlingual homographs in sentence contexts and compared their reading times to those of matched controls (e.g., Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Libben & Titone, 2009; Schwartz & Kroll, 2006; Van Hell & De Groot, 2008). Cognates in sentence contexts produced facilitation compared with noncognates, whereas homographs typically had longer reading times than matched controls. Furthermore, these cross-language effects depended in part on degree of sentence constraint. For example, Libben and Titone (2009) showed that in low-constraint sentences, cognate facilitation and homograph inhibition occur across all fixation duration measures. In high-constraint sentences, these effects were observed in initial fixation measures (i.e., first-fixation duration, first-pass gaze duration), but not in the later measures (i.e., go-past time, total gaze duration; see Schotter & Rayner, this volume, for a discussion of these measures). Libben and Titone concluded that language nonselective activation occurs initially, but that semantic ambiguity resolves rapidly in constraining contexts. Jouravlev and Jared (2014) found an interlingual homograph effect even when the alphabets of the bilinguals' two languages used mostly different letters (Russian and English), indicating that clues from the script do not limit activation to a single language.

Further evidence for cross-language activation of words in sentence contexts comes from studies using homonym cognates (see Schwartz & Van Hell, 2012). For example, Schwartz, Yeh, and Shaw

(2008) showed that the strength of activation of the meanings of a homonym (e.g., *novel,* which refers to a story but also to something new) is influenced by knowledge of a cognate that shares a meaning (*novela* in Spanish shares the meaning 'story'). On critical trials, sentence contexts biased the subordinate meanings of English homonyms, which either did (e.g., *She is an original thinker and her ideas are **novel***) or did not have a Spanish cognate (e.g., *In observance of the religious holiday the family observed the **fast***). After the homonym was presented, a target word appeared that on critical trials was related to the dominant meaning of the homonym (e.g., *book, speed*), and Spanish-English bilingual participants had to indicate whether the target word was related to the sentence. Participants were slower and made more errors in rejecting target words that followed cognate homonyms than noncognate homonyms (e.g., it was harder to reject *book* after reading the preceing sentence ending in *novel* than to reject *speed* after reading the sentence ending with *fast*). This finding suggests that the dominant meaning of the cognate homonym (e.g., *novel*) received a boost in activation from its Spanish cognate mate (e.g., *novela*), and provides further evidence that the language of a sentence does not constrain activation to words of that language.

fMRI Studies

The studies just reviewed provide evidence that the languages of a bilingual influence one another when reading. Other studies have used fMRI to investigate whether reading words in a second language (L2) recruits the same brain network as is used to read words in the first language (L1), or whether different brain structures are recruited (for a review, see Cao, Tao, Liu, Perfetti, & Booth, 2013). Bilinguals who know Chinese and English are of particular interest because the two languages are so different. Intuitively, one would expect that very different languages would be least likely to involve the same brain structures. However, Cao et al. (2013) found that Chinese-English bilinguals who had begun to acquire English around age 11, when performing a rhyme judgment task with printed English words, had activation patterns similar to those shown by Chinese monolinguals when performing a rhyme judgment task with printed Chinese words. The main difference was that the bilinguals showed reduced activation in the right middle occipital gyrus, an area involved in visual word form processing. Stronger activation in this area in reading Chinese may be due to the greater visual complexity of Chinese characters. Higher proficiency in English was associated with greater involvement of the Chinese network and reduced involvement of the English network. In contrast, a study with English-Chinese bilinguals (Cao et al., 2012) found greater activation of the left medial frontal gyrus when reading Chinese than when reading English. This area is associated with lexical retrieval and integration in reading Chinese. Cao et al. (2013) speculated that a reader who has first learned to read in a language with a relatively arbitrary relationship between orthography and phonology, such as Chinese, may be able to assimilate a new language with more transparent spelling-sound correspondence in the same brain structures, but that a reader who has first learned to read in a more transparent language may need to recruit new structures when learning to read in a less transparent language. This intriguing idea needs further study.

In fMRI studies of early and fluent bilinguals, some differences in brain areas activated by the two languages have been found, both in studies that tested bilinguals whose languages are quite dissimilar, such as in Bick, Goelman, and Frost's (2011) study of Hebrew-English bilinguals and Das, Padakannaya, Pugh, and Singh's (2011) study of Hindi-English bilinguals, and also when bilinguals' languages are more similar, as in a study by Jamal, Piche, Napoliello, Perfetti, and Eden (2012) of Spanish-English bilinguals. Although a number of brain areas are activated by both of a bilingual's languages, these studies reveal at least some language-specific representations for word recognition in proficient early bilinguals that are attributable to differences in the orthographic properties and spelling-sound transparency of the two languages.

The Bilingual Interactive Activation Model

One theoretical model has dominated the field of bilingual word recognition and can account for many of the findings just reviewed. The bilingual interactive-activation model (BIA+; Dijkstra & Van Heuven, 2002; see Figure 12.1) is an extension of McClelland and Rumelhart's (1981) interactive-activation monolingual model of word recognition. Briefly, the model assumes that representations for the two languages of a bilingual are stored together, that representations from both of a bilingual's languages are activated when reading in one language, and that bilinguals cannot inhibit representations belonging to a language that is not needed for a task.

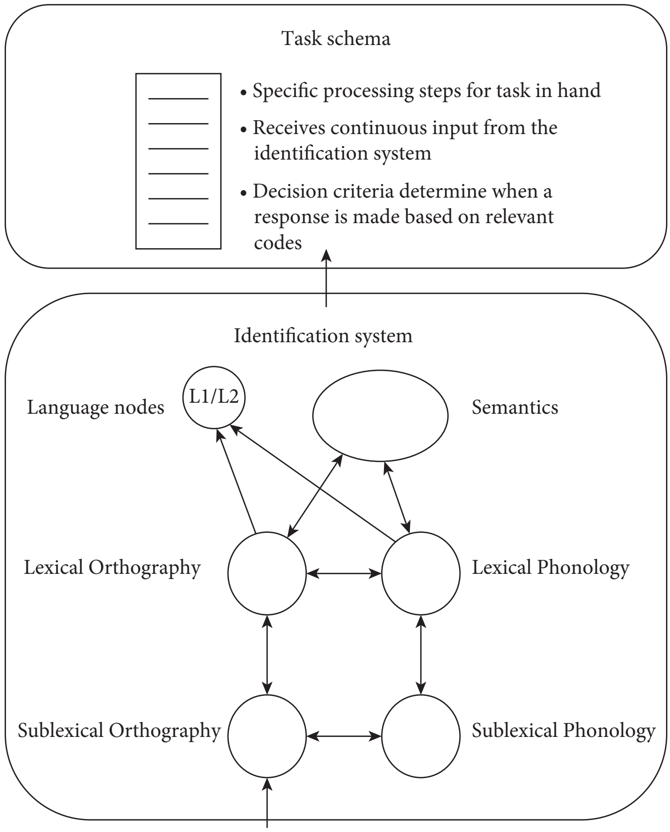


Fig. 12.1 The BIA+ model. Reprinted from "The Architecture of the Bilingual Word Recognition System: From Identification to Decision," by T. Dijkstra and W. J. B Van Heuven, 2002, *Bilingualism: Language and Cognition, 5,* Fig. 2. Reprinted with the permission of Cambridge University Press.

More specifically, there are pools of nodes representing sublexical orthographic information and sublexical phonological information, as well as pools of lexical orthographic nodes and lexical phonological nodes; each of these pools contains representations from both languages. If a bilingual's languages use the same alphabet, the sublexical orthographic nodes will be entirely shared between the two languages. In other cases, such as English and Russian, some sublexical nodes will represent shared features and some will represent features that are unique to each language, and in still other cases, such as English and Chinese, each language will have its own sublexical orthographic nodes within the same pool. The same is true for representations of sublexical phonological information. Sublexical orthographic representations connect to lexical orthographic representations and to sublexical phonological representations, and the latter connect to lexical phonological representations. There are connections between lexical orthographic nodes and lexical phonological nodes, and from lexical nodes to nodes representing semantic information, as well as to one of two language nodes. The connection from a lexical representation to a language node is how the model codes the language status of a word. Connections feed activation forward and backward, with the exception of the language nodes, which can only feed activation forward. Activated sublexical representations send activation to any lexical representations that they are consistent with, regardless of language. Nodes within a pool compete with one another. The resting level of word nodes depends on subjective frequency, with the consequence that words from a less proficient language will typically be activated more slowly than words from a more proficient language, and will be less likely to influence the activation of words in a more proficient language than the reverse. This assumption is consistent with findings that cross-language effects are larger when bilinguals perform tasks in their L2 than in their L1, and indeed studies often only test bilinguals in their L2 (see Van Hell & Tanner, 2012, for a review of how proficiency modulates coactivation). Because there are feedback connections from the semantic nodes to the lexical nodes, the prior semantic context of a sentence can influence the activation level of lexical nodes, but the language of the sentence cannot.

If bilinguals activate representations from both languages when reading in one, how do they select the representation in the appropriate language? Selecting the representation in the correct language is potentially more of a challenge when reading in L2 than in L1 because L1 representations would be stronger competitors for L2 than the reverse. Response selection is a particularly active topic of research in the bilingual speech production literature (e.g., Misra, Guo, Bobb, & Kroll, 2012). In the original BIA (Dijkstra & Van Heuven, 1998), the language nodes collected activation from nonlinguistic sources of information (e.g., task instructions) as well as from word representations, and could send inhibition to word-level representations in the language that was not needed for the task. In the revised version of the model, the BIA+, language nodes collect activation from within the word identification system only, and top-down language-to-word inhibition was eliminated. A task schema system was added to the model that makes use of the output of the word recognition system. When the task requires responses from a single language, the task schema at the decision stage excludes responses from the other language. Dijkstra and Van Heuven (2002) made this change because representations from the nontarget language appear to influence activation levels of target language representations regardless of task demands or participant strategies.

The BIA+ is broadly consistent with the findings in the literature on bilingual word recognition. However, a full implementation of the model would allow for a better evaluation. For example, it is unclear whether the assumption of competing lexical phonological representations will allow the model to produce facilitatory phonological effects such as those found for homophones and in priming studies. Simulations of human data were essential for refining computational models of English word recognition (e.g., Coltheart et al., 2001; Plaut, McClelland, Seidenberg, & Patterson, 1996). Dijkstra and colleagues (Dijkstra, Hilberink-Schulpen, & Van Heuven, 2010; Lam & Dijkstra, 2010) reported simulations of form priming with just the orthographic part of the model, and they also reported some success in extending the framework to include phonological representations in a monolingual version of the model called SOPHIA. However, they acknowledged difficulties in extending the model to include phonological representations for two languages. A problem with the interactive-activation framework in general concerns the way in which orthography is coded. There is a separate letter position channel for each letter in a stimulus. Studies of letter transposition effects in monolingual English readers have provided evidence that this assumption is not correct (e.g., Frost, this volume; Perea & Lupker, 2006). Another limitation of the model is that it does not include a learning mechanism, in contrast to distributed connectionist models of English word recognition (e.g., Seidenberg & McClelland, 1989). Thomas and Van Heuven (2005) presented some ideas on how a distributed bilingual model of word recognition could be constructed. One such model is presented in the later section of this chapter on biliteracy development in children. A challenge for progress in the field is that computational models will have to be developed for each pair of languages being studied.

The BIA+ postulates integrated representations, but future research could examine this assumption more closely to see whether the extent of language integration depends on timing and context of learning and the characteristics of the bilingual's two languages (e.g., Zhao & Li, 2010). Another possibility is that sublexical orthographic and sublexical phonological representations are integrated across languages but lexicons for each language are separate (see Kroll, Van Hell, Tokowicz, & Green, 2010, for arguments in favor of separate but interconnected lexicons).

Sentence Processing

A small but growing body of literature has investigated bilingual sentence processing (for reviews, see Clahsen & Felser, 2006; Tolentino & Tokowicz, 2011; Van Hell & Tokowicz, 2010). This is an area where research on reading and oral language comprehension intersect, with studies in both modalities addressing similar questions and using similar materials. A central question in this research is whether bilinguals process sentences in their L2 similarly to natives in that language. To stay with the theme of this chapter on interactions between languages in bilingual reading, the more specific question considered here is whether sentence comprehension in one language is influenced by knowledge of another language. The studies to be discussed use visually presented stimuli.

THE COMPETITION MODEL

A source of inspiration for a number of studies on sentence processing in bilinguals is MacWhinney's competition model. This model was first developed to account for sentence processing in monolinguals, and later extended to bilinguals (MacWhinney, 1997, 2005). In the competition model, language learning involves learning cues (e.g., word order, animacy) that help determine form-function relations (e.g., *who is the actor?*) when interpreting sentences. The strength of the cues depends on their availability and reliability. When the cues provide conflicting interpretations, they compete. The more exposure one has to one's L1, the stronger the relevant cues become. This view predicts that learning an L2 will be easier if the same cues that are relevant in L1 are also relevant in L2, and will be harder if different cues are relevant in the L2. For structures that are unique in L2 there will be no competition from L1.

Tokowicz and MacWhinney (2005) provided support for these predictions of the competition model in an ERP study of English learners of Spanish. Their critical stimuli involved three types of morphosyntactic violations in Spanish sentences. One type involved a Spanish construction that is similar to English. Both languages form the progressive tenses by placing the auxiliary before the participle, and so omitting the auxiliary is a syntactic violation in both languages (e.g., *Su abuela cocinando muy bien* 'His grandmother cooking very well'). The second type involved a construction that differed in Spanish and English. In Spanish, different determiners are used for singular (e.g., *el*) and plural (e.g., *los*), but in English the same determiner is used for both (e.g., *the*). Syntactic violations in Spanish sentences involved mismatches in number between the determiner and the noun (e.g., *El niños están jugando* 'The boys are playing'; *El* should be *Los*). The third type involved a construction that is unique to Spanish. Violations of gender agreement between the determiner and the noun occur only in Spanish (e.g., *Ellos fueron a un fiesta* 'They went to a party'; *fiesta* should have the feminine determiner *una*). The researchers compared the size of the P600, a positive-going wave that is presumed to reflect a late controlled process of re-analysis and repair of effortful syntactic integration, when their participants read sentences with the three types of syntactic violations with the waveform produced for matched grammatically correct sentences. Participants produced a greater P600 in response to morphosyntactic violations in Spanish sentences for Spanish constructions that are formed similarly to English and for constructions that are unique to Spanish, but not for constructions where Spanish differs from English. These findings indicate that knowledge of English influenced participants' ability to detect morphosyntactic violations in Spanish. Tolentino and Tokowicz (2011) review further evidence for the competition model.

A concern about using anomalous sentences to understand how bilinguals process L2 sentences was raised by Frenck-Mestre (2005a). She argued that it is not evident that the ability to detect illegal structures is a sensitive measure of L2 readers' ability to understand legal sentences. Other research uses correct but ambiguous sentences—for example, sentences with relative clauses such as *Someone shot the son of the actress who was on the balcony*. This type of sentence is interesting because English speakers prefer *the actress* as the subject of the relative clause (*Who was on the balcony?*) whereas Spanish and French speakers prefer *the son* (low and high attachment, respectively). Such differences allow researchers to examine the influence of the parsing preferences in one language on comprehension of the other language. In two such studies, participants read sentences as their eye movements were monitored.

Frenck-Mestre (2002, 2005b) presented French sentences to Spanish and English learners of French, proficient English-French bilinguals, and French monolinguals. The correct attachment was cued by the second verb in the sentence. Like the French monolinguals, the Spanish-French participants had shorter first-pass gaze durations on the disambiguating verb when it signaled that high attachment was the correct solution, whereas the English learners of French had shorter gaze durations on this verb when low attachment was the correct solution, indicating that beginning bilinguals showed the same attachment preference as in their L1. However, the proficient English-French bilinguals had shorter first-pass gaze durations in the disambiguating region when high attachment was the correct solution. This result provided evidence that they had adopted the same preferences as native French speakers and indicates that L2 parsing strategies change as proficiency develops.

Conversely, Dussias and Sagarra (2007) showed that L2 preferences could impact L1 processing. They presented sentences in Spanish to Spanish-English bilinguals who were either immersed in a Spanish or an English environment as well as to Spanish monolinguals. The sentences had a relative clause construction and used morphological gender to cue either high or low attachment. The results for Spanish-English bilinguals who were immersed in a Spanish environment were similar to those for monolingual Spanish speakers: Both had shorter total fixation times for sentences in which high attachment was correct. In contrast, Spanish-English bilinguals immersed in an English environment showed shorter fixation times when low attachment was correct.

The studies just reviewed have provided evidence that the languages of a bilingual can influence one another at the sentence level when reading. More specifically, the studies provide evidence that the ease of sentence processing in one language is influenced by the similarity of specific constructions to those in the other language.

The final study in this section on sentence processing used fMRI to investigate the related issue of whether reading sentences in an L2 recruits the same brain network as is used to read sentences in L1, or whether different brain structures are recruited (for reviews, see Abutalebi, 2008; Abutalebi, Cappa, & Perani, 2005). Wartenburger et al. (2003) considered whether age of acquisition (AOA) of the L2 and proficiency in the L2 influences the similarity of the brain structures that are recruited for the two languages. They compared the performance of three groups of Italian-German bilinguals—early AOA-high proficiency, late AOA-high proficiency, and late AOA-low proficiency—on tasks that involved judging whether sentences contained semantic and grammatical anomalies. The anomalies in the grammatical judgment task included disagreements of number, gender, or case. For early AOA bilinguals the extent of neural activity was the same for L1 and L2, but for late bilinguals it was different. The early and late AOA bilinguals who were both high in proficiency did not differ in the behavioral data, and they also had similar brain activation patterns when performing the semantic judgment task. However, the late AOA high proficiency bilinguals engaged more extended areas of the prefrontal cortex when they performed the syntactic judgment task. The latter finding provides evidence that AOA affects the neural representations involved in grammatical processing. The brain activation patterns for the two late AOA bilingual groups differed on the semantic task, suggesting that proficiency impacts cerebral organization of semantic information.

TEXT PROCESSING

Very little research has investigated text reading processes in bilinguals (see O'Brien & Cook, this volume, for a review of research with monolinguals). One question is whether comprehension processes that are developed through reading in one language transfer to the reading of another language. Kintsch and Van Dijk's (1978) theory of text processing in monolinguals postulates three separate levels of text representation. These representations are hierarchical in nature, with surface form, textbase, and situation model constituting the levels from lowest to highest, respectively. The surface form of a text includes its wording and syntax. The textbase is the meaning of the text represented in the form of propositions. These propositions are not contingent on the exact wording of the text. The situation model is created from the textbase and a reader's background knowledge to form an overall impression of the text. It also includes any inferences drawn from the text. Using this framework, Raney, Obeidallah, and Miura (2002) speculated that the ability to create textbase representations from the surface form is likely to depend largely on language-specific reading experience, whereas the ability to develop situation models from textbase representations may benefit from reading experience in either language. This view is consistent with that of Cummins (1991), who proposed that lower-level comprehension processes (e.g., lexical and syntactic analyses) are language specific and higher-level processes (e.g., integration and comprehension) are language nonspecific.

Friesen and Jared (2007) investigated message-level and word-level transfer effects from a passage in one language to a passage in the other language. One question they addressed was whether a cognate presented in a text in one language would facilitate the subsequent reading of the cognate when it appeared in a text in the other language. Raney (2003) suggested that whether or not word-level transfer effects would be seen across passages depended on the quality of the situation model that is formed on reading the first text and the overlap between that situation model and the model formed from reading the second text. When a good quality situation model is formed, a cognate in one text should influence the reading time of that cognate in another text in a different language only when a similar situation model is formed on reading the second passage. However, if a poor situation model is developed, perhaps because of lack of fluency in the language of that text, the surface form and textbase are not tightly bound to the text representation. Therefore, cognates in the first text should influence reading times on the same cognates in a second text in a different language regardless of the similarity between the texts. The results of Friesen and Jared (2007) provided some support for this view. Clearly, much more is to be learned about how bilinguals represent and process texts.

L2 Literacy Acquisition in Adults

There is a long history of research on individual differences in L2 acquisition (for a review, see Dewaele, 2009). Here I focus on a new initiative by Frost (2012; Frost, Siegelman, Narkiss, & Afek, 2013) on statistical learning ability, because it is related to the theme of this chapter on interactions between languages in bilingual reading. Connectionist models of monolingual word recognition (e.g., Seidenberg & McClelland, 1989) assume that skilled word reading involves learning the statistical properties of the language, and empirical studies support this assumption (e.g., Jared, McRae, & Seidenberg, 1990; Treiman, Mullenix, Bijeljac-Babic, & Richmond-Welty, 1995). The statistical relationships include the co-occurrence of letters in printed words in the language, the relationships between letters and sounds, and the relationships between letter clusters that represent morphemes and semantic representations. Frost (2012) noted that these statistical relationships differ between languages. Skilled readers of an L1 will have developed a sensitivity to the statistical properties of that language. When acquiring literacy in an L2, they have to implicitly learn a new set of statistical regularities. In this view, the ease of learning to read in a new language depends in part on the similarity of the statistical properties of the two languages and on the individual's statistical learning ability. The latter is thought to be a domain-general cognitive ability. Frost et al. (2013) sought evidence for this view in a study of adult English learners of Hebrew, whose two languages have very different statistical properties. Participants were given a visual statistical learning task, which assessed their ability to detect the implicit transitional probabilities embedded in a continuous stream of visual shapes, and they also completed three Hebrew reading tasks in both the first and second semester of their Hebrew course. Scores on the visual statistical learning task were correlated with the amount of improvement on the Hebrew reading tasks. The authors concluded that the ability to pick up co-occurrences in the environment is the common underlying ability relating performance on the visual statistical learning task and the learning of the structural properties of Hebrew. This intriguing hypothesis and the idea that the similarity of the statistical properties of the two languages influences L2 learning deserve further study.

Biliteracy Development in Children

More children than ever are acquiring a second language, partly due to increases in migration but also due to a desire by many parents that their children be prepared to participate in a global economy. Until relatively recently, the vast majority of the literature on reading development investigated monolingual speakers of English (Share, 2008). However, there has been a rapidly growing interest in reading development in children who are exposed to more than one language.

Reading Outcomes

One question that researchers have addressed is whether reading outcomes for students who are learning to read in their second language (or in two languages at once) differ from outcomes observed in monolingual students. Studies have shown that children who learn to read in a language that is new to them typically develop word-reading skills at a similar pace as children who are native speakers (see Lesaux, Koda, Siegel, & Shanahan, 2006, for a review). However, there is a wide variation in the development of text reading skills (e.g., Nakamoto, Lindsey, & Manis, 2007, 2008; Siegel, 2011) that is largely due to differences in the development of L2 oral language skills (see Leider, Proctor, Silverman, & Harring, 2013, for a review). Acquisition of L2 oral language skills is facilitated by interactions with native speaking peers, but may be delayed if many other children in the learning context speak the same L1 (e.g., Oller, Jarmulowicz, Pearson, & Cobo-Lewis, 2010).

Predictors of L2 Reading Ability

A much larger body of research explores whether individual differences in L2 reading outcomes can be predicted by various tests of cognitive and language abilities (see August & Shanahan, 2006, for a review). A practical aim of this research is to identify students who are at risk for reading difficulties in L2. Studies that investigate the ability of L1 language measures to predict children's reading development in their L2 are most relevant to the current focus on how knowledge of one language influences reading in the other (for reviews, see Genesee & Geva, 2006;

Genesee, Geva, Dressler, & Kamil, 2006). Three examples of longitudinal studies that first tested children in kindergarten in L1 and assessed subsequent reading ability in L2 are a study by Manis and colleagues of Spanish-English bilinguals (Lindsey, Manis, & Bailey, 2003; Manis, Lindsey, & Bailey, 2004; Nakamoto, Lindsey, & Manis, 2007, 2008), a study of English-French bilinguals by Jared, Cormier, Levy, and Wade-Woolley (2011), and a study of Chinese-English bilinguals by Pan et al. (2011). The conclusions that are discussed next are supported by these and other shorter-term studies.

Skills closely linked to word reading ability in studies of monolingual English children have provided clearer evidence of cross-language relationships than those related to reading comprehension. One ability that has repeatedly been shown to be a cross-language predictor of word reading skill is phonological awareness. Once children have understood that speech can be segmented into smaller units based on experience in one language, then that understanding carries over to a new language and facilitates the learning of relationships between printed and spoken words in the new language. Similarly, knowledge of L1 spelling-sound correspondences has been found to be a cross-language predictor of word reading ability. Although letter–sound correspondences may differ across languages, children appear to be able to transfer across languages the more general understanding that letters correspond to sounds in spoken words. Such an understanding would contribute to the development of efficient decoding skills. Other measures that have been shown to be cross-language predictors of word reading ability are rapid automatized naming (RAN), the ability to rapidly name sequences of alphanumeric characters, and morphological awareness (e.g., Deacon, Wade-Woolley, & Kirby, 2007), the understanding that some words can be decomposed into morphemes. L1 vocabulary and grammatical skills are more related to L2 reading comprehension than to L2 word identification, with L1 grammatical skills typically found to be more strongly related to L2 reading comprehension than L1 vocabulary knowledge. The latter finding suggests that there is an aspect of the ability to understand grammatical structures that is not specific to a particular language, but is more general, such that children who are better able to comprehend grammatical structures in their native language more readily learn to understand a second language. A possible underlying link is statistical learning ability. In summary, research on L1 predictors of L2 reading ability indicates that knowledge of one language can facilitate learning to read in another. Once reading development is underway in L2, evidence suggests that skills tested in L2 become better predictors of subsequent L2 reading skills (Leider et al., 2013; Mancilla-Martinez & Lesaux, 2010; Manis et al., 2004; Nakamoto et al., 2008).

Corroborating evidence that at least some of the skills that facilitate learning to read words are not language specific are fairly strong correlations between word reading abilities in both of a biliterate child's languages. Manis and colleagues (Lindsey et al., 2003; Nakamoto et al., 2007) found that Spanish and English word identification scores had a correlation of .66 in first grade and .68 in third grade, and in Jared et al.'s (2011) study the correlations between English and French word identification were .65, .66, and .65 for first through third grades respectively. In Pan et al.'s (2011) study of Chinese-English readers, the correlations were .37 at age 8 and .51 at age 10 (see also Jared & Kroll, 2011, Table 1).

Models of Biliteracy Development

Only a relatively small number of studies have investigated learning mechanisms and mental representations in young bilingual readers. The BIA+, discussed earlier, is a model of word recognition in proficient adult bilinguals and does not include any learning mechanisms. In the BIA+, a bilingual's two languages are integrated into single stores for each of the types of representations, allowing reading in one language to be influenced by representations in the other language. Developmental questions not addressed by this theory are whether these stores are integrated from the beginning of reading acquisition or become more integrated over time, and whether integrated representations would occur under all L2 learning situations. Developmental models of bilingual word recognition will need to be informed by data from bilingual children in a variety of learning situations and using a variety of language pairs.

Paradis's (1981, 2004) model of bilingual language processing (not reading specifically) does posit a developmental process. Paradis proposes that bilinguals have two subsets of neural connections, one for each language, within the same language system. The connections between the units within a subset are formed by their co-occurrence in the language input. That is, representations of words that belong to a specific language become strongly interconnected because they typically occur together in speech and print. The interconnectedness of the two subsystems and the degree of inhibition between them is assumed to depend on age of acquisition of the second language, the way in which language is learned, and the extent to which the individuals engage in language mixing. Several small-scale connectionist models (French, 1998; Li & Farkas, 2002; Thomas, 1997) have shown that it is possible for a system to develop language subsets from input in both languages (see De Groot, 2011; French & Jaquet, 2004, and Thomas & Van Heuven, 2005, for discussions of these models).

A larger-scale, self-organizing neural network model of the development of spoken language understanding in bilinguals has been developed by Zhao and Li (2010, 2013). Interestingly, the model produced different representational maps depending on the timing of the introduction of L2. When L2 was learned simultaneously with L1 or after a short delay, the model developed fairly separate representations for words in the two languages. However, when L2 was introduced after a substantial delay, representations for translation equivalent words were much closer together. Zhao and Li claimed that these findings were due to differences in network plasticity at the time L2 was introduced. When L2 was introduced early, the network had enough plasticity that it could reorganize the lexical space to dedicate resources to L2. When L2 was introduced later, the system had already organized the space based on L1, so L2 had to use existing structures and associative connections that were set up for L1. The model predicts that cross-language interactions are more likely in later bilinguals than in early bilinguals, and that in later bilinguals L1 is more likely to impact processing in L2 than the reverse. Although this is a model of spoken language development, the same principles could apply to the organization of orthographic and phonological lexicons in a reading model.

The first computational model of biliteracy development simulates learning to read in English and Chinese (Yang, Shu, McCandliss & Zevin, 2013). The goal was to show that these two very different written languages could be learned within the same connectionist architecture and using the same learning principles. Yang et al.'s view is that "reading skill is acquired by the application of statistical learning rules to mappings among print, sound and meaning, and that differences in the typical and disordered acquisition of reading skill between writing systems are driven by differences in the statistical patterns of the writing systems themselves, rather than differences in cognitive architecture of the learner" (p. 354). English has more reliable mappings between print and sound than Chinese, whereas Chinese has semantic radicals that provide more information about the mapping of print to meaning than is found in English monomorphemic words. These differences between languages are assumed to result in a different division of labor between orthographic-phonological and orthographic-semantic-phonological pathways in naming words in the two languages. The computational model is similar to Harm and Seidenberg's (1999, 2004) connectionist model of English reading development, except that it also includes a second orthographic input layer for Chinese and more units in the phonological layer to be able to encode Chinese pronunciations. Training on the two languages was interleaved, with each language being presented equally often. The model learned to produce correct English pronunciations faster than Chinese, which captures differences in children's rates of learning each language. To simulate reading impairments, a decay was applied to either the hidden units from orthography to phonology or the hidden units from orthography to semantics. The phonological impairment had a greater impact on the model's accuracy in English, whereas the semantic impairment had a greater impact on Chinese accuracy. This result provided support for the hypothesis of Yang et al. that there is a different division of labor between the two reading pathways for English and Chinese. This model is an important advance in the field of biliteracy acquisition. The next steps would be to explore how reading in one language is influenced by knowledge of the other language, and to develop a model for biliteracy acquisition in two languages that share the same alphabet and to compare the behavior of that model with the Yan et al. model.

Relevant Empirical Research

There has been little research on bilingual children comparable to the research on the adult bilingual word recognition system with which to inform a developmental model of bilingual reading acquisition. Just a few studies have investigated whether reading words in one language is influenced by knowledge of another language.

Two studies with children that have used cross-language words have provided evidence that representations in both languages were activated when reading in one and have shown that such cross-language effects are particularly evident when reading in L2. Brenders, Van Hell, and Dijkstra (2011) asked Dutch learners of English in fifth, seventh, and ninth grades to perform a lexical decision task that included Dutch-English cognates and matched control words. When the participants were to decide whether stimuli were English words, they responded to cognates faster than to English control words. In a Dutch version of the task, there was no difference in the responses to cognates and Dutch control words. A subsequent English experiment found that interlingual homographs were responded to more slowly than control words. In a study with younger L2 learners, Jared, Cormier, Levy, and Wade-Woolley (2012) asked native English-speaking children in third grade who were in French immersion programs to name cognates, interlingual homographs, and interlingual homophones and matched controls, either in English or in French. When the naming task was in French (L2), there were facilitatory cognate and interlingual homophone effects and an inhibitory interlingual homograph effect. Only the interlingual homograph effect was present when the naming task was in English. The results of these studies suggest that the word recognition systems of young bilinguals are interconnected. Cognates facilitate word reading due to shared meaning across languages, interlingual homophones facilitate naming due to shared phonology across languages, and interlingual homographs interfere with word reading due to conflicting meanings and pronunciations across languages. Even in the younger readers, L1 word recognition processes were sufficiently established to show little influence of L2 except in the case of homographs, which have the greatest cross-language conflict.

A prominent theory of reading development by Share (1995) proposes that word-specific orthographic representations are built up from successful decoding encounters. Schwartz, Kahn-Horwitz, and Share (2014) recently provided evidence that the acquisition of such representations in a new language is facilitated by the ability to read in a similar orthography. The three groups of sixth-grade English language learners in their study all read and spoke Hebrew. Two of the groups also spoke Russian, but only one of the groups could read it. The researchers hypothesized that because Russian orthography is more similar to English orthography than is Hebrew orthography, the participants who could read Russian should have an advantage in orthographic learning in English over the other two groups. Including a group who could speak Russian but not read it ensures that any advantages found for the group who could read Russian could be attributed specifically to knowledge of Russian orthography and not to knowledge of the language in general. Participants read short English passages that each contained a pseudoword; one week later they were shown pairs of pseudowords, one of which they had seen and one that was a homophone of the target, and they had to select which they had seen. All participants performed above chance, but the participants who read Russian were significantly more accurate than the other two groups, who did not differ. The results provide evidence that the orthographic distance between a learner's L1 and L2 influences the ease with which they learn the orthography of words in L2.

Children who are learning to read in two languages that use the same alphabet need to learn about the letter sequences that occur in each language. For example, although both English and French have words with the letter sequence *ain* (e.g., *main, pain*), the French word *beurre* contains a sequence of letters, *eurre*, that is not seen in English and the English word *lawn* contains a sequence of letters, *awn*, that is not found in French. Research has shown that by the end of first grade, monolingual children are sensitive to the co-occurrence of letters in their language (Cassar & Treiman, 1997; Pacton, Perruchet, Fayol, & Cleeremans, 2001). Pacton et al. argued that this sensitivity results from a statistical learning process in which children implicitly extract information about the frequency and co-occurrence of letters from the printed material to which they are exposed. A question of interest for children who are learning to read in two languages is whether this learning process in one language is influenced by exposure to the other language.

Sensitivity to orthographic patterns is often assessed in monolinguals by giving participants pairs of pseudowords (e.g., *filv-filk*) and asking them to choose which could be an acceptable word (Siegel, Share, & Geva, 1995). Van der Leij, Bekebrede, and Kotterink (2010) compared the performance of two groups of Dutch children in third grade one of which received reading instruction in both Dutch and English and the other of which was learning to read in Dutch only, on both Dutch and English versions of the task. The biliterate group outperformed the monolingual group on the English version, but no differences were found between the groups on the Dutch version. These findings indicate that the bilingual group had begun to pick up on the orthographic patterns typical of English and that exposure to English orthography did not interfere with their learning of Dutch letter patterns.

Deacon, Commissaire, Chen, and Pasquarella (2013) showed that by the end of first grade, English native speakers in French immersion programs could discriminate between pseudowords with illegal letter patterns in both English and French and pseudowords with legal letter patterns in these languages, and were more accurate when the letter patterns were legal in both languages than when they were legal in just one language. This latter finding suggests either that exposure to letter co-occurrences in one language facilitated learning letter co-occurrences in the other language or that children did not distinguish the co-occurrences for the two languages. In another study of French immersion students, Jared, Cormier, Levy, and Wade-Woolley (2013) showed that children in second grade could differentiate pseudowords with English-specific spelling patterns and French-specific spelling patterns at above-chance accuracy and that accuracy on the task improved by third grade. However, most children were still much less accurate than bilingual adults after three years of exposure to print, indicating that learning word-specific letter patterns for each language is a long process. Discrimination ability was strongly related to word identification ability in each language, and not to a nonverbal IQ test involving explicit analysis of patterns, providing some evidence of implicit learning from print. Stronger evidence for the statistical learning view would be a demonstration that children's knowledge of orthographic patterns was related to the frequency of the patterns occurring in the text to which they were exposed in each language and that performance on a test of statistical learning ability (e.g., Frost et al., 2013) was related to discrimination performance. Such a learning mechanism is consistent with connectionist models (e.g., Yang et al., 2013), and indeed data from children on the development of sensitivity to co-occurring letter patterns would be useful in the development of such models.

Summary

At present there is only a small body of literature that can inform cognitive models of the representations and processes involved in the development of word reading skills in biliterate children. Even less is known about the cognitive representations and developmental processes involved when children learn to read and comprehend sentences and texts in two languages. For example, do children find it easier to read sentences in L2 if the grammatical structure is similar to the structure in L1 than if it is unique to L2 and, if so, how long does the influence of L1 grammatical knowledge persist? How much language-specific information is in the mental representations that children form when reading texts in one language? The field now needs to move beyond the type of correlational studies reviewed at the start of this section and produce more research that elucidates what is happening in the mind of a child who is becoming biliterate.

Conclusion

As this review has shown, the study of bilingual reading and biliteracy acquisition has made great progress in the last decade or so, and has highlighted the complexity of the cognitive processing involved. It is now clear that a bilingual is not "two monolinguals in one head" (Grosjean, 1998)—that is, the two languages of a bilingual do not operate in isolation from one another. As interactions among people across the globe increase, the need to be able to read in more than one language increases. An understanding of the learning mechanisms for acquiring literacy in a new language may help more people successfully become biliterate. Literacy in another language in addition to one's native language not only can have benefits with respect to future employment but also can open up a person's view to another culture.

References

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica, 128,* 466–478.

Abutalebi, J., Cappa, S. F., & Perani, D. (2005). What can functional neuroimaging tell us about the bilingual brain? In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 497–515). New York, NY: Oxford University Press.

Altarriba, J., & Basnight-Brown, D. M. (2007). Methodological considerations in performing semantic- and translation-priming experiments across languages. *Behavior Research Methods, 39,* 1–18.

Altarriba, J., & Basnight-Brown, D. M. (2009). An overview of semantic processing in bilinguals: Methods and findings. In A. Pavleko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 79–98). Buffalo, NY: Multilingual Matters.

Ando, E., Matsuki, K., Sheridan, H., & Jared, D. (2015). The locus of Katakana-English masked phonological priming effects. *Bilingualism: Language and Cognition, 18,* 101–117.

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth*. Mahwah, NJ: Erlbaum.

Basnight-Brown, D. M., & Altarriba, J. (2007). Differences in semantic and translation priming across languages: The role of language direction and language dominance. *Memory & Cognition, 35,* 953–965.

Bick, A., Goelman, G., & Frost, R. (2011). Hebrew brain vs English brain: Language modulates the way it is processed. *Journal of Cognitive Neuroscience, 23,* 2260–2270.

Brenders, P., Van Hell, J. G., & Dijkstra, T. (2011). Word recognition in child second language learners: Evidence from cognates and false friends. *Journal of Experimental Child Psychology, 109,* 383–396.

Brysbaert, M., Van Dijk, G., & Van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 137–148.

Cao, F., Tao, R., Liu, L., Perfetti, C. A., & Booth, J. R. (2013). High proficiency in a second language is characterized by greater involvement of the first language network: Evidence from Chinese learners of English. *Journal of Cognitive Neuroscience, 25,* 1649–1663.

Cao, F., Vu, M., Chan, D. H. L., Lawrence, J. M., Harris, L. N., Guan, Q., . . . Perfetti, C. A. (2012). Writing affects the brain network of reading in Chinese: A functional magnetic resonance imaging study. *Human Brain Mapping, 34,* 1670–1684.

Carrasco-Ortiz, H., Midgley, K. J., Frenck-Mestre, C. (2012). Are phonological representations in bilinguals language specific? An ERP study on interlingual homophones. *Psychophysiology, 49,* 531–543.

Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology, 89,* 631–644.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27,* 3–42.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. C. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 70–89). Cambridge, England: Cambridge University Press.

Das, T., Padakannaya, P., Pugh, K. R., & Singh, N. C. (2011). Neuroimaging reveals dual routes to reading in simultaneous proficient readers of two orthographies. *NeuroImage, 54,* 1476–487.

Deacon, S. H., Commissaire, E., Chen, X., & Pasquarella, A. (2013). Learning about print: The development of orthographic processing and its relationship to word reading in first grade children in French immersion. *Reading and Writing: An Interdisciplinary Journal, 26,* 1087–1109.

Deacon, S. H., Wade-Woolley, L., & Kirby, J. (2007). Crossover: The role of morphological awareness in French Immersion children's reading. *Developmental Psychology, 43,* 732–746.

De Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals*. New York, NY: Psychology Press.

De Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in a bilingual and a monolingual task: Support for nonselective access to bilingual memory. *Quarterly Journal of Experimental Psychology, 53,* 397–428.

Dewaele, J.-M. (2009). Individual differences in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 623–646). Bingley, England: Emerald.

Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). New York, NY: Oxford University Press.

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41,* 496–518.

Dijkstra, T., Hilberink-Schulpen, B., & Van Heuven, W. J. B. (2010). Repetition and masked form priming within and between languages using word and nonword neighbors. *Bilingualism: Language and Cognition, 13,* 341–357.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language, 62,* 284–301.

Dijkstra, T., & Van Heuven, W. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–225). Mahwah, NJ: Erlbaum.

Dijkstra, T., & Van Heuven, W. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175–197.

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition, 1,* 51–66.

Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011). Phonology by itself: Masked phonological priming effects with and without orthographic overlap. *Journal of Cognitive Psychology, 23,* 185–203.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2010). Masked translation priming effects with highly proficient simultaneous bilinguals. *Experimental Psychology, 57,* 98–107.

Dussias, P. E, & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism: Language and Cognition, 10,* 101–116.

Duyck, W. (2005). Translation and associative priming with cross-lingual pseudohomophones: Evidence for nonselective phonological activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1340–1359.

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 663–679.

Finkbeiner, M., Forster, K. I., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming, *Journal of Memory and Language, 51,* 1–22.

French, R. M. (1998). A simple recurrent network model of bilingual memory. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the twentieth annual conference of the cognitive science society* (pp. 368–373). Mahwah, NJ: Erlbaum.

French, R. M., & Jaquet, M. (2004). Understanding bilingual memory: Models and data. *Trends in Cognitive Science,* 8, 87–93.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 217–236). Amsterdam, the Netherlands: Elsevier.

Frenck-Mestre, C. (2005a). Ambiguities and anomalies: What can eye movements and event-related potential revel about second language sentence processing? In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 268–281). New York, NY: Oxford University Press.

Frenck-Mestre, C. (2005b). Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research, 21,* 175–198.

Friesen, D. C., & Jared, D. (2007). Cross-language message- and word-level transfer effects in bilingual text processing. *Memory & Cognition, 35,* 1542–1556.

Friesen, D. C., & Jared, D. (2012). Cross-language phonological activation of meaning: Evidence from category verification. *Bilingualism: Language and Cognition, 15*, 145–156.

Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences, 35,* 263–279.

Frost, R., Siegelman, N., Narkiss, A., & Afek, L. (2013). What predicts successful literacy acquisition in a second language? *Psychological Science, 24*, 1243–1252.

Genesee, F., & Geva, E. (2006). Cross-linguistic relationships in working memory, phonological processes, and oral language. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth* (pp. 175–195). Mahwah, NJ: Erlbaum.

Genesee, F., Geva, E., Dressler, C., & Kamil, M. (2006). Synthesis: Cross-linguistic relationships. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth* (pp. 153–174). Mahwah, NJ: Erlbaum.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23,* 1122–1139.

Grainger, J., & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 207–220). Amsterdam, the Netherlands: Elsevier.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition, 1*, 131–149.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103,* 518–565.

Haigh, C. A., & Jared, D. (2007). The activation of phonological representations by bilinguals while reading silently: Evidence from interlingual homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33,* 623–644.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106,* 491–528.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111,* 662–720.

Heredia, R. R., & Altarriba, J. (Eds.). (2002). *Bilingual sentence processing*. Amsterdam, the Netherlands: Elsevier.

Hoshino, N., Midgley, K. J., Holcomb, P. J., & Grainger, J. (2010). An ERP investigation of masked cross-script translation priming. *Brain Research, 1344,* 159–172.

Jamal, N. I., Piche, A. W., Napoliello, E. M., Perfetti, C. A., & Eden, G. F. (2012). Neural basis of single-word reading in Spanish-English bilinguals. *Human Brain Mapping, 33,* 235–245.

Jared, D., Cormier, P., Levy, B. A., & Wade-Woolley, L. (2011). Early predictors of biliteracy development in children in French immersion: A 4-year longitudinal study. *Journal of Educational Psychology, 103,* 119–139.

Jared, D., Cormier, P., Levy, B. A., & Wade-Woolley, L. (2012). Cross-language activation of phonology in young bilingual readers. *Reading and Writing: An Interdisciplinary Journal, 25,* 1327–1343.

Jared, D., Cormier, P., Levy, B. A., & Wade-Woolley, L. (2013). Discrimination of English and French orthographic patterns by biliterate children. *Journal of Experimental Child Psychology, 114,* 469–488.

Jared, D., & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language, 44,* 2–31.

Jared, D., & Kroll, J. F. (2011). Cognitive processes in bilingual reading. In P. McCardle, B. Miller, J. R. Lee, & O. J. L. Tzeng (Eds.), *Dyslexia across languages* (pp. 262–280). Baltimore, MD: Brookes.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29,* 687–715.

Jared, D., & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Bilingualism: Language and Cognition,* 5, 225–239.

Jouravlev, O., & Jared, D. (2014). Reading Russian-English homographs in sentence contexts: Evidence from ERPs. *Bilingualism: Language and Cognition, 17,* 153–168.

Jouravlev, O., Lupker, S. J., & Jared, D. (2014). Cross-language phonological activation: Evidence from masked onset priming and ERPs. *Brain and Language, 134,* 11–32.

Kerkhofs, R., Dijkstra, T., Chwilla, D. J., de Bruijn, E. R. (2006). Testing a model for bilingual semantic priming with interlingual homographs: RT and N400 effects. *Brain Research, 1068,* 170–183.

Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language, 49*, 484–499.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363–394.

Kroll, J., Gullifer, J. W., & Rossi, E. (2013). The multilingual lexicon: The cognitive and neural basis of lexical comprehension and production in two or more languages. *Annual Review of Applied Linguistics, 33,* 102–127.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism: Language and Cognition, 13*, 373–381.

Lam, K. J., & Dijkstra, T. (2010). Word repetition, masked orthographic priming, and language switching: Bilingual studies and BIA+ simulations. *International Journal of Bilingual Education and Bilingualism, 13*, 487–503.

Leider, C. M. Proctor, C. P., Silverman, R. D., & Harring, J. R. (2013). Examining the role of vocabulary depth, cross-linguistic transfer, and types of reading measures on the reading comprehension of Latino bilinguals in elementary school. *Reading and Writing: An Interdisciplinary Journal, 26,* 1459–1485.

Lemhöfer, K., & Dijkstra, A. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language specific and generalized lexical decision. *Memory & Cognition, 32,* 533–550.

Lesaux, N., Koda, K., Siegel, L. S., & Shanahan, T. (2006). Development of literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth* (pp. 75–1122). Mahwah, NJ: Erlbaum.

Li, P., & Farkas, I. (2002). A self-organizing connectionist model of bilingual processing. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 58–85). Amsterdam, the Netherlands: Elsevier.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 381–390.

Lindsey, K. A., Manis, F. R., & Bailey, C. E. (2003). Prediction of first-grade reading in Spanish-speaking English-language learners, *Journal of Educational Psychology, 95,* 482–494.

MacWhinney, B. (1997). Second language acquisition and the competition model. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 113–142). Mahwah, NJ: Erlbaum.

MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). New York, NY: Oxford University Press.

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners. *Journal of Educational Psychology, 102,* 701–711.

Manis, F. R., Lindsey, K. A., & Bailey, C. E. (2004). Development of reading in Grades K-2 in Spanish-speaking English-language learners. *Learning Disabilities Research and Practice, 19,* 214–224.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

Midgley, K. J., Holcomb, P. J., Van Heuven, W. J. B., Grainger, J. (2008). An electrophysiological investigation of cross-language effects of orthographic neighborhood. *Brain Research, 1246,* 123–135.

Misra, M., Guo, T., Bobb, S. C., & Kroll, J. F. (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language, 67,* 224–237.

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading and Writing: An Interdisciplinary Journal, 20*, 691–719.

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2008). A cross-linguistic investigation of English language learners' reading comprehension in English and Spanish. *Scientific Studies of Reading, 12,* 351–371.

Nakayama, M., Sears, C. R., Hino, Y., & Lupker, S. J. (2012). Cross-script phonological priming for Japanese-English bilinguals: Evidence for integrated phonological representations. *Language and Cognitive Processes, 27,* 1563–1583.

Oller, D. K., Jarmulowicz, L., Pearson, B. Z., & Cobo-Lewis, A. B. (2010). Rapid spoken language shift in early second-language learning: The role of peers and effects on the first language. In A. Y. Durgunoğlu & C. Goldenberg (Eds.), *Language and literacy development in bilingual settings* (pp. 94–120). New York, NY: Guilford.

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General, 130,* 401–426.

Pan, J., McBride-Chang, C., Shu, H., Liu, H., Zhang, Y., & Li, H. (2011). What is in the naming? A 5-year longitudinal study of early rapid naming and phonological sensitivity in relation to subsequent reading skills in both native Chinese and English as a second language. *Journal of Educational Psychology, 103,* 897–908.

Paradis, M. (1981). Neurolinguisitic organization of a bilingual's two languages. In J. E. Copeland & P. W. Davis (Eds.), *The seventh LACUS Forum* (pp. 486–494). Columbia, SC: Hornbeam.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Philadelphia, PA: Benjamins.

Paulmann, S., Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2006). Is bilingual lexical access influenced by language context? *NeuroReport, 17,* 727–731.

Peeters, D., Dijkstra, T., & Grainger, J. (2013). The representation and processing of identical cognates by late bilinguals: RT and ERP effects. *Journal of Memory and Language, 68,* 315–332.

Perea, M., Duñabeitia, J. A., & Carreiras, M. (2008). Masked associative/semantic priming effects across languages with highly proficient bilinguals. *Journal of Memory and Language, 58,* 916–930.

Perea, M., & Lupker, S. J. (2006). Does *jugde* prime COURT? Transposed-letter similarity effects in masked associative priming. *Memory & Cognition, 31,* 829–841.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Raney, G. E. (2003). A context-dependent representation model for explaining text repetition effects. *Psychonomic Bulletin & Review, 10,* 15–28.

Raney, G. E., Obeidallah, S. M., & Miura, T. K. (2002). Text comprehension in bilinguals: Integrating perspectives on language representation and text processing. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 165–183). Amsterdam, the Netherlands: Elsevier.

Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2009). Semantic and translation priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition, 37,* 569–586.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55,* 197–212.

Schwartz, A. I., Kroll, J. F., & Diaz, M. (2007). Reading words in English and Spanish: Mapping orthography to phonology in two languages. *Language and Cognitive Processes, 22,* 106–129.

Schwartz, A. I., & Van Hell, J. G. (2012). Bilingual visual word recognition in sentence context. In J. Adelman (Ed.), *Visual word recognition: Vol. 2. Meaning, context, individuals and development* (pp. 131–150). New York, NY: Psychology Press.

Schwartz, A. I., Yeh, L., & Shaw, M. P. (2008). Lexical representation of second language words: Implications for second language vocabulary acquisition and use. *The Mental Lexicon, 3*, 309–324.

Schwartz, M., Kahn-Horwitz, J., & Share, D. (2014). Orthographic learning and self-teaching in a bilingual and biliterate context. *Journal of Experimental Child Psychology, 117,* 45–58.

Seidenberg, M. S. & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Share, D. L. (1995). Phonological recoding and self-teaching: *Sine qua non* of reading acquisition. *Cognition, 55,* 151–218.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134,* 584–615.

Siegel, L. S. (2011). Reducing reading difficulties in English L1 and L2: Early identification and Intervention. In P. McCardle, B. Miller, J. R. Lee, & O. J. L. Tzeng (Eds.), *Dyslexia across languages* (pp. 294–304). Baltimore, MD: Brookes.

Siegel, L. S., Share, D., & Geva, E. (1995). Evidence for superior orthographic skills in dyslexics. *Psychological Science, 6,* 250–254.

Smits, E., Martensen, H. E., Dijkstra, A. F. J., & Sandra, D. (2006). Naming interlingual homographs: Variable competition and the role of the decision system. *Bilingualism: Language and Cognition, 9*, 281–297.

Thomas, M. S. C. (1997). Distributed representations and the bilingual lexicon: One store or two? In J. Bullinaria, D. Glasspool, & G. Houghton (Eds.), *Proceedings of the fourth annual neural computation and psychology workshop* (pp. 240–253). London, England: Springer.

Thomas, M. S. C., & Van Heuven, W. J. B. (2005). Computational models of bilingual comprehension. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 202–225). New York, NY: Oxford University Press.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition, 27,* 173–204.

Tolentino, L. C., & Tokowicz, N. (2011). Across languages, space, and time: The role of cross-language similarity in L2 (morpho)syntactic processing as revealed by fMRI and ERP. *Studies in Second Language Acquisition, 33,* 91–125.

Treiman, R., Mullinnex, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology, General, 124,* 107–136.

Van der Leij, A., Bekebrede, J., & Kotterink, M. (2010). Acquiring reading and vocabulary in Dutch and English: The effect of concurrent instruction. *Reading and Writing: An Interdisciplinary Journal, 23,* 415–434.

Van Hell, J. G., & de Groot, A. M. B. (2008). Sentence context affects lexical decision and word translation. *Acta Psychologica, 128,* 431–451.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9*, 780–789.

Van Hell, J. G., & Tanner, D. S. (2012). Effects of second language proficiency on cross-language activation. *Language Learning, 62* (Issue Supplement s2), 148–171.

Van Hell, J. G., & Tokowicz, N. (2010). Event-related potentials and second language learning: Syntactic processing in late L2 learners at different proficiency levels. *Second Language Research, 26,* 43–74.

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39,* 458–483.

Van Heuven, W. J. B., Schriefers, H., Dijkstra, T., & Hagoort, P. (2008). Language conflict in the bilingual brain. *Cerebral Cortex, 18,* 2706–2716.

Van Wijnendaele, I., & Brysbaert, M. (2002). Visual word recognition in bilinguals: Phonological priming from the second to the first language. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 616–627.

Voga, M., & Grainger, J. (2007). Cognate status and cross-script translation priming. *Memory & Cognition, 35,* 938–952.

Von Studnitz, R. E., & Green, D. W. (2002). Interlingual homograph interference in German-English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition, 5,* 1–23.

Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron, 37,* 159–170.

Williams, J. N. (1994). The relationship between word meanings in the first and second language: Evidence for a common, but restricted, semantic code. *European Journal of Cognitive Psychology, 6,* 195–220.

Wu, Y. J., & Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *Journal of Neuroscience, 30,* 7646–7651.

Yang, J., Shu, H., McCandliss, B. D., & Zevin, J. D. (2013). Orthographic influences on division of labor in learning to read Chinese and English: Insights from computational modeling. *Bilingualism: Language and Cognition, 16,* 354–366.

Zhao, X., & Li, P. (2010). Bilingual lexical interactions in an unsupervised neural network model. *International Journal of Bilingual Education and Bilingualism, 13*, 505–524.

Zhao, X., & Li, P. (2013). Simulating cross-language priming with a dynamic computational model of the lexicon. *Bilingualism: Language and Cognition, 16,* 288–303.

Zhou, H., Chen, B., Yang, M., & Dunlap, S. (2010). Language nonselective access to phonological representations: Evidence from Chinese-English bilinguals. *Quarterly Journal of Experimental Psychology, 63,* 2051–2066.

PART 3

Reading Sentences and Texts

CHAPTER

13 The Role of Sound in Silent Reading

Alexander Pollatsek

Abstract

This chapter discusses whether sound is involved in the process of skilled (and apparently silent) reading of words and texts, and, if so, how. The term "phonological coding" encompasses a broad variety of phenomena, including inner speech and subvocalization. In the research on single-word encoding, the focus has largely been on the level of phonemic coding, and the controversies have largely been about whether readers do this encoding with something like a rule-governed process or by learning correlations between visual and auditory patterns. The chapter also reviews the large literature that examines phonological coding in reading sentences and text using eye-movement methodology, including display change techniques. Other aspects of phonological coding discussed include its role to mark stress and its role in short-term memory to facilitate the reading of text. In addition, the chapter attempts to clarify the relationship between phonological coding and subvocalization.

Key Words: phonological coding, subvocalization, phonemes, stress, eye movements, display change techniques

This chapter focuses on skilled adult reading and the extent to which coding the text into sound is involved in the process. This has been a contentious issue, partly because there is never complete agreement about what coding into sound means. Another reason that it is contentious is that opinions differ as to whether coding text into sound in reading is a good thing or a bad thing. Obviously, if this means coding the text all the way up to saying the words out loud as you read them, it would not be optimal for most purposes, as it would slow reading down below the normal reading rate of about 300 words per minute for skilled readers. However, there are many intermediate levels in which the auditory and speech systems could be involved in the reading process in skilled adult reading, short of overt speech. To introduce the topic, let me describe a simple uncontrolled experiment I performed on myself while reading an earlier draft of this first paragraph.

I read the paragraph three times: (1) out loud (but not as if I were presenting it to an audience); (2) almost silently, but in a whisper; and (3) silently but making sure that the sounds of the words went by in my head as I read. The results of the "experiment" were that my reading speeds in conditions (1), (2), and (3) were 137, 205, and 326 words per minute, respectively. Clearly, these numbers are not to be taken seriously; among other things, the reading speeds are undoubtedly too high, as the reader knew the content of the text. However, the results make some important points. First, the overt speech in condition (1) slows reading enormously, and even the subtle involvement of the speech tract in (2) still slows reading quite a bit. Second, surprisingly, reading was quite rapid in the condition where I forced myself to mentally sound out every word. (I ran a silent reading control condition where the reading rate was about 500 words per minute, but I was unconvinced that I was really reading everything.) Any of the things I did in my experiment could be described either as phonological coding or subvocalization, and there are undoubtedly many more processes that range from small movements of the lips and/or vocal tract to other representations of phonological codes that do not have any necessary relationship for the skilled reader to overt speech that could also be described by either of these terms. I will adopt the relatively neutral term *phonological coding* for some sort of attempt, either conscious or unconscious, of the reader to recode the visual information into some auditory and/or speech format. However, I think my informal experiment makes clear that it is implausible that anything like overt speech can be involved with fluent adult reading.

The issue of phonological coding in reading has been explored in many different paradigms. It has perhaps been most extensively explored with adult normal readers processing single words, usually using either brief presentation paradigms or speeded classification tasks. A related body of studies has used many of the same tasks on people with varieties of dyslexia. The third type of paradigm, which will be the primary focus of this chapter, employs experiments that examine adults reading sentences or short passages of text using eye-movement technology. This is perhaps the best way to study a close approximation of normal skilled reading in the laboratory and still have a moment-to-moment record of the reading process and be able to exert control of the reading process as well (Rayner & Pollatsek, 1989; Rayner, Pollatsek, Ashby, & Clifton, 2012; also see Schotter & Rayner, this volume).

This chapter will take as its primary goal to document that phonological coding occurs when skilled readers are reading text, and that the involvement of phonological coding in skilled reading goes beyond its involvement in the identification of individual words and that it plays a large role aiding the comprehension of sentences and discourse. There are two related but somewhat distinct roles that phonological codes could play in facilitating comprehension at these higher levels. The first is a memory function that is distinct from the encoding of the printed or written words. That is, as it is likely that the visual memory codes for arbitrary symbols such as the printed letter are quite short-lived (e.g., Craik & Lockhart, 1972), phonological coding of the text would serve a short-term memory function that should aid the reading process. The second memory function (which may be more active than the first) could help the reader to recreate prosodic aspects of the spoken language such as phrase and clause boundaries that would aid in text comprehension. This latter memory may not be strictly auditory, but could also involve suppressed motor components. In the next section, however, the focus is on phonological processing in visual word encoding. As will be seen, this is where most of the research on phonological coding in reading has been done. Moreover, in the following sections on the role of phonological coding in short-term memory and some sort of inner speech in reading, questions about phonological short-term memory and inner speech are often intertwined with questions about phonological processing in visual word encoding.

Phonological Coding in Word Identification in Reading

Perhaps a good place to start discussing this topic is to review the literature on single-word identification tasks. This is because the large majority of studies on phonological coding in word identification come from such studies.

Phonological Encoding of Words Examined in Single-Word Identification Tasks

Most of the literature on the role of phonological coding in lexical access has been on the coding of phonemes. This may be largely due to the fact that the written code in alphabetic languages has largely been devised to convey phonemic information (see Kessler & Treiman, this volume). The key question, of course, is how this phonemic information is obtained from the letters on the page. This was originally framed in terms of a relatively simple dual-route model (e.g., Coltheart, 1981). In this model, one of the routes, *assembled phonology*, used a constructive process that went through the letters of a word from left to right (in English and other European alphabetic languages), and the process was assumed to be mostly context free. Although the conversion process in this assembly was posited to go (in English) more or less from left to right and from letter to letter producing phonemes, it did not do so completely, as larger letter units, *graphemes* (e.g., ‹ch›, ‹th›, ‹ng›) were also posited to have unique mappings to phonemes. The other route, *addressed phonology*, was basically a look-up process in which there is an assumed mental lexicon where the phonological representations of known words are stored that can be quickly accessed (or addressed) when the word is seen. This conception of word processing led to the idea that is widely used in the word recognition literature that words can largely be divided into *regular* words (i.e., those whose pronunciations would be correctly given by applying a set of rules prescribed by some assembly process) and *irregular* words (i.e., whose pronunciations would be incorrectly given by applying the rules).

Controversies arose about whether any set of simple rules exists. However, at an intuitive level, it does seem as though a spelling like ‹bat› is regular in that these letters are pronounced the same way in almost all three-letter words containing them. In contrast, ‹have› seems irregular in that it is unlikely that simple, widely applicable rules of the language would yield the appropriate pronunciation for it (i.e., most one-syllable words ending in ‹ave› are pronounced to rhyme with ‹save›). More generally, according to a dual-route theory, the phonological representation of a regular word can be obtained either through the assembled route or through the addressed route, whereas the correct phonological representation of the irregular word can only be obtained through the addressed route. This leads to the prediction that it will take longer to obtain the correct phonological representation for irregular words than for regular words when using the assembly process, This type of theory makes a second prediction: an interaction of regularity with word frequency. That is, if one assumes that the addressed route is generally fairly fast for high-frequency words, it may usually win the race with the assembled route. Thus, regularity might be largely irrelevant for high-frequency words. On the other hand, for low-frequency words, the addressed and assembled routes may operate at about the same speed. Thus, one would expect to see a larger effect of regularity on the time to access the phonological code for them. In addition, for low-frequency words, there would be a question of which phonological code gets accessed (i.e., the assembled or the addressed code), so there can be appreciable errors for low-frequency words as well.

For the moment, I will continue to assume this regular-irregular dichotomy of words to illustrate the two primary paradigms in which phonological processing has been extensively studied in experiments using single words as stimuli and having subjects make speeded responses to them. These two paradigms are the *speeded naming task,* where naming onset latency is measured, and the *lexical decision task* (Meyer, 1970), where the nonword distractors are usually *pseudowords* (i.e., pronounceable nonwords) (see also Yap & Balota, this volume). One of the primary designs in many of the experiments was to factorially manipulate the regularity and frequency of the words. These studies found the interactive pattern predicted earlier (i.e., little regularity effect for high-frequency words and a large regularity effect for low-frequency words) in the naming task. A similar, although somewhat weaker, pattern was obtained in the lexical decision task. That the pattern is weaker for the lexical decision task might be expected as the task is not as inherently connected to phonology as the naming task is (see Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004).

In these experiments, there are two related but distinct issues: (1) Does early analysis of the written code exemplified by assembled phonology in the original Coltheart (1981) model play an important role in creating a phonological representation of a written word? (2) Does early analysis of the written code that results in producing a phonological code play an important role in accessing the word in the reader's orthographic lexicon and also in accessing its meaning? Presumably the naming task is primarily answering the first question, whereas the lexical decision task is more directed at answering the second question. In either case, these experiments argue that early analysis of the written code appears to be involved in activating the phonological code of a word, which in turn activates its lexical entry and probably its meaning.

So far, I have greatly oversimplified the issues about how phonology is derived from printed words; however, I think the preceding discussion has served to give both the first basic theory that most of the other models were reactions to and also the types of data that researchers collected to test their theories. Much of the ensuing literature could, in some sense, be considered a reaction and response to the relatively simple hypothesis of Coltheart's (1981) original model. One set of later models questioned the whole framework of Coltheart's (1981) model and similar models (e.g., Coltheart, Davelaar, Jonasson, & Besner, 1977). A distinguishing feature of these models was that they had no lexicon—either visual or auditory. Originally, these models were considered single-process models (e.g., Seidenberg, Waters, Barnes, & Tanenhaus, 1984), as they did not have a distinct lexical look-up route to the phonological code. Instead, the models had many low-level subword detectors operating in parallel. These subword detectors (both visual and auditory) did not necessarily correspond to any particular linguistic entities such as phoneme, syllable, or grapheme (although they could). Instead, they were patterns that were learned by the reader over time that are perceptually salient (i.e., either the correlations of the stimuli within the visual or auditory modality or between modalities are high). Newer versions of these models (e.g., Seidenberg & McClelland, 1989; Plaut, McClelland, Seidenberg, & Patterson, 1996) are often called *triangle* models (see Woollams, this volume, Yap & Balota, this volume) as they explicitly represent semantic representations interacting with both orthography and phonology to arrive at a phonological representation of a word or nonword. The reader learns correlations between the visual and auditory representations and these semantic representations. In sum, the essence of these models is the claim that people have not learned rules to convert orthography to phonology. Instead, they have in some sense learned the correlations between the orthographic structure of the writing system and the phonological structure of the spoken language (and possibly the semantics of the spoken language as well in the later models).

Other modelers did not make such a radical break with Coltheart's (1981) original conception. Instead, their ideas of what constituted something like an assembly process differed substantially from a simple rule-based grapheme-phoneme conversion scheme. This was partly because they were skeptical that a rule-based system relying only on graphemes could be very successful for a language like English. This kind of system seems a bit too austere, as it would require one pronunciation for letters such as ‹c› and ‹g› even in a language like Spanish, where the spelling-to-sound mapping is much more intuitively regular than English. Thus some higher-order rules are needed, because in (Latin American) Spanish, ‹c› has the sound /s/ when preceded by ‹e› or ‹i› but has the sound /k/ otherwise. (English is quite similar, though its rules are not quite as regular as those in Spanish.) Thus, even at the grapheme-phoneme level, it would intuitively seem that rules must be sensitive to context to some extent. A major way in which these models differed from Coltheart's original conception is that they emphasized syllables and their subunits, *onsets* and *rimes,* as units of pronunciation (e.g., Treiman & Baron, 1983; Treiman, Kessler, & Bick, 2003) (e.g., for ‹bat›, the onset is /b/ and the rime is /æt/). A related idea is that perhaps classifying words as regular or irregular is not the key comparison of interest, and that instead classifying them as *consistent* or *inconsistent* is of greater relevance (Cortese & Simpson, 2000). This is related to analyzing words at the syllable and onset-rime level, as a pronunciation of a letter cluster would be completely consistent if it led to the same rime pronunciation in every syllable or word in which it appeared (e.g., *bake, cake, lake*). In contrast, a letter cluster like ‹ave› would lead to inconsistent pronunciations at the word/syllable level (e.g., *gave, have, save*). However, the term is generally used statistically in that a word would be classified as consistent if most words sharing its rime were pronounced the same way and as inconsistent if most words sharing its rime were pronounced differently (but similarly to each other). One important piece of data arguing for consistency as a relevant variable is that Treiman et al. (2003) have shown that the pronunciation of a vowel in a naming task can be influenced both by the consistency of the onset that precedes it in the syllable and the coda that follows it (e.g., ‹w› is the onset and ‹n› is the coda in ‹wan›). Note that this is a somewhat more sophisticated definition of consistency in that it not only uses the syllable but also the onset of the syllable.

Although the models that have been discussed differed as to the reason for the effect, the most important empirical finding that emerged from this research is that the time to name a word (and to a lesser extent, the time to judge whether a string of letters is a word or not) can be predicted by the properties of its spelling pattern. In some views it was regularity, in other views it was consistency, and in the third class of models it was largely orthography-phonology correlations that had been learned by the system. This strongly suggests that encoding of print involves an automatic conversion from an orthographic code to a phonological code. (However, it should also be noted that most of this research was limited to studying relatively short monosyllabic words.) Moreover, virtually all of the experiments (e.g., Seidenberg, 1985; Seidenberg et al., 1984) that manipulated a variable (e.g., regularity, consistency) that was designed to demonstrate phonological involvement in naming time for printed words obtained significantly larger phonological effects for low-frequency words than for high-frequency words. However, there is at least one report of a significant consistency effect for high-frequency words (Jared, 2002). On the other hand, as indicated earlier, the consistency effects in lexical decision times were weaker, and were generally only observed for certain kinds of low-frequency words with highly unusual spelling patterns (Seidenberg et al., 1984).

One might think that the fact that phonological effects are small for high-frequency words means that phonological coding of words may be largely restricted to low-frequency words. However, another simple paradigm involving speeded responses to single words argues strongly that the phonological code of a word is intimately involved in activating its meaning for more than low-frequency words. In this paradigm (Van Orden, 1987), subjects were asked to make judgments about whether words are members of a semantic category, such as the category of foods. Surprisingly, the error rate was quite high (about 30%) for homophones of members of the category (e.g., *meet*), but low (about 5%) for control words (*melt*) that also differed by one letter from the true member of the category (*meat*). Correct *no* responses to these pseudomembers of the category were also very long, again indicating that subjects' responses were strongly influenced by the phonological code that had been encoded that they could not ignore. In a follow-up experiment, Van Orden, Johnston, and Hale (1987) found virtually the same sized effect for pseudowords, as when *sute*, for example, was judged falsely as an article of clothing. A major reason this categorical judgment paradigm is of interest is because the effect is so large. The paradigm also strongly argues for an important place for phonological coding in word encoding and reading, as the task it employs is getting at the access of meaning rather than using a judgment about wordness or the sound properties of a letter string. Although one might argue that the response times in this task are a bit slow (they are usually on the order of 500–600 ms) and possibly not related to the rapid processing of words that goes on during silent reading of skilled readers, they are not much slower than lexical decision times when pseudowords are used as foils (which are often close to 500 ms). Importantly, these experiments make clear that people don't go directly from print to the orthographic lexicon; otherwise, subjects wouldn't misclassify fairly common words in the semantic judgment task and even misclassify pseudowords.

A last set of studies with single words that are important in the literature on phonological coding are those with adult populations with severe reading impairments (*acquired dyslexics*) (e.g., Marshall & Newcombe, 1973, 1981). Two groups of these dyslexics are of particular interest, because they appeared to be relatively pure cases of having one or the other the two routes of Coltheart's (1981) dual-route mechanism grossly impaired and the other route relatively intact. One of these groups of patients, *phonological dyslexics*, to some approximation, has the following syndrome. They can basically read almost any word within their current vocabulary correctly but have great trouble reading any new word or pseudoword. That is, it appears that they can find the pronunciation of a letter string if it matches one in their mental lexicon (i.e., by something like Coltheart's [1981] addressed route). If not, they appear to be lost. The other group, *surface dyslexics*, can render a pronunciation for both words and pseudowords. However, the pronunciation for irregular words is often wrong. For example, they will pronounce *island* something like /ɪzlænd/. Similarly, these patients appear to rely completely on something like an assembled route to achieve a pronunciation of a word and the lexical route is rarely (if ever) there to correct them.

Although these data are certainly suggestive of these two separate routes (addressed and assembled) to a mental lexicon, there has been considerable dispute in the literature about whether this is the best way to explain these clinical data. This dispute is tied in with whether the dual-route model is the best way to explain the interactive pattern of frequency and regularity on naming time with normal skilled readers. First, at a local level, the picture presented in the prior paragraph is somewhat of an oversimplification of the two real syndromes. Second, it is important to stress that there are many other populations of acquired dyslexics who do not look at all like these two groups (see Woollams, this volume). Allied with this, there still is considerable controversy about whether something like a dual-route model or a triangle model is the best way to explain these acquired dyslexic syndromes (see Plaut et al., 1996; Woollams, this volume), as well as the best way to explain the interactive pattern of phonological regularity by frequency in naming times in normal skilled readers.

Phonological Encoding of Words Examined in Paradigms Using Eye Movements

So far, this chapter has concentrated on the literature having people either name or make judgments about single words in isolation. These single-word experiments have the advantage of allowing for very tight control over aspects of the target stimuli (i.e., the words of interest) such as word frequency, word length, and regularity, because one can use many different words in a 1-hour experiment. However, if one wants to know the role of sound coding and phonology in reading, then it seems advisable to also study it in paradigms that more closely approximate natural reading. Accordingly, most of the rest of the chapter will discuss studies that rely on using eye-movement technology, the basics of which are covered in Schotter and Rayner (this volume). However, a brief summary of terminology and measures used in eye-movement experiments will be provided here.

In reading, as in viewing any static display, the eyes, to a very good approximation, are still and both are pointing at the same location most of the time. These still, periods are called *fixations*, and in reading they typically last about 200 to 300 ms. The eyes move from fixation to fixation in extremely rapid movements called *saccades*, during which, to all practical purposes, nothing is seen. Most eye movements are forward through the text (and these will be the focus here, unless noted otherwise). The eye-movement measures that are of greatest interest in reading with respect to processing of a target word are the following: (1) *first-fixation duration*, the duration of the first fixation on the target word; (2) *gaze duration*, the duration of all fixations on the target word before the target word is exited to the right; and (3) *total fixation time*, the sum of all fixation durations on the target word after it is first fixated. (These measures are only taken if the word was not skipped when the reader went through the text for the first time, i.e., on the *first pass*.) Other measures that are used are the probability of skipping a word on the first pass through the text, the probability of *regressing* back (i.e., going back to *refixate* a target word), and *saccade length*.

Many of these experiments simply have people read text in which key words or regions of the text have been designated as targets (e.g., a target word is either a homophone or a control word). However, other experiments have used sophisticated display change techniques described in Schotter and Rayner (this volume) in which the display on the computer monitor changes during a saccade between one fixation and the next. This allows one to understand the time course of events during the reading process in great detail. A bonus of many of these techniques is that the reader is unaware of the change, so that one has reasonable confidence that the experimenter has not disturbed the reading process. Ideally, it would be most natural to have people read paragraphs or whole stories to study reading in a completely natural setting, but that is usually impossible given the time constraints of running experiments. Thus, most of studies described in this chapter have people reading a sentence or two per trial.

One of the issues that isn't really resolved by the single-word studies is the time course of phonological encoding (and the time course of processing more generally) in reading. One interpretation of the interactive pattern of regularity and frequency in the naming task (and especially in the lexical decision task) is that phonological coding plays a relatively minimal role in lexical access, and that it is mainly restricted to low-frequency words. The categorization experiments reviewed earlier presented a different picture; however, the response times were relatively long. Reading experiments using eye-movement measures as dependent variables can help to provide a better timeline for the involvement of phonology in the reading process to determine whether phonology enters into the early stages of processing of all words when people read. We begin with a paradigm that controls the visual display in such a way that one can make more precise statements about the timing of phonological processing.

Boundary Studies

One important paradigm employed in reading is the *boundary paradigm* developed by Rayner (1975a, 1975b). This is a special case of the more general parafoveal preview contingent eye-movement paradigm (e.g., Rayner, McConkie, & Ehrlich, 1978). The fovea is the region of clear vision in the retina, which is well defined anatomically (it extends 1 degree in all directions); although there is no clear physical demarcation of the parafovea, the term is generally applied to the next 5 degrees or so on the retina—extending horizontally in both directions. (As this chapter deals with the reading of languages written or printed horizontally, we'll omit the vertical anatomical details of the retina.) In a boundary experiment, there is a selected region—usually one word—that is changed (or not changed) between when it is viewed parafoveally and when it is viewed foveally, and there is only one display change.

There are variations of the technique, but the one we will focus on is the following (see Figure 13.1): There is an invisible *boundary* after a certain location in the sentence, and before the reader's eyes cross the boundary (assuming the reader is going left to right), the text is in one form (see line 1 of Figure 13.1). When the reader's eyes cross this boundary, the text is changed to the second form (see line 2 of Figure 13.1), and it remains that way until the end of the trial, when the subject presses a key to indicate that the sentence has been read. The saccadic eye movements crossing the boundary are extremely rapid and, as nothing is seen during a saccade, the reader is almost always unaware of both the content of the parafoveal preview and that there was any change in the display. Occasionally,

Fixation *n*

```
                 *    |
This bird prefers beach trees for nesting.
```

Fixation *n* + 1

```
                      |  *
This bird prefers beech trees for nesting.
```

Fig. 13.1 Illustration of the boundary technique in reading. The stars on the first and second lines indicate that the subject was fixating on the word *prefers* on fixation *n* and then on *beech* on fixation *n + 1* when the display change was made. The vertical line indicates the position of the (invisible) boundary that triggers the display change when the saccade crosses it. Other previews could be an identical preview (*beech*), an orthographic control (*bench*), or a dissimilar preview (*stock*).

if either the fixation prior to the change or after the change was too near to the boundary (and thus that the change was probably not successfully made during the saccade), the reader may be aware of the change. These trials are usually removed from the analyses (see Schotter & Rayner, this volume, for more details about eye-movement analyses).

As can be seen in Figure 13.1, similar to priming experiments with single words, the design of preview experiments is that there are many possible parafoveal previews of the target word just before the word is fixated. Of greatest interest here is the comparison between a homophonic preview and a word preview that is as orthographically similar to the target as the homophonic preview. The primary finding is that the fixation time on the target word is less in the homophonic preview condition than in the orthographic control condition. It should be noted that "shorter fixation time" on the target word means that (except if stated otherwise) both the first-fixation duration and the gaze duration on the target word are shorter. The E-Z Reader model (e.g., Pollatsek, Reichle, & Rayner, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998) makes a strong argument for why both these measures capture something close to word processing time, or at least that differences in these measures capture differences in word processing time.

These preview experiments thus indicate that phonological coding of words begins even before a word is fixated in text, because the phonological information is being extracted from the preview of the word before it is fixated. However, because the relevant fixation time measures only indicate processing after the word has been fixated, these experiments cannot adjudicate between whether the phonological information was purely extracted from the parafoveal glance or whether it only was used in conjunction with the phonological information extracted from the word when it was later fixated. This has bearing on the issue of whether the phonological coding in the parafovea was going on at a prelexical level. In terms of the models discussed earlier, a prelexical phonological level would be like assembled phonology in a Coltheart (1981) type of model. For the triangle types of models, a prelexical phonological level would be the boxes in the auditory word recognition system between the sensory organ and the response (as there is no auditory lexicon in these models). That is, if the parafoveal preview word has already been fully identified, then one can't rule out the hypothesis that the phonological coding in the parafovea is merely some sort of look-up process. On the other hand, if the data indicate that the preview word is only partially identified, then many of the complexities of phonological encoding that apply to single words in isolation must also apply to words that the reader hasn't yet fixated.

Rayner (1975a) also employed a different boundary paradigm with a single word or pseudoword in the parafovea and just a fixation cross in the fovea. The reader makes a saccade to the parafoveal location and there is the same kind of display change during the saccade as in the boundary experiments illustrated in Figure 13.1. In these experiments, the usual measure of performance, however, is naming time of the target (which is almost always a word). The patterns of results from this paradigm in English mirror those in the reading studies quite closely, including the sizes of the effects. The reason for mentioning this single-word boundary paradigm is that it was the first that was used to study homophonic parafoveal priming effects in Chinese (Pollatsek, Tan, & Rayner, 2000). Significant priming effects were obtained, so that it appears that sound codes are extracted in the parafovea even in a nonalphabetic language. However, the priming effects in Chinese were more equivocal in boundary experiments that used the sentence reading paradigm (the one displayed in Figure 13.1). Liu, Inhoff, Ye, and Wu (2002) did find a benefit from a homophonic parafoveal preview; however, it was not on the fixation time on the target word that was previewed. Instead, the effect was registered as readers having fewer regressions back to the region of the target word. On the other hand, a later experiment (Tsai, Lee, Tzeng, Hung, & Yen, 2004) did find reliable fixation time effects on the target word when the preview character and its radical component were both homophonic to the target character.

There are two possible explanations for this difference between this somewhat apparently delayed or less reliable effect of phonology in Chinese and the more immediate effect in English. The first explanation is a somewhat artifactual one. That is, since Chinese words tend to be short (on average two characters) and since the lines of text are sequences of characters unbroken by any markers to indicate words or phrases, it is highly unlikely that Chinese readers use a strategy for targeting their eye movements that involves targeting an individual word (Wei, Li, & Pollatsek, 2013). The second and possibly more interesting explanation for the differences between the results in the two languages is the nature of the orthography. Chinese orthography does not, in a consistent way, represent the sounds of words, whereas the orthography in alphabetic languages does to a greater extent. Thus, one might expect the sound codes to be more quickly activated in English or other alphabetic languages than in Chinese.

PARAFOVEAL FAST PRIMING

A related paradigm examines slightly later stages of phonological coding in reading: the *parafoveal fast priming paradigm* (Lee, Binder, Kim, Pollatsek, & Rayner, 1999; Lee, Rayner, & Pollatsek, 1999; Rayner, Sereno, Lesch, & Pollatsek, 1995). Here, there are two display changes after the boundary is crossed. Prior to the crossing the boundary, the target region is a nonsense string (a sequence of random letters). Then, when the boundary is crossed, for the first 35 ms the prime is presented, and then it is followed by the target which remains visible for the remainder of the trial (see Figure 13.2). This paradigm is quite like the classic fast priming paradigm devised by Forster and Davis (1984). Forster and Davis's paradigm first displayed something like a row of Xs in the fovea for approximately a second followed by a brief prime and then the target—all in the same location (but the subjects presumably did not move their eyes). As in the other display change experiments described here, subjects are virtually never aware of the first display change but, as in the Forster and Davis paradigm, subjects are often aware of the change from prime to target as something like a flicker. However, they are virtually never aware of identity of the prime. Using the paradigm to assess phonological effects in the parafovea, similar kinds of primes were used as in the parafoveal preview experiments described in the earlier section (e.g., identical,

Fixation *n*

```
                 * |
This bird prefers dilst trees for nesting.
```

First 35 ms on Fixation *n* + 1

```
                   | *
This bird prefers beach trees for nesting.
```

After the First 35 ms on Fixation *n* + 1

```
                   | *
This bird prefers beech trees for nesting.
```

Fig. 13.2 Illustration of the fast priming technique in reading. The stars on the lines indicate that the subject was fixating on the word *prefers* on fixation *n* and then on *beach* on fixation *n* + 1 when the first display change was made. The second display change is made while the subject is fixating on the target word location. The vertical line indicates the position of the (invisible) boundary that triggers the display change when the saccade crosses it. Other primes could be an identical prime (*beech*), an orthographic control prime (*bench*), or a dissimilar prime (*stock*).

homophone, orthographic control, and different). The results were similar, in that fixation times on the target word were shorter when the prime was a homophone of the target word than when it was an orthographically related control prime.

These phonological fast priming experiments indicate that extraction of phonological codes begins within the first 35 ms after a word is fixated in text. The explanation of these priming phenomena is complex. It is not necessarily the case that the prime in these experiments is fully processed in the first 35 ms of the fixation and then facilitates processing of the target. Instead, it is likely that a much more complex interactive process involving partial activation of the letter nodes from both sources could explain the faster processing of the target word in the homophonic prime conditions. Nevertheless, both the parafoveal preview paradigm and the parafoveal fast priming paradigm indicate that activation of phonological codes is an integral part of word encoding that starts early in the reading process. Specifically, this activation starts before a word is fixated and during the first 35 ms of the fixation—even when interrupted by a display change. The figure of 35 ms here is approximate, as the display change is made when the eye crosses the boundary and the fixation begins a few milliseconds after that. It should be noted that, although the fast priming results described earlier were reliably observed over several experiments at approximately 35 ms, they were not observed either at appreciably shorter or longer nominal prime durations. The explanation given by the authors of the fast priming studies (Lee, Binder, et al., 1999; Lee, Rayner, et al., 1999; Rayner et al., 1995) for this is the following: For shorter prime durations, not enough of the prime had been processed to have an effect. For longer prime durations, the prime had been consciously processed on at least some trials, so that there was likely to be a complex mixture of trials with facilitation from the prime when it wasn't consciously processed and inhibition from the prime when it was.

Although the rapid activation of phonological codes during reading that has been demonstrated seems compatible with a view that assembled phonology is an integral part of the reading process, the data presented so far do not make a strong case for it, as there were no reliable pseudoword priming effects. However, a more traditional masked priming experiment by Pollatsek, Perea, and Carreiras (2005) in Spanish makes a good case for early involvement of phonological coding in word identification. Here the pronunciation of the initial consonant of the nonword masked prime was altered by the following vowel. For example, ‹cinal› and ‹conal› were primes for the target ‹canal›. (In Castillian Spanish, ‹c› has a /θ/ sound when followed by ‹e› or ‹i› and a /k/ sound when followed by any other letter, including other vowels.) Lexical decision response times were faster when the initial phonemes of the prime matched. This result indicates that phonological coding of the prime must have been occurring in the 30 or so ms before the target appeared. (There were orthographic control conditions to ensure that the differences were not due to differences between the forms of the vowel letters.)

This particular manipulation has not been done in sentence reading in Spanish, but a similar one has been done by Ashby, Treiman, Kessler, and Rayner (2006) in English using the boundary paradigm in reading. Instead of using the vowel to manipulate the sound of the initial consonant, they used the final consonant or consonants of the syllable (the coda) to manipulate the vowel sound that preceded it. An example would be that for the target word ‹rack›, there would be the two previews, ‹raff› or ‹rall›, where presumably ‹raff› biases the vowel sound to be consistent with the vowel sound in ‹rack›, whereas ‹rall› biases the vowel sound to be as in ‹call›. They obtained reliably shorter fixation times on the target word when the vowel sound that was most likely for the preview agreed with the vowel sound for the target. This is one piece of evidence that this early use of phonological codes is largely an assembly process to help arrive at word identification. The fact that these effects occur across saccades strengthens this conclusion. This result indicates that discrete pieces of visual information are integrated in some format to arrive at the identification of the word. Given that it is unlikely that the raw visual representation of a word lasts across the saccade from the preview to when it is fixated, the integration of the information almost has to be occurring either in an abstract letter format, a phonological format, or both. Given that the phonological format has desirable short-term memory properties (Baddeley, 1979), it would seem implausible that it wasn't involved in the integration process.

Beyond Phonemes

So far, the discussion of phonological encoding in word identification has focused on the phoneme. This is because most of the research has been on the phoneme. There has been some research, however, about whether syllables are involved in the early encoding stages of sound coding in reading. In an attempt to answer this question, Ashby and Rayner (2004) conducted parafoveal preview experiments in which the previews were initial syllables whose length either matched or mismatched the length of the initial syllable of the target word. As in all the experiments described previously, the target word is in a sentence and subjects are asked to read the sentence silently for comprehension and press a key when they have come to the end of the sentence. For example, there would be two target words in the experimental design, ‹magic› or ‹magnet›. The previews for ‹magic› could either be ‹maxxx› or ‹magxx›, and the previews for ‹magnet› would either be ‹maxxxx› or ‹magxxx›. In the first case, the ‹mag› preview can be classified as a mismatch, as the preview has one too many letters for the first syllable, whereas in the second case, the ‹ma› preview would be the mismatch as the preview has one too few letters for the first syllable. Ashby and Rayner (2004) found that fixation times on the target word were less when the preview matched than when it mismatched. A variant of this experiment using a masked priming paradigm with a single word and a lexical decision task (Ashby, 2010) indicated that these syllable-matching effects occurred as early as 100 ms after the onset of the target word in the record of event-related potentials (ERPs). These ERPs are electrical signals measured at various points on the scalp in a situation where subjects can perform laboratory tasks naturally.

Another aspect of phonological coding in word encoding besides the syllable that goes beyond the phoneme is stress. For example, some words are distinguished by having different stress patterns even though they have either identical or very similar phonemes. This is true whether they are semantically and morphologically related (e.g., the noun and verb forms of *increase*) or if their meanings have little to do with each other (e.g., the noun and verb meanings of *entrance*). Stress will be discussed in the next section.

Phonological Coding and Short-Term Memory

We have now established that there is some sort of phonological coding in reading that is involved in the encoding of words. However, phonological coding can serve other useful purposes in the reading process. This section focuses on one important one: aiding short-term memory. As we will see, phonological encoding is invariably intertwined with phonological short-term memory in reading. Considerable work in cognitive psychology has shown that much of what we think of as short-term memory is a phonological store (e.g., Baddeley, 1979). As it is clear that the visual codes of arbitrary symbols such as printed words or letters have short-lived memories (Craik & Lockhart, 1972), it seems reasonable that having the sound of the word in a phonological short-term memory store is a good back-up place to have it preserved. As we will also see, there is quite good evidence that readers do this phonological short-term memory coding of words, even when it may be counterproductive at times.

Heterophones

One paradigm that has been employed to assess the role of phonological short-term memory codes in reading, especially in connected discourse, has been to use words with phonologically ambiguous properties as target words. The paradigm is straightforward: A homographic or a control word is inserted into a sentence as a target word. Then eye-movement measures are recorded—usually both when the reader encounters the target word and at a point in a second sentence at a point of disambiguation that makes clear which meaning of the word was intended. (The second sentence is not presented in the examples given later.) What is of interest here with respect to phonological short-term memory are the later effects, which indicate that the phonological code of the word had been stored in the reader's short-term memory.

A key experiment by Folk and Morris (1995) compared several types of potentially ambiguous target words (each with its own control word). Of central interest here is the comparison between a homographic heterophone such as *tear* in Sentence (1a) (i.e., a word with two distinct meanings and two different pronunciations) and a homographic homophone such as *calf* in Sentence (2a) (i.e., a word with two distinct meanings but one pronunciation). The ambiguous word in the experiment was presented in the first sentence that is shown, and its meaning was only disambiguated in a second sentence that is not shown. There were also control sentences exemplified by sentences (1b) and (2b), which had phonologically unambiguous synonyms substituted for the key target words. Obviously the design was a counterbalanced one, where no subject saw both an (a) and a (b) version with the same sentence frame. (It also might be worth mentioning here that in this experiment, as in most of the eye-movement experiments described in the chapter, that there is no display change; the subjects simply read the text for comprehension as their eye movements were monitored.) The important question is what effect the ambiguity of the heterophone *tear* had on the reading process above and beyond the semantic ambiguity for the homophone *calf*. The answer is that the effect was quite large. Whereas there was literally no cost due to the meaning ambiguity (relative to the control condition) in gaze duration on the target word for the homophones, there was an 81-ms cost (in gaze duration on the target word) in the heterophone condition! The effect was this large in spite of the fact that the prior context was compatible with either meaning of the heterophone or homophone, so that readers, by guessing, had a 50% chance of coming up with the correct sound and meaning of the word. (Actually, the experimenters chose the more frequent meaning of the word so that readers were biased in favor of initially coming up with the correct meaning/sound of the target word.) Moreover, when readers fixated the word that disambiguated the meaning of the target word, there were twice as many regressions back to the target word region in the heterophone condition as in the control condition, whereas there was virtually no difference in the number of regressions between the homophone condition and its control condition.

(1a) There was a tear in Jim's shirt after he caught his sleeve on a thorn. (Heterophone)

(1b) There was a hole in Jim's shirt after he caught his sleeve on a thorn. (Heterophone Control)

(2a) When Ann hurt her calf after she fell down the stairs, she cried. (Homophone)

(2b) When Ann hurt her shin after she fell down the stairs, she cried. (Homophone Control)

Folk and Morris's (1995) finding indicates that both representations of the heterophone—involving phonological and semantic codes—are activated initially at least some of the time, and the conflict between them slows the reading process. The conflict between the phonological codes is over and above the conflict that occurs between the two semantic codes in the homographic homophone condition. In addition, the need for regressions back to the target word indicates that the wrong phonological coding persists at least some of the time, necessitating a need to try to repair the damage. This latter finding indicates that the phonological codes play a fairly major role in maintaining information in short-term memory, and that readers use this memory to help tie the meaning of the text together.

Breen and Clifton (2011, 2013) provided similar evidence for the use of stress information. They employed ambiguous noun-verb words such as *abstract* as in sentences (3a) and (3b), with a later region of the sentence that disambiguated which meaning was possible in the sentence. (In English, the stress is on the first syllable for the noun version for these pairs and on the second syllable for the verb version.) Sentences (3c) and (3d) are analogous noun-verb pairs such as *report*, except that these noun-verb pairs have the same stress pattern. Moreover, the preceding adjective in these sentences strongly biased the word to be interpreted as the noun meaning. In spite of the prior semantic bias, fixation times on the words with ambiguous stress were about 25 ms longer than on the ones with unambiguous stress even though there was the same semantic ambiguity in the two types of sentences. Thus, as in Folk and Morris (1995), having phonological ambiguity in encoding a word causes lengthening of encoding time beyond having semantic or syntactic ambiguity. The effect on the target word in the Breen and Clifton study may have been appreciably smaller than in the Folk and Morris study because the difference between the two meanings of the words in Breen and Clifton was much less than in the Folk and Morris. (In the latter study, the two word meanings were usually completely unrelated.)

(3a) The brilliant *abstract* |the best ideas from the things they read.

(3b) The brilliant *abstract* |was accepted at the prestigious conference.

(3c) The brilliant *report* |the best ideas from the things they read.

(3d) The brilliant *report* |was accepted at the prestigious conference.

As in the Folk and Morris (1995) study, Breen and Clifton (2011, 2013) examined the later cost of the effect of the wrong stress. Here again, one would expect that having the wrong sound of the noun meaning of *abstract* in sentence (3a) with stress on the first syllable would induce a greater cost when it was clear the verb meaning was correct (with stress on the second syllable) than the cost needed to change the interpretation from noun to verb in sentence (3c), where there would be no difference in stress or sound between noun and verb. This is indeed what was found in both reading times in the region that disambiguated the meaning of the ambiguous word and in the number of regressions back to the target ambiguous word.

Spoken-Word Priming

Another paradigm that indicates that phonological coding can interfere with reading is the *spoken-word priming* paradigm (Inhoff, Connine, Eiter, Radach, & Helter, 2004; Inhoff, Connine, & Radach, 2002). In this paradigm, the subject reads a sentence silently while his or her eyes are being monitored, and when a target word is fixated, a spoken word is played. The word is either identical to, phonologically similar to, or different from the target word. Inhoff and colleagues found that fixation times on the target word were less when the spoken word was identical to the target word than in the other two conditions, which did not differ from each other. Thus, hearing the actual sound strengthened the phonological code. What is perhaps more interesting is that there were interference effects after fixating the target word. For example, readers had longer fixation durations on the post-target word in the phonologically similar condition than in the phonologically dissimilar condition. It thus seems that the wrong phonological auditory prime could not be dismissed in the phonologically similar condition, and it led to an uncertain representation in phonological short-term memory that had to be dealt with (see also Eiter and Inhoff, 2010).

Tongue-Twisters

Another type of experiment that has been used to demonstrate reliance on inner speech has used some variant of tongue-twisters—but in an experiment in silent reading rather than in spoken language. That is, the experiments have people read materials that are putatively difficult to pronounce and determine whether this also makes reading of the materials more difficult. One variation of the paradigm is to use a phonemic sequence over and over again in successive words in a sentence, as in *Crude rude Jude ate stewed prunes*. Another is to use words with similar spellings but mismatching pronunciations adjacent to each other in sentences, such as *nasty* and *hasty*. In both cases, reading times on the sentences—and on target regions in the sentences—are slower than on control sentences (Ayres, 1984; Haber & Haber, 1982; McCutchen & Perfetti, 1982). There is related evidence from reading studies using repeated initial consonants on words, which is a mild form of a tongue-twister. Kennison and colleagues (Kennison, 2004; Kennison, Sieck, & Briesch, 2003) used a self-paced reading method in which readers pressed a key each time they wanted to see the next phrase displayed, and found a slowing in reading due to the repeated initial consonants—but it was delayed. It was only registered in key presses requesting to see the prior phrase. (This somewhat corresponds to regressions in an eye-movement record.) Such regression effects are usually not interpreted as effects on the difficulty of initial encoding processes; instead, they are thought to involve reanalyses of the text, which would involve holding the text in auditory short-term memory. However, in a later study using eye movements, Warren and Morris (2009) had people read similar tongue-twister materials and found both immediate effects due to the tongue-twisters (i.e., longer fixation times on the critical words) and longer sentence-reading times, which involved regressions back to the critical word region. One may worry that these tongue-twister materials are somewhat artificial, and that after a while people may not be really reading them for meaning but viewing the experiment as a game. This problem may be minimized by including many filler sentences in the experiment, so that the experiment is less obviously about these tongue-twisters—nonetheless, the problem remains.

Pauses in Text Coded by Inner Speech

In spoken language, pauses between segments of speech (usually between words) convey a type of stress. Pauses not only help to stress certain syllables but also mark phrase and clause boundaries, mark sentences as questions, and convey emotion and serve other functions. It is beyond the bounds of this chapter to cover all of these topics. However, the use of pauses to mark aspects of the spoken language such as phrase and clause boundaries strongly suggests that another aspect of phonological coding is to preserve these aspects of spoken language in the mind of the reader. This obviously also serves an important short-term memory function. As it seems clear that the written language only provides hints about how stress should be indicated in a spoken sentence (consider how different actors deliver a speech from a Shakespeare play), this discussion also suggests that inner speech is not a passive storage device for sounds but something like a recreation of the spoken language.

One ubiquitous feature of written language related to pausing in the spoken language is the comma, and the obvious question is how readers' eye-movement behavior in the text in the region of the comma can be interpreted in terms of an ongoing inner-speech process. However, commas can serve functions other than to mark pauses in the spoken discourse, such as enumeration (e.g., "red, green, and blue"). The presence of commas in the text facilitates reading even when the parsing of the phrases or clauses in the text is unambiguous (Hirotani, Frazier, & Rayner, 2006). For example, in sentences like (4a) and (4b), Staub (2007) found that reading times for the sentences as a whole were shorter when commas were present. On the other hand, there were slightly longer fixation times in the regions of the sentences when there was a comma than when there was no comma, possibly due to the extra character.

(4a) When the dog arrived(,) the vet and his assistant went home.

(4b) When the dog arrived at the clinic(,) the vet and his assistant went home.

These findings of Hirotani et al. (2006) and Staub (2007) are of interest for two reasons. First, the comma is logically redundant, as the clause boundary is clear from the sentence structure. Second, the facilitation effect of the comma was the same size for the longer (4b) sentences even though the comma was even more redundant. Thus, it appears that readers use commas to add pauses in their inner speech soon after they see them in the text. Although this act of covertly adding the pause to the phonological code or inner speech is slightly time consuming, putting these pauses in to mark the phrase boundaries aids the overall reading process. This is another indication that inner speech is not an epiphenomenon, but a functional part of reading.

Relation of Inner Speech to Outer Speech

This leaves the difficult questions of what exactly inner speech is and what its relation to outer speech is. Part of the problem is that there are so many things that one could mean by inner speech, as I hope my experiment at the beginning of the chapter illustrated. At one extreme it could mean something like the voice in your head, which is close to what people mean by phonological coding. However, perhaps the more usual meaning of inner speech is some version of *subvocalization*, which could involve actual movements in the vocal tract and possibly the mouth (perhaps something like my whispering condition). Inner speech, however, usually implies that its output may not be audible to any listener besides the reader. The meanng of inner speech that people usually think about (subvocalization) would probably have negative consequences on reading in that it would slow the reading process considerably if it went on all the time. An unanswered question is whether there really is any real inner speech that doesn't involve actual movements of the vocal tract.

Subvocalization

One of the major paradigms for studying subvocalization is using electromyographic (EMG) recording. This is usually done either by putting electrodes on an organ related to speech (e.g., lips, tongue, throat, larynx) or placing a needle electrode inside a muscle that is related to speech. This recording provides a continuous record of the electrical activity in the muscle of interest that is usually accurate on the order of milliseconds. The difference in activity in these muscles during reading and some nonreading activity is then compared to the difference in activity during reading and some nonreading activity between muscles that are not plausibly related to reading, such as forearm muscles.

The usual question that has been studied using this technology is whether skilled readers use subvocalization when they are (apparently) silently reading. Researchers generally agree that the answer to the question is that they do subvocalize when they are apparently silently reading (see a review by McGuigan, 1970). The two main issues that have dominated this area are the following. The first issue is whether subvocalization is a necessary part of skilled reading. That is, one view is that subvocalization may be something like a bad habit that is either irrelevant to skilled reading or may even slow it down without aiding comprehension, whereas the opposite view is that subvocalization (which could be construed as some people's view of phonological coding) is a necessary functional part of skilled reading. The second issue is whether subvocalization for skilled readers goes on at a constant rate or mainly at certain specifiable places. That is, perhaps skilled readers chiefly subvocalize only when the text is difficult, or when stress is needed, or to mark phrase boundaries, or at other specifiable locations or for certain types of reading materials.

The paradigm that has been employed to test whether subvocalization is a necessary part of skilled reading is to monitor subvocalization with EMG equipment while people read in a preliminary phase of the experiment. Then, in the training phase of the experiment, subvocalization is monitored as they read and they are given feedback (usually a tone) which is a signal to suppress their subvocalization. Finally, in the test phase, their reading performance is assessed compared either with a control group or with their reading performance in the preliminary phase of the experiment. McGuigan (1970) reported that the general pattern of results from this literature is reasonably consistent. The training does work in that people can, to a great extent, suppress their subvocalization in the experimental session. However, the effects of the training generally do not last very long. The effects of suppression on reading and comprehension were examined in what is perhaps the best-known subvocalization study (Hardyck & Petrinovich, 1970), and the outcome was relatively clear. For the easy reading material there was little decrement in comprehension due to suppression, but for the difficult reading material there was a significant decrement in comprehension. (Comprehension was measured by objective questions asking about details in the texts.)

In summary, the subvocalization suppression studies showed that (1) skilled readers do appear to subvocalize some of the time, at least as a habit, but (2) perhaps only when the material became difficult. It is still an open question what the subvocalization process is that is being suppressed here by this technique. Is it one that involves actual motor movements, such as my whispering condition, or does it only involve suppression of more covert vocalization? This question is far from answered. Another problem with drawing firm conclusions from these suppression studies is that the decrement in performance caused by making people suppress their subvocalization might not be due to suppressing subvocalization but to a general task competition effect. That is, the subjects in these experiments were concentrating on suppressing a behavior that they have done for many years and that had become largely automatized. Thus, the act of suppression itself was likely to demand attention and may have taken cognitive resources away from the reading task. The fact that the suppression effect mainly appeared for the difficult text is compatible with the hypothesis that subvocalization (and phonological coding) was mainly involved in difficult places in the text, and perhaps the difficult text had more phrase boundaries and more difficult phrasing. However, the pattern of data is equally compatible with the hypothesis that reading easy text would not suffer as much from having to concentrate on not having to subvocalize (or in general, not do another task). (See Garrity, 1977, for a review of the relevant literature and a more complete version of the task competition argument.)

As indicated at the beginning of this section, the concepts of phonological coding and inner speech do not appear to be logically identical: The former only implies that some mental representation of the phonology of the text has been involved in the reading process, whereas the latter suggests that more mental representations related to the speech system (and possibly small motor movements) have been involved. However, it does seem that one can separate speech from the voice in your head. To see this, try saying something over and over to yourself out loud such as "blah blah blah…" while reading text (or even poetry), and you will find that you can still hear the little voice in your head. However, it is not clear that anyone so far has successfully been able to clearly disentangle the two concepts (i.e., demonstrate that phonological coding occurs without any involvement—either overt or covert—from the speech system). Nonetheless, as this chapter has hopefully clearly demonstrated, phonological coding is an important activity that occurs during reading, and the focus of the concluding part of this chapter will be to try to further characterize what the nature of these codes is.

Back to Phonological Coding

It has often been argued that phonological coding is not a full code of the spoken language that it represents. That is, if normal skilled adult reading rates are about double those of normal speaking rates, it could be argued that is implausible that phonological coding can be using all the information that is in normal spoken language. Although this argument is reasonable, it is not necessarily true. Thus it is worth discussing whether there is really any solid evidence for it and, if there is, what types of information would be left out that could help to make the phonological code speedier or more efficient.

One proposal for such efficiency has been put forth by Perfetti and McCutchen (1982; see also McCutchen & Perfetti, 1982). They proposed that the phonological code for a word is biased heavily toward having the word's initial phoneme together with the word's general phonetic contour (which mainly has the distinctive features of the remaining consonants of the word)—but it has little vowel information. Obviously, leaving out much of the vowel information would help to make the phonological code a kind of shorthand, as vowel durations are considerably longer than most consonant durations in the spoken language. Perfetti and McCutchen (1982) also posited that little information about function words (i.e., articles, prepositions, pronouns) is preserved in the phonological code. Part of their proposal seems consistent with certain writing systems such as Arabic and Hebrew that do not represent many vowels. Nonetheless, although there may be less information in vowels than in consonants in most written languages, it is hardly the case that there is no vowel information, especially in shorter words (e.g., *bat, bit, bet, beet, bait, boat*). Thus it seems implausible that, a priori, the reader would decide to leave the vowel information out of the ongoing phonological code. It is possible that the reader could, for example, decide to leave in the vowel information for shorter words but use a phonological code of the type posited by Perfetti and McCutchen (1982) for longer words. However, such an active decision mechanism seems unlikely. The hypothesis that function words are not fully represented in phonological short-term memory seems more plausible, but to date there is little firm evidence for it.

This leads to the question of whether there is any other evidence besides the homophone effects discussed earlier that phonological coding time directly influences the efficiency of the reading process. One piece of evidence comes from a lexical decision study by Abramson and Goldinger (1997), who compared words such as *game*, which has a phonetically long vowel because it precedes a voiced consonant, with words such as *tape*, which has a phonetically short vowel because it precedes a voiceless consonant. They found significantly longer lexical decision times in the former case and concluded that this was because there was an inner-speech intermediary in the lexical decision task that took longer to form. Lexical decision times are not necessarily good indicators of phonological coding time in reading. A more direct measure would come from eye-movement measures. One question is whether the syllabic properties of a word predict the time that readers spend on a word in reading sentences. The data are mixed. Ashby and Clifton (2005) found that fixation times were no longer on two-syllable words than on one-syllable words that were matched on number of letters, even though the two-syllable words took longer to pronounce. However, words with two stressed syllables (e.g., *RAdiAtion*) took longer to pronounce than words with only one stressed syllable (e.g., *geOMetry*), and there were reliably longer gaze durations on the former. The difference between the two-stressed syllable words and the one-stressed syllable words was not significant for first-fixation durations on the words, but because they are long words, there are many reasons for this (including that the whole word may not have been processed on many of the first fixations).

In sum, there is only a little evidence for relationships between phonological characteristics of the text and the speed of reading other than at the phoneme level. First, there is the evidence that reading slows when commas are inserted, even if they are redundant, presumably signaling a greater break or pause in the inner speech stream (Hirotani et al., 2006; Staub, 2007). Second, there is the evidence that words with two stressed syllables receive longer gaze durations than words with one stressed syllable (Ashby & Clifton, 2005). Nonetheless, there are puzzling null effects, such as why fixation times on two-syllable words are not longer than those on one-syllable words even though they take longer to pronounce (Ashby & Clifton, 2005). Clearly, we are far from a complete understanding of the nature of phonological coding and its relation to outer speech.

Summary

Although there is much to be learned both about phonological coding and its relation to subvocalization in skilled reading, there are also some clear findings that are quite important, both to guide future research in adult reading and to guide thinking about reading in children. Perhaps the most important finding is that phonological coding is a ubiquitous phenomenon and is almost certainly an important and functional part of the reading process. (This is somewhat contrary to the literature on subvocalization suppression, which suggested that phonological coding—or at least subvocalization—was only needed for difficult passages of text.) As a consequence, teachers of reading should be informed that what appears to be silent reading is far from silent. That is, because skilled readers are in some sense sounding out words as they read, learning to sound out words is clearly part of learning to read. Sound coding is part of the word—encoding process and also probably serves a short-term memory role. The use of eye-movement techniques helps to pinpoint the immediacy of phonological involvement in the silent reading process. Among other things, eye-movement techniques have shown that the phonological encoding process begins early in the word identification process: as it is engaged before readers fixate a word.

References

Abramson, M., & Goldinger, S. D. (1997). What the reader's eye tells the mind's ear: Silent reading activates inner speech. *Perception & Psychophysics, 59,* 1059–1068.

Ashby, J. (2010) Phonology is fundamental in skilled reading: Evidence from ERPs. *Psychonomic Bulletin & Review*, *17,* 95–100.

Ashby, J., & Clifton, C., Jr. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition*, *96,* B89–B100.

Ashby, J., & Rayner, K. (2004). Representing syllable information in word recognition during silent reading: Evidence from eye movements. *Language and Cognitive Processes*, *19,* 391–426.

Ashby, J., Treiman, R., Kessler, B., & Rayner, K. (2006). Vowel processing during silent reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32,* 416–424.

Ayres, T. J. (1984). Silent reading time for tongue-twister paragraphs. *American Journal of Psychology*, *97,* 605–609.

Baddeley, A. D. (1979). Working memory and reading. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Processing of visible language* (Vol. 1, pp. 355–370). New York, NY: Plenum.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133,* 283–316.

Breen, M., & Clifton, C. Jr. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of Memory and Language*, *64,* 153–170.

Breen, M., & Clifton, C., Jr. (2013). Stress matters revisited: A boundary change experiment. *Quarterly Journal of Experimental Psychology*, *66,* 1896–1909.

Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language*, *15,* 245–286.

Coltheart, M., Davelaar, E., Jonasson, J., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance* Vol. 6, (pp. 535–555). Hillsdale, NJ: Erlbaum.

Cortese, M. J., & Simpson, G. B. (2000). Regularity effects in word naming: What are they? *Memory & Cognition*, *28,* 1269–1276.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11,* 671–684.

Eiter, B. M., & Inhoff, A. W. (2010). Visual word recognition in reading is followed by subvocal articulation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36,* 457–470.

Folk, J. R., & Morris, R. K. (1995). Multiple lexical codes in reading: Evidence from eye movements, naming time, and oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21,* 1412–1429.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10,* 680–698.

Garrity, L. I. (1977). Electromyography: A review of the current status of subvocal speech research. *Memory & Cognition*, *5,* 615–622.

Haber, R. N., & Haber, L. R. (1982). Does silent reading involve articulation? Evidence from tongue-twisters. *American Journal of Psychology*, *95,* 409–419.

Hardyck, C. D., & Petrinovich, L. F. (1970). Subvocal speech and comprehension level as a function of the difficulty level of reading material. *Journal of Verbal Learning and Verbal Behavior*, *9,* 647–652.

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language*, *54,* 425–443.

Inhoff, A. W., Connine, C., Eiter, B., Radach, R., & Heller, D. (2004). Phonological representation of words in working memory during sentence reading. *Psychonomic Bulletin & Review*, *11,* 320–325.

Inhoff, A. W., Connine, C., & Radach, R. (2002). A contingent speech technique in eye movement research on reading. *Behavior Research Methods, Instruments, & Computers*, *3,* 471–480.

Jared, D. (2002). Spelling-sound consistency and regularity effects in word naming. *Journal of Memory and Language, 46,* 723–750.

Kennison, S. M. (2004). The effect of phonemic repetition on syntactic ambiguity resolution: Implications for models of working memory. *Journal of Psycholinguistic Research*, *33,* 493–516.

Kennison, S. M., Sieck, J. P., & Briesch, K. A. (2003). Evidence for a late-occurring effect of phoneme repetition in silent reading. *Journal of Psycholinguistic Research*, *32,* 297–312.

Lee, Y., Binder, K. S., Kim, J., Pollatsek, A., & Rayner, K. (1999). Activation of phonological codes during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *25,* 948–964.

Lee, H., Rayner, K., & Pollatsek, A. (1999). The time course of phonological, semantic, and orthographic coding in reading: Evidence from the fast priming technique. *Psychonomic Bulletin & Review*, *5,* 624–634.

Liu, W., Inhoff, A. W., Ye, Y., & Wu, C. (2002). Use of parafoveally visible characters during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance*, *28,* 1213–1227.

Marshall, J. C., & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research, 2,* 175–199.

Marshall, J. C., & Newcombe, F. (1981). Lexical access: A perspective from pathology. *Cognition, 10,* 209–214.

McCutchen, D., & Perfetti, C. A. (1982). The visual tongue-twister effect: Phonological activation in silent reading. *Journal of Verbal Learning and Verbal Behavior*, *21,* 672–687.

McGuigan, F. J. (1970). Covert oral behavior during the silent performance of language tasks. *Psychological Bulletin*, *74,* 309–326.

Meyer, D. E. (1970). On the representation and retrieval of stored semantic information. *Cognitive Psychology*, *1,* 242–300.

Perfetti, C. A., & McCutchen, D. (1982). Speech processes in reading. In N. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 7, pp. 237–269). New York, NY: Academic Press.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Pollatsek, A., Perea, M., & Carreiras, M. (2005). Does *conal* prime CANAL more than *cinal?* Masked phonological priming effects with the lexical decision task. *Memory & Cognition*, *33,* 557–565.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology,* 52, 1–52.

Pollatsek, A., Tan, L.-H., & Rayner, K. (2000). The role of phonological codes in integrating information across saccadic eye movements in Chinese character identification. *Journal of Experimental Psychology: Human Perception and Performance*, *26,* 607–633.

Rayner, K. (1975a). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7,* 65–81.

Rayner, K. (1975b). Parafoveal identification during a fixation in reading. *Acta Psychologica*, *39,* 271–282.

Rayner, K., McConkie, G. W., & Ehrlich, S. F. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and Performance*, *4,* 529–544.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C., Jr. (2012). *Psychology of reading* (2nd ed.). New York, NY: Taylor and Francis.

Rayner, K., Sereno, S. C., Lesch, M. F., & Pollatsek, A. (1995). Phonological codes are automatically activated during reading: Evidence from an eye movement priming paradigm. *Psychological Science*, 26–32.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Towards a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, *19,* 1–30.
Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.
Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior*, *23,* 383–404.
Staub, A. (2007). The parser doesn't ignore intransitivity, after all. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33,* 550–569.
Treiman, R., & Baron, J. (1983). Individual differences in spelling: The Phoenician-Chinese distinction. *Topics in Learning & Learning Disabilities*, *3,* 33–40.
Treiman, R., Kessler, B., & Bick, S. (2003). Influence of consonantal context on the pronunciation of vowels: A comparison of human readers and computational models. *Cognition, 88,* 49–78.
Tsai, J., Lee, C., Tzeng, O. J. L., Hung, D. L., & Yen, N. (2004). Use of phonological codes for Chinese characters: Evidence from processing of parafoveal preview when reading sentences. *Brain and Language*, *91,* 235–244.
Van Orden, G. C. (1987). A rose is a rows: Spelling, sound, and reading. *Memory & Cognition*, *15,* 181–198.
Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1987). Word identification in reading proceeds from spelling for sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14,* 371–386.
Warren, S., & Morris, R. K. (2009, August). *Phonological similarity effects in reading*. Paper presented at the meeting of the European Conference on Eye Movements, Southampton, England.
Wei, W., Li, X., & Pollatsek, A. (2013). Properties of fixated words affect outgoing saccade length in Chinese reading. *Vision Research*, *22,* 1–6.

CHAPTER

14 Reading Sentences: Syntactic Parsing and Semantic Interpretation

Adrian Staub

Abstract

Understanding what we read requires constructing a representation of each sentence's syntactic structure and meaning. We are generally not consciously aware of the complex cognitive operations that underlie these processes. However, researchers have been able to learn a great deal about these operations by using methods such as eye tracking during reading and recording of event-related potentials (ERPs). This chapter provides a broad overview of some of the main theoretical issues and empirical conclusions in this area, focusing on the following questions: How rapidly do we perform syntactic and semantic analysis in the course of reading? Do we construct more than one syntactic analysis at a time? Does syntactic analysis always precede semantic interpretation? What role does memory play in sentence processing? Finally, how do these sentence-level processes relate to the process of visual word recognition itself?

Key Words: sentence processing, parsing, semantic interpretation, eye movements, event-related potentials

Much of the research reviewed in this handbook is about how readers recognize words. But recognizing each word is only part of what readers must do. They must also analyze the grammatical structure of each sentence, and must identify each word's place in this structure. This process is known as *syntactic parsing*. In addition, readers must combine the meaning of individual words and phrases to construct a representation of the sentence's overall meaning, determining what the various expressions in the sentence refer to, and who did what to whom. We will call this process *semantic interpretation*.

In general, parsing and interpretation are quick, automatic, and effortless in both reading and spoken language processing. Indeed, we are usually not consciously aware that these processes are even taking place. However, it is easy to construct cases in which this effortless sentence processing breaks down. These special cases have provided some of the most interesting evidence about how the system works. A useful analogy is found in visual illusions. Cases where visual perception breaks down have proven to be informative about the architecture of the visual system, and similarly, cases where sentence processing breaks down have proven to be informative about the human sentence processing mechanism.

Consider sentence (1), which is probably the most famous example in the history of psycholinguistics:

(1) The horse raced past the barn fell.
(Bever, 1970)

If you've never seen this sentence before, it probably seems to you that it is not a grammatical sentence of English. But in fact, it is perfectly grammatical. To see this, consider the following two sentences that are similar to (1) in their structure:

(2) The story told by the scout leader scared everyone.
(3) The toy bought at the store broke.

You probably found these two sentences much easier to understand. Now try to impose on sentence (1) the structure of (2) and (3). Do you understand it now?

According to most theories, the trouble with (1) is as follows. The syntactic analysis of the verb *raced* is initially ambiguous: It can be a main verb, as in *The horse raced down the track*, or it can be the beginning of a structure called a reduced relative clause, which modifies *horse*. The parser tends to analyze *raced* as a main verb; the question of why it does so has been the subject of extensive research and theorizing. But once the parser has adopted this analysis, there's no way to attach the word *fell* into the sentence. At this point, the parser seems to have trouble going back and revising its initial analysis in favor of the relative clause analysis. Again, the question of why this reanalysis is so hard has been much discussed (Fodor & Ferreira, 1998). In (2) and (3), by contrast, either the relative clause analysis is adopted in the first place (Trueswell, Tanenhaus, & Garnsey, 1994), or it is easier to recover that analysis after an initial main verb analysis than it is in (1) (Clifton et al., 2003).

Sentence (1) is an especially vexing example of what is known as a *garden path* sentence. The metaphor comes from the fact that the parser seems to be "led down the garden path" to an attractive but incorrect syntactic analysis of the sentence. But in fact, syntactic ambiguity is everywhere. Sentence (4), for example, is globally ambiguous: It has at least five different possible syntactic analyses that correspond to five different meanings. Can you identify them?

(4) Put the block in the box on the table in the kitchen. (Church & Patil, 1982)

Even in apparently simple sentences, the parser must make decisions about which analysis to adopt.

Now consider sentences (5a and 5b):

(5)

a. The doctor that the nurse that the administrator knows hates resigned.

b. The doctor that everyone I know hates resigned.

If you are like most people, sentence 5a is extremely difficult to understand—maybe even impossible. But in fact, this too is a perfectly grammatical sentence of English. You can assure yourself of this by considering sentence 5b, which is like 5a in terms of its structure, but is easier to understand. Unlike in (1), the difficulty with (5a) does not seem to have anything to do with ambiguity. Rather, there is something about sentences like 5a, which are called *double center-embedded* sentences, that overwhelms a resource or capacity involved in sentence processing (e.g., Gibson, 1998). The substitution of different kinds of noun phrases in 5b (*everyone* in place of *the nurse*, and *I* in place of *the administrator*) somehow resolves this problem (Warren & Gibson, 2002).

It is easy to see from these cases that complex cognitive mechanisms are involved in understanding sentences during reading, even if we are not usually aware of their operation. This chapter will discuss several issues that are at the forefront of research on these mechanisms. At the outset, it is important to note some limitations on this chapter's scope. We are concerned with how sentence processing works in fluent adult readers; we will not consider sentence processing in children or in people with acquired or developmental language disorders. In fact, there is notable work on both of these topics (see Snedeker, 2013, and Caplan, 2013, for reviews). In addition, we will primarily discuss research on readers of English and other European languages. The vast majority of sentence processing research is carried out in these languages, although interesting research comparing sentence processing across languages has been emerging (e.g., Bornkessel & Schlesewsky, 2006). Because the focus of this handbook is reading, we will primarily discuss research in which sentences are presented to subjects in written rather than spoken form, even though many of the same issues arise in the two modalities. Finally, the only research method from cognitive neuroscience that is discussed here is the event-related potential (ERP) method, as fMRI and magnetoencephalography, or MEG, have just begun to make substantial contributions to our understanding of online sentence processing.

The topics covered in this chapter are as follows. First, we consider the question of how rapidly we perform syntactic and semantic analysis in the course of reading. Do we compute a syntactic parse and a semantic interpretation on a word-by-word basis, or is there some delay between reading a word and assigning it a place in the sentence's syntactic structure and meaning? Second, we consider the question of whether a reader entertains only one analysis of a sentence at a time or multiple analyses at once. Third, we ask how syntactic and semantic analyses are related to each other. Does parsing come first, or might some kinds of semantic interpretation proceed in synchrony with, or even in advance of, a full syntactic analysis? Fourth, we address the role played by memory in sentence processing. Finally, we consider the question of how sentence-level processing in reading is related to visual word recognition itself. Does the visual word recognition system pass its output to a separate system (or systems) devoted to sentence-level processing, or is there interaction between these levels, such that word recognition may be affected by the syntactic or semantic context in which a word appears?

Incrementality in Sentence Processing

What evidence is there that readers actually do construct a syntactic parse and a semantic interpretation incrementally—that is, as each word is read, as opposed to waiting for the end of a clause or a sentence? An influential early proposal by Fodor, Bever, and Garrett (1974; see also Marcus, 1980) suggested that parsing and interpretation decisions are made only on a clause-by-clause basis, not on a word-by-word basis, with the sentence processing system delaying decisions in cases of ambiguity until the end of a clause. But the last few decades of research have established beyond a doubt that this view is not correct. Intuitive evidence has already been provided, as our explanation of the difficulty of sentence (1) suggested that you were committed to a specific, incorrect analysis before encountering the word *fell*. More convincingly, there is abundant evidence of such incremental processing from behavioral and electrophysiological experiments. It appears that there is no measurable lag between recognizing a word and attempting to integrate it into a sentence-level syntactic and semantic representation.

Perhaps the best evidence for this conclusion comes from studies in which readers' eye movements are monitored. In a seminal study, Frazier and Rayner (1982) presented readers with sentences like (6a–b), among others:

(6)

a. Since Jay always jogs a mile this seems like a short distance to him.

b. Since Jay always jogs a mile seems like a short distance to him.

Frazier's (1978, 1987) garden path theory, one of the first explicit theories of the operation of the parser, predicted that readers should apply a structural principle that results in initially attaching the phrase *a mile* as the object of *jogs* rather than as the subject of the main clause of the sentence. In 6a, this analysis turns out to be correct, but in 6b, this analysis turns out to be wrong. Frazier and Rayner found that immediately upon encountering the word *seems* in 6b, which is the point at which the object analysis of *a mile* is ruled out, readers' eye fixations increased in their duration, and readers were also more likely to make a regressive (i.e., leftward) saccadic eye movement to an earlier region of the sentence. This result suggests that by the time readers had encountered the word *seems*, they had already committed to one analysis of the sentence, which in the case of 6b happened to be the incorrect one. This basic pattern, of an immediate increase in fixation durations or an increase in regressive eye movements upon encountering material that rules out an attractive initial syntactic analysis, has since been obtained in dozens of studies (see Clifton, Staub, & Rayner, 2007, for a review).

A sizable literature has also directly investigated the time course with which the plausibility of a word as a sentence continuation affects eye movements in reading, in the absence of syntactic ambiguity (Cohen & Staub, 2014; Filik, 2008; Rayner, Warren, Juhasz, & Liversedge, 2004; Staub, Rayner, Pollatsek, Hyönä, & Majewski, 2007; Warren & McConnell, 2007). Implausibility cannot affect eye movements until the reader has actually constructed, or at least entertained, an implausible semantic interpretation; thus, these studies place an upper bound on how long it takes to integrate a word into a representation of the sentence's meaning. In the first of these studies, Rayner et al. (2004) compared reading times on a critical word (e.g., *carrots*) and subsequent material when this word is plausible, given a preceding context, and when the word is implausible:

(7)

a. John used a knife to chop the large carrots for dinner.

b. John used a pump to inflate the large carrots for dinner.

In general, increased reading times appear in early measures of processing on the implausible word, as early as the reader's first fixation on this word in some studies (Staub et al., 2007; Warren & McConnell, 2007). In addition, more severe plausibility violations have earlier effects than less severe violations (Rayner et al., 2004; Warren & McConnell, 2007).

Taken together, these studies suggest that, within the span of a reader's first eye fixation on a word, which is generally shorter than 300 ms, the reader has not only extracted a great deal of lexical information from the fixated word but has also at least begun the process of integrating that word into a representation of the sentence's structure and meaning. An error signal from the latter processes can, at least in some circumstances, affect the duration of that very first fixation (see Reichle, Warren, & McConnell, 2009, for an explicit computational model of how this might work).

Several decades of research using ERPs has also found evidence of word-by-word processes of parsing and interpretation. Event-related potentials are a form of electroencephalography (EEG), in which researchers examine the effect of a specific stimulus on the pattern of electrical potentials measured at the scalp. In the present case, the EEG is recorded as a reader encounters a critical word in a sentence. Because eye movements cause artifacts in the EEG record, words are usually presented one word at a time in the center of a computer screen, using a method called rapid serial visual presentation (RSVP). A basic conclusion of this research is that it is possible to see effects of the relationship between an input word and its sentence context within a few hundred milliseconds. A word that is semantically anomalous in its context typically elicits an increase in the amplitude of the N400 component (Kutas & Hillyard, 1980), a negative deflection in the EEG record that peaks around 400 ms after word onset (see Kutas & Federmeier, 2007, for a review). When a word is syntactically anomalous is its context, the pattern is somewhat more complex, as this has been shown to elicit both a relatively late positive deflection, known as the P600 (Osterhout & Holcomb, 1992), and an earlier negative deflection, similar in latency to the N400 but with a different scalp distribution, known as the left anterior negativity (LAN; see Kutas & Federmeier, 2007, for review). Some researchers have reported an especially early version of the LAN, known as the ELAN (Friederici, Hahne, & Mecklinger, 1996; Neville, Nicol, Barss, Forster, & Garrett, 1991), and there is ongoing controversy about the circumstances in which this very early component arises.

Taken together, eye movement and ERP studies establish quite definitively that both syntactic parsing and semantic interpretation happen incrementally, at least most of the time. As we read each word, we try to fit that word into the representation of the sentence's syntactic structure and meaning that we have been building up. If this process runs into difficulty, the consequences can be seen within a few hundred milliseconds in both the eye-movement and EEG records.

Some recent results suggest, however, that at least some of the time, the incremental representation that readers construct may not be fully faithful to the input, or may be left unspecified in certain ways. For example, it appears that readers may sometimes fail to correct initial syntactic misanalyses (e.g., Christianson, Hollingworth, Halliwell, & Ferreira, 2001) and that they may strategically adopt a relatively superficial mode of reading, which enables them to avoid committing to an interpretation of an ambiguous sentence (Swets, Desmet, Clifton, & Ferreira, 2008). It also appears that readers may occasionally analyze a string of adjacent words in a manner that is not consistent with the sentence as a whole (Tabor, Galantucci, & Richardson, 2004). While there remain important questions about the scope and proper interpretation of these findings, they have certainly generated a great deal of interest in the sentence-processing community.

We have concluded that parsing is highly incremental. But in fact, there is evidence that parsing may be *hyper*incremental: Readers not only structure the input as they receive it but also actively anticipate upcoming structure. Some of this evidence comes from processing of structures known as *long distance dependencies,* in which an element appearing at one point in the sentence receives its thematic role (e.g., agent, patient) at a later point in the sentence. Consider sentences (8a–b), from Pickering and Traxler (2003):

(8)

a. That's the truck that the pilot landed carefully behind in the fog.

b. That's the plane that the pilot landed carefully behind in the fog.

In these sentences the initial noun phrase (*the truck/the plane;* known in psycholinguistic terminology as a *filler*) is semantically—and on some theories, syntactically—linked to a position after the word *behind* (known as the *gap*), where it is assigned a role in the sentence's meaning. In reading such a sentence, the reader first encounters the filler, and then later encounters the gap.

But there is evidence that the parser does not actually wait for direct evidence as to the location of the gap, instead actively anticipating the gap in the first location at which it could appear (e.g., Crain

& Fodor, 1985; Frazier & Clifton, 1989; Garnsey, Tanenhaus, & Chapman, 1989; Stowe, 1986; Traxler & Pickering, 1996). This evidence takes the form of processing disruption when the first possible gap site turns out not to contain the gap after all. An elegant demonstration comes from Pickering and Traxler (2003). In 8a and 8b, the parser could initially posit a gap after *landed*, which is the gap location in a sentence such as *That's the plane that the pilot landed on runway number three*. But in the sentences presented by Pickering and Traxler (2003), the gap was not in this location. Instead, as noted earlier, the gap was later in the sentence, after the word *behind*. Pickering and Traxler found that reading time on the words *landed carefully* was inflated in 8a compared to 8b. In 8a, it is at this point in the sentence that the reader is able to conclude that the gap can't be after *landed*, as you can't land a truck. In 8b, however, the gap after *landed* is still plausible until the reader reaches the word *behind*; and indeed, Pickering and Traxler found that on that next region of the sentence, reading times were longer in 8b than in 8a. In short, it appears that readers posited the gap in the first possible location, after *landed*, and incurred a processing cost at the point in each sentence at which this analysis was ruled out.

There may be other structures, in addition to long distance dependencies, in which the parser actively predicts a grammatically required element. For example, Staub and Clifton (2006) demonstrated that readers anticipate a disjunction after reading the word *either*. Staub and Clifton had subjects read sentences like *The team took the train or the subway to get to the game*, and found that including the word *either* before *the train or the subway* reduced reading time on *or the subway*. But several recent models have made an even stronger suggestion regarding predictive or anticipatory processes in parsing (e.g., Hale, 2001; Levy, 2008). These models propose that readers maintain expectations for how a sentence will continue based on the statistics of their language experience; more likely continuations are more expected, and less likely continuations are less expected. These models also propose that these statistically based expectations determine the difficulty of processing the material that is actually encountered.

Seriality Versus Parallelism

We now know that the sentence processing system constructs a syntactic analysis and a semantic interpretation in a highly incremental manner. But does it construct just one syntactic analysis and interpretation? Or might the system entertain multiple analyses at the same time in cases of ambiguity? In the parlance of cognitive psychology, this is a question about whether the system is *serial* or *parallel.* A serial model holds that at any moment the system maintains only a single analysis of a sentence, and that it replaces this analysis if it proves to be incorrect. The paradigmatic serial model is Frazier's (1978, 1987) garden path model; more recent serial models include Lewis and Vasishth's (2005) ACT-R-based parser, and Hale's rational left-corner parser (2011). A parallel model, on the other hand, holds that the system may maintain multiple analyses at the same time, with these analyses varying in their degree of activation. Another way of putting this is to say that the system will construct more than one analysis of ambiguous material and will remain in a state of indecision between these analyses until clearly disambiguating material is encountered. This view is represented by the constraint-based approach to sentence comprehension (MacDonald, Perlmutter, & Seidenberg, 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998) and more recently by the surprisal model of Levy (2008). The unrestricted race model of van Gompel and colleagues (Traxler, Pickering, & Clifton, 1998; van Gompel, Pickering, Pearson, & Liversedge, 2005; van Gompel, Pickering, & Traxler, 2000, 2001) is a kind of hybrid; it proposes that at a point of ambiguity, multiple analyses race to be constructed, with the comprehender adopting the analysis that finishes first. Thus, only one complete analysis is entertained.

A critical empirical issue that has the potential to distinguish serial from parallel models was noted at the very beginning of psycholinguistics, by Fodor et al. (1974):

> Practically any model which is parallel in spirit will differ in an important way from any model which is fundamentally serial: parallel theories predict that the computational difficulty of a sentence ought, in general, to increase as a function of the number of unresolved ambiguities it contains; serial theories do not. This is because parallel theories claim that each time we encounter an *n*-ways ambiguous portion of a sentence, we must compute and store *n* paths of analysis, with *n* reducing to one only if disambiguating material is encountered. On the serial model, however, ambiguity should cause increased computational loads only when the first analysis assigned turns out to be the wrong one; i.e.,

only when the character of the subsequent input is incompatible with whatever reading was assigned to the ambiguous material. (p. 362)

In fact, there is essentially no evidence that syntactic ambiguity itself is costly in terms of reading time; the data are reviewed at length in Clifton and Staub (2008; see also Lewis, 2000). As predicted by serial models, there is abundant evidence that disambiguation toward an initially dispreferred analysis is costly, as in (1) and (6b). And, as noted earlier, in eye movement studies this cost often takes the form of an increase in the probability of a regressive eye-movement to an earlier point in the sentence. But reading is not disrupted by syntactic ambiguity itself. Indeed, van Gompel and colleagues (Traxler et al., 1998; van Gompel et al., 2005 2000, 2001; van Gompel et al.,) showed that some globally ambiguous sentences are actually read faster than their unambiguous counterparts. Consider 9a–c:

(9)

a. The driver of the car with the moustache was pretty cool.
b. The car of the driver with the moustache was pretty cool.
c. The son of the driver with the moustache was pretty cool.

In 9c, the reader may treat *with the moustache* as modifying either *the son* or *the driver*; but in 9a and 9b, plausibility dictates which of the preceding nouns the prepositional phrase must modify. In fact, Traxler et al. (1998) found that sentences like 9c were read faster, not slower, than sentences like 9a and 9b.

This lack of difficulty associated with syntactic ambiguity stands in contrast to the abundant evidence that encountering a word with two distinct meanings (e.g., *bank, port*) may indeed have a reading time cost, depending on the relative frequency of these meanings and on the sentence context (Duffy, Morris, & Rayner, 1988; Rayner & Duffy, 1986). Thus, models of visual word recognition can assume that word meanings are accessed in parallel, with these meanings competing for selection. The fact that reading time cost is easily demonstrated in the case of lexical ambiguity, but not in the case of syntactic ambiguity, presents a challenge for parallel sentence processing models. See Green and Mitchell (2006) for an argument that parallel sentence processing models need not actually predict an ambiguity cost, and see Clifton and Staub (2008) for a rejoinder to this argument.

What is the current status of the seriality versus parallelism debate? Parallel constraint-based models were initially cast in theoretical opposition to the dominant serial model of the 1980s and 1990s, the garden path model of Frazier (1978, 1987). Unlike the garden path model, constraint-based models allowed for multiple nonstructural factors to influence initial parsing decisions. For example, the verb *remember* tends to be followed by a direct object, as in he *remembered the story*, while the verb *suspect* tends to be followed by a sentence complement, as in *he suspected the story is false*. Several studies (e.g., Trueswell, Tanenhaus, & Kello, 1993) have shown that such lexical biases appear to influence parsing decisions quite rapidly. As another example, the garden path model predicts that there should be difficulty associated with a sentence such as *She'll implement the plan she proposed tomorrow*, because of a structural preference to attach a modifier (*tomorrow*) to the most recent phrase. However, Altmann, van Nice, Garnham, and Henstra (1998) showed that in reading this preference may be overridden by a sufficiently supportive previous discourse context. In addition to allowing nonstructural factors to play a role in initial parsing decisions, constraint-based models also allow that such decisions may not be deterministic, with comprehenders constructing different analyses on different occasions. There is abundant evidence that this is the actual state of affairs, at least for certain kinds of parsing decisions (e.g., Carreiras & Clifton, 1999).

It is important to note, however, that these issues are actually orthogonal to the issue of whether multiple fully formed syntactic analyses are simultaneously entertained. This fact has often been obscured in the literature, with structural parallelism being embraced together with the use of nonstructural parsing heuristics, in whole-hearted endorsement of the constraint-based framework. This is not universally the case, as the parsing models of van Gompel and colleagues (van Gompel et al., 2005; van Gompel et al., 2000, 2001) and Lewis and Vasishth (2005) simultaneously hold that only one fully formed analysis is constructed in cases of ambiguity and that the parse that is constructed depends on a variety of nonstructural factors. Moreover, they hold that parsing decisions are probabilistic rather than deterministic. Thus these models account for the influence of nonstructural factors on parsing, but also (correctly) avoid predicting an ambiguity cost.

Independence of Semantic Processing

This chapter is about two processes that are obviously related: syntactic parsing and semantic interpretation. But exactly how are they related? In this section we discuss a theoretical motivation for the idea that syntactic parsing must precede semantic interpretation and some interesting empirical evidence that bears on this issue.

In linguistic semantics, the meaning of a sentence is modeled as a function of the meaning of the words that it contains and the way in which these words are combined syntactically. The principle that the meaning of an expression is a function of its parts and their particular mode of combination is called *compositionality* and is often attributed to the logician Gottlob Frege (see Heim and Kratzer, 1998, for a thorough implementation of this idea). To take a simple example, the fact that (10a) and (10b) are roughly equivalent in meaning, and the fact that (10c) does not mean the same thing as (10a), despite the fact that in both sentences *the dog* is in subject position, arises from the fact that the passive structure in (10b) and (10c) puts the verb's theme in subject position, and the verb's agent in the by-phrase.

(10)

a. The dog bit the man.
b. The man was bitten by the dog.
c. The dog was bitten by the man.

Applied to online processing, this idea would suggest that semantic interpretation follows, both temporally and logically, from the construction of a syntactic parse; it is the syntactic analysis of a sentence that determines, for the comprehender, how the meanings of the individual words are to be combined. Thus syntactic attachment of an input word into the current parse of the sentence should precede any attempt to determine semantic relationships between this word and previous material, and semantic interpretation should be constrained by this syntactic parse. But in fact, many psycholinguists have suggested that this may not be the case (e.g., Bornkessel & Schlesewsky, 2006; Kuperberg, 2007; Townsend & Bever, 2001). Rather, these theorists have proposed versions of the idea that in at least some circumstances, semantic interpretation need not await syntactic analysis, relying instead on quick, automatic nonsyntactic heuristics for combining words in a sensible way. For example, the comprehender may initially arrive at an interpretation that is plausible given world knowledge, even if that interpretation is not licensed by the syntactic parse of the sentence.

An ERP study by Kim and Osterhout (2005) has been influential in support of this idea. As mentioned earlier, a plausibility violation generally elicits an increased amplitude N400 waveform in the ERP record, while syntactic violations such as errors in agreement, phrase-structure violations, or incorrect verb morphology elicit an increased P600. Kim and Osterhout presented sentences such as (11a–c) in RSVP format:

(11)

a. The hearty meal was devouring the kids.
b. The hearty meal was devoured by the kids.
c. The hungry boy was devouring the cookies.

Sentence 11a is syntactically well formed, but implausible: A meal cannot devour anything. But in fact, the word *devouring* in 11a elicited an increased P600, not an increased N400, compared with the corresponding words in 11b and 11c. It appears that the brain responds to this word as if it constituted a syntactic anomaly, rather than a semantic one. Why might this be? Kim and Osterhout suggested that in reading a sentence like (11a), we use our knowledge that a hearty meal is very likely to be devoured to initially construct a semantic interpretation on which *the hearty meal* is the theme of *devouring,* not the agent. Because we are treating *the hearty meal* as the theme, we expect the participle to be *devoured*, rather than *devouring*. Thus the problem with the participle *devouring* appears to be incorrect morphology (*-ing* rather than *-ed*); that is, a syntactic error. Kim and Osterhout concluded that "at least under some circumstances, semantic processing operates independently of and perhaps even controls syntactic analysis" (p. 210). They also demonstrated in a second experiment that the critical result is obtained only when the sentence's subject is an attractive theme for the verb, as they found an N400, not a P600, for *The dusty tabletops were devouring*. This result is consistent with the hypothesis that initial semantic interpretation is based on real-world plausibility. This work has generated much discussion and several follow-up studies (Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb, 2006; Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2007; Stroud & Phillips, 2012; Van Herten, Kolk, & Chwilla, 2005).

On the other hand, another line of ERP research has addressed the issue of the potential independence of semantic interpretation by asking what happens when a reader encounters a word that is both syntactically anomalous and constitutes a semantic violation (see Bornkessel-Schlesewsky & Schlesewsky, 2009, and Friederici, 2002, for a review of this work). The following are examples from a German study by Hahne and Friederici (2002); the critical word is in italics:

(12)

a. Das Brot wurde *gegessen*. 'The bread was eaten'
b. Das Vulkan wurde *gegessen*. 'The volcano was eaten'
c. Das Eis wurde im *gegessen*. 'The ice cream was in-the eaten'
d. Das Turschloß wurde im *gegessen*. 'The door lock was in-the eaten'

Sentences 12b through 12d are all unacceptable at the point of reaching the participle *gegessen*. In 12b, this word is semantically implausible. An increased N400 is expected in this condition compared with 12a. In 12c, this word is immediately preceded by a preposition, resulting in a phrase structure violation. A combination of an early negativity and a P600 is expected in this condition, compared with 12a. The critical question was whether 12d, in which both types of violation are combined, would still elicit an N400. If so, this would suggest that the processing system noted the semantic implausibility of eating a door lock, even though the participle couldn't be attached syntactically to the preceding preposition in any case. But in fact, while the early negativity and P600 occurred in sentences like 12d, no N400 effect was detected. Similar results have been obtained in several other studies, suggesting that the presence of a phrase structure violation effectively blocks any attempt at compositional semantic interpretation.

In recent studies, Zhang, Yu, and Boland (2010) and Wang, Mo, Xiang, Xu, and Chen (2013) have found that the situation may be different for Chinese readers. Both studies found that semantic violations elicited an N400 effect even when the critical word also constituted a syntactic category violation. These authors suggested that Chinese speakers may adopt fundamentally different processing strategies from speakers of the other languages in which these studies have been carried out (German, Dutch, and French). Chinese lacks overt cues to syntactic structure such as marking for number, gender, and case, which may induce reliance on a semantics first processing strategy.

In sum, despite a large number of studies using ERP methodology that have addressed the question of whether semantic interpretation can proceed without a full syntactic parse, it is not possible at this point to provide a definitive answer to this question. It is likely that this will remain a fertile area for research in the future.

The Role of Memory

Understanding sentences makes demands on memory. Consider (13):

(13) Which horse did Mary think that her daughter wanted to ride?

This sentence is fairly easy to understand, considering that there are several points at which you had to retrieve some linguistic element from your memory of an earlier part of the sentence. That is, there are several points at which an input word, as you read the sentence from left to right, had to be linked with an element that did not immediately precede it:

- the pronoun *her* had to be linked to its antecedent *Mary*
- the verb *ride* had to be linked to its subject, *her daughter*
- the verb *ride* also had to be linked to its object, *Which horse*

Note that the language itself does not place an upper limit on the amount of material over which an element can be held in memory, for later retrieval; consider the following variant of (13), where *Which horse* must again be retrieved upon encountering *ride*:

(14) Which horse did the woman who got out of the gray Mercedes behind the house think that her daughter wanted to ride?

While this sentence is more difficult, it is still comprehensible. Evidently, we are fairly skilled at holding linguistic elements in memory as we read or hear a sentence, and then retrieving them at the appropriate points.

However, holding a linguistic element in memory and retrieving it later are not entirely without cost. A recent eye-tracking study by Bartek, Lewis, Vasishth, and Smith (2011), following earlier work by Grodner and Gibson (2005), is suggestive. Bartek et al.'s subjects read sentences like those in (15):

(15)

a. The nurse supervised the administrator while…
b. The nurse from the clinic supervised the administrator while…
c. The nurse who was from the clinic supervised the administrator while…

In these sentences, the verb, *supervised*, varies in its distance from the subject, *the nurse*. Bartek et al.'s critical finding was that reading time on the verb was slightly, though reliably, longer when the verb was not adjacent to the subject. They argued that this increase in reading time reflects an increase in the difficulty of retrieving the subject when the verb is encountered, which could arise for two reasons: decay of the memory trace of the subject and interference from intervening material during the process of retrieving the subject from memory (Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006). However, this conclusion must be reconciled with other experiments that have failed to find effects on reading times of distance between elements in a dependency (e.g., Martin & McElree, 2008).

There is also reason to believe that memory limitations play a role in processing more complex structures. In the introduction it was noted that sentence (5a), repeated here, is extremely difficult to understand:

(5) a. The doctor that the nurse that the administrator knows hates resigned.

It appears that holding three noun phrases in memory (*the doctor, the nurse, the administrator*) at the same time, and then retrieving them in rapid succession to fill roles as subject and object of three verbs (e.g., retrieving *the nurse* as the subject of *hates*, and *the doctor* as the subject of *resigned*) is quite difficult. The difficulty goes away almost entirely if there is only a single embedded clause, so that there are only two noun phrases, and two verbs.

(16) The doctor that the nurse hates resigned.

This dramatic increase in difficulty as the number of embeddings increases from one to two was observed many years ago (Miller & Chomsky, 1963; Yngve, 1960). Gibson (1998, 2000) attributed this difficulty to the number of new discourse referents that are introduced in (5a) between *the doctor* and its verb, *resigned,* which the reader must mentally represent and hold in memory. These include *the nurse, the administrator*, and the two intervening verbs. Warren and Gibson (2002) found that this difficulty is greatly alleviated when the noun phrases in the embedded clauses do not, in Gibson's terms, introduce new discourse referents, as in (5b), repeated here:

(5) b. The doctor that everyone I know hates resigned.

Gordon and colleagues (Gordon, Hendrick, Johnson, & Lee, 2006; Gordon, Hendrick, & Levine, 2002) have also suggested that difficulty of sentences like (5a) is due in part to the fact that the noun phrases are all of a similar type and hence interfere with each other in memory. This problem is alleviated in (5b).

Memory limitations have also been invoked to explain one of the oldest findings in psycholinguistics (e.g., Wanner & Maratsos, 1978). This is the finding that object relative clauses (ORCs; 17b) are more difficult to understand than subject relative clauses (SRCs; 17a), even when these two constructions contain exactly the same words:

(17)

a. The reporter that attacked the senator admitted the error.
b. The reporter that the senator attacked admitted the error.

The critical difference in meaning between these sentences is that in (17a), the noun that the relative clause modifies (*the reporter*) is also the agent of the relative clause verb (*attacked*), while in (17b), this noun is the theme of the relative clause verb. Many theorists (e.g., Gibson, 1998; Gordon et al., 2006; Grodner & Gibson, 2005; Lewis & Vasishth, 2005) have proposed that the difficulty of ORCs can be attributed to the operations of working memory. The basic idea is that in (17b), *the reporter* must be held in memory until *attacked* is encountered, at which point it must be retrieved. Critically, the retrieval of *the reporter* is likely to be more difficult when this element is farther back in the sentence, with other material intervening. The presence of this intervening material may be more disruptive when it is similar to the target, as in (17b), where both *the reporter* and *the senator* are noun phrases that denote occupations. Consistent with this account, Gordon et al. (2006) found that reading times for sentences like (17b) were reduced when one of the critical noun phrases was replaced by a name.

Recently, Staub (2010) investigated exactly where the difficulty first appears in the course of reading sentences like (17b), and found that the specific pattern is not consistent with an exclusively memory-based account. Readers do show longer reading times on the verb *attacked* in (17b) than in (17a), which is predicted by memory-based accounts. However, there is even more difficulty on *the senator* in (17b), which is before the reader is required to engage in any difficult memory retrieval. Staub (2010) suggested that some part of the difficulty of ORCs may be due to an initial preference, upon encountering the word *that*, to build an SRC structure. Complicating matters further, recent work by Roland, Mauner, O'Meara, and Yun (2012) suggests that ORCs may not be particularly difficult for readers in the actual contexts in which they occur in real discourse, where the subject of the relative clause (*the senator* in 17b) is likely to have been previously mentioned. In sum, the role of memory limitations in ORC difficulty is still under debate.

Researchers have also asked a broader question about the memory system used in sentence comprehension: Is it the same system used in other tasks, such as in the explicit verbal memory tasks used by cognitive psychologists, or is it a separate system? Caplan and Waters (1999) have argued that it cannot be the same system, on the basis of several kinds of evidence. First, they argued that individual differences in verbal working memory, as established by tests such as the reading span task (Daneman & Carpenter, 1980), which requires subjects to remember the final word of sequentially presented sentences, are not correlated with the ability to comprehend ORCs or garden path sentences. Consistent with this finding, Sprouse, Wagers, and Phillips (2012) also failed to find a relationship between an individual's performance on explicit memory tasks and processing of a different kind of complex sentence. Second, Caplan and Waters argued that imposing an additional memory load during the processing of sentences does not specifically impair comprehension of the kinds of sentences that are thought to make demands on working memory. Third, they argued that patients with working memory disorders do not show impaired comprehension of complex sentences.

The second of these three claims has recently been challenged by Fedorenko, Gibson, and Rohde (2006, 2007). Fedorenko et al. conducted experiments in which subjects read sentences of varying complexity in a self-paced reading paradigm (i.e., pressing a button to reveal each word of the sentence), while either simultaneously remembering several words for an unrelated memory task or performing an arithmetic or spatial cognition task. Reading was slowed when subjects had to simultaneously remember words that were related in meaning to words in the sentence or perform an arithmetic task. The critical result, however, was that the effect of the memory task and the effect of sentence complexity interacted, with the memory task having a larger effect when the sentence was complex. Fedorenko et al. argued, on the basis of these results, that the working memory processes involved in understanding complex sentences are also involved in other memory tasks, including nonverbal tasks such as arithmetic. Still more recently, however, Evans, Caplan, and Waters (2011) have argued that the Federenko et al. dual-task methodology may recruit unusual processing strategies, undermining the generality of their conclusions.

This section has only begun to touch on the many empirical debates, methodologies, and theoretical frameworks related to the role of memory in sentence comprehension. This is a very active research area, and is likely to remain so in the future.

Sentence Processing and Visual Word Recognition

Many of the chapters in this handbook address, from one perspective or another, the topic of visual word recognition (e.g., Yap & Balota, Taft, and Andrews). These chapters generally discuss the recognition of isolated words. But it is important to ask whether word recognition is sufficiently autonomous that it makes sense to discuss this topic without considering the syntactic and semantic context in which a word appears. Can we assume that there is a separate (or *modular*, in the well-known terminology of Fodor, 1983) visual word recognition system and that this system delivers an output that becomes the input to the system or systems that perform syntactic parsing and semantic interpretation? Or is word recognition itself affected by a word's syntactic and semantic context? It is clear that the resolution of lexical ambiguity is affected by context (Duffy et al., 1988), with context helping the reader interpret a letter string such as *bank* as denoting a financial institution or the edge of a river. It also appears that context can influence word recognition when a word is easily confusable with a different word, as when two words differ by one letter (e.g., *trial* and *trail*; Johnson, 2009;

Slattery, 2009). But what about the more usual case, where there is no genuine ambiguity as to a word's identity? Is the process of recognizing a word influenced by context in this case as well? Several decades ago Stanovich (1980) asserted, in the context of understanding differences between good and poor readers, that good readers' behavior in particular is characterized by "rapid context-free word recognition" (p. 32). But is this assertion correct?

Consider sentences (18a–b):

(18)

a. The woman took the warm cake out of the oven and frosted it.

b. The woman walked over to the oven and opened it.

In (18a), the word *oven* is highly predictable, given the preceding context, while in (18b), it is not. This can be demonstrated empirically with a *cloze* task, in which subjects guess the next word of a sentence (Taylor, 1953). It is well established in controlled experiments that the time the eyes spend on a word in reading is affected by a word's predictability (e.g., Ehrlich & Rayner, 1981; Rayner & Well, 1996). Moreover, Smith and Levy (2013) have recently found a very general logarithmic relationship between lexical predictability and reading time in a corpus of eye movement data. The amplitude of the N400 component in ERP research is also affected by predictability (e.g., Federmeier & Kutas, 1999; Kutas & Hillyard, 1984). These effects seem to suggest that a word is, in fact, easier to recognize when it is predictable. Comprehenders may predict specific words that are likely to be coming up in the sentence on the basis of world knowledge (Altmann & Kamide, 1999; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005) or because they maintain probabilistic expectations on the basis of statistical regularities in the language (Smith & Levy, 2013), and these predictions may facilitate word recognition, perhaps by reducing the amount of bottom-up evidence required for lexical identification (e.g., Norris, 2006).

However, it is not mandatory to regard effects of predictability as effects on the process of word recognition. As discussed in this volume (e.g., Yap & Balota) word frequency has a reliable influence on word recognition difficulty in a variety of paradigms. If predictability also influences word recognition, then the effects of word frequency and word predictability might interact, as both factors influence a common processing stage (e.g., McDonald & Shillcock, 2003). More specifically, many theorists have found it intuitive that there should be a smaller or nonexistent word frequency effect when the context renders a word highly predictable. This hypothesis has been tested in several eye-movement studies in which the frequency and predictability of a target word have both been varied. Rather surprisingly, the effects of word frequency and word predictability on measures such as first-fixation duration (the duration of a reader's very first eye fixation on a word) and gaze duration (the sum of all fixation durations on the reader's first inspection of a word) are consistently additive, not interactive (e.g., Altarriba, Kroll, Scholl, & Rayner, 1996; Ashby, Rayner, & Clifton, 2005; Gollan et al., 2011; Hand, Miellet, O'Donnell, & Sereno, 2010; Rayner, Ashby, Pollatsek, & Reichle, 2004). Even when predictability is high, the effect of word frequency is not attenuated. The lack of interaction between frequency and predictability is a challenge for models proposing that predictability and frequency affect common processes.

As noted earlier, reading times on a word are also inflated when a word is a poor fit in its context, either because it is inconsistent with the reader's syntactic analysis at that point in the sentence or because it is implausible. Because these effects arise so rapidly, it is again tempting to regard them as reflecting effects on the difficulty of word recognition. However, the most recent version of the E-Z Reader model of eye movement control in reading (Reichle et al., 2009) implements the idea that these effects actually arise postlexically, after a word has already been identified. Staub (2011) tested the detailed predictions of this model by examining readers' eye movements in several experiments, including with sentences like those in (19). In these experiments a critical word was either high or low in frequency (*walked* vs. *ambled*) and occurred in a location where it was likely to disambiguate a garden path (versions c and d), or where the garden path was eliminated due to the presence of a comma (versions a and b):

(19)

a. While the professor lectured, the students walked across the quad.

b. While the professor lectured, the students ambled across the quad.

c. While the professor lectured the students walked across the quad.

d. While the professor lectured the students ambled across the quad.

These experiments confirmed the model's central predictions. The probability of a regressive eye movement from the critical word was higher in sentences like 19c and 19d. Word frequency, however, affected reading times but not the rate of regressions. When both manipulations did affect reading times, they had additive rather than interactive effects.

In sum, the large community of researchers studying the recognition of single words can take some comfort from the current state of research on word recognition in context. While a word's sentence context does influence very early processing measures, it is far from clear that these effects of context are, in fact, effects on word recognition itself. This should also provide comfort to sentence-processing researchers, who have generally regarded complete word representations as the input to the processes of syntactic parsing and semantic interpretation.

Conclusions

The main goal of this chapter has been to provide a sense of the issues and debates that characterize research on sentence processing in reading. We have argued that the current state of this research supports the following conclusions: First, both syntactic parsing and semantic interpretation are highly incremental, with readers updating syntactic and semantic representations on a word-by-word basis as they read. Second, it appears that readers maintain only a single syntactic representation at a time and reanalyze when necessary; there is no convincing evidence for parallelism in syntactic parsing. Third, some results do suggest that readers may construct a semantic interpretation in advance of, or in contradiction to, a syntactic parse of the sentence, though the literature is equivocal on this point. Fourth, comprehending sentences makes demands on a form of short-term memory, though the exact circumstances in which it does so are not clear and it is also unclear whether the form of memory used in sentence comprehension is shared with other tasks. Fifth, the evidence to date does not undermine the assumption of a modular architecture in which visual word recognition is functionally separate from processes of sentence comprehension.

The qualified nature of these conclusions may lead to the impression that there are as many open questions in this area as there are answers. This impression is intended. We know much less about how readers put words together, in real time, than we do about how readers recognize words themselves. Only in the last thirty years have we had methods for investigating truly online processing of written sentences (taking Frazier and Rayner, 1982, as a starting point), and while much has been learned in this time, thirty years is not a very long time, given the pace of scientific discovery.

References

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed language sentences: Evidence from eye fixation and naming times. *Memory & Cognition*, *24,* 477–492.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73,* 247–264.

Altmann, G., van Nice, K. Y., Garnham, A., & Henstra, J. A. (1998). Late closure in context. *Journal of Memory and Language*, *38,* 459–484.

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *Quarterly Journal of Experimental Psychology*, *58A,* 1065–1086.

Bartek, B., Lewis, R. L., Vasishth, S., & Smith, M. R. (2011). In search of on-line locality effects in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37,* 1178–1198.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–352). New York, NY: Wiley.

Bornkessel, I., & Schlesewsky, M. (2006). The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, *113,* 787–821.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). *Processing syntax and morphology: A neurocognitive perspective.* Oxford, England: Oxford University Press.

Caplan, D. (2013). Sentence processing and comprehension in atypical populations. In R. P. G. van Gompel (Ed.), *Sentence processing* (pp. 247–268). Hove, England: Psychology Press.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22,* 77–126.

Carreiras, M., & Clifton, C. (1999). Another word on parsing relative clauses: Eyetracking evidence from Spanish and English. *Memory & Cognition*, *27,* 826–833.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, *42,* 368–407.

Church, K., and R. Patil. 1982. Coping with syntactic ambiguity, or how to put the block in the box on the table. *American Journal of Computational Linguistics*, *8,* 139–149.

Clifton, C., Jr., & Staub, A. (2008). Parallelism and competition in syntactic ambiguity resolution. *Language and Linguistics Compass*, *2,* 234–250.

Clifton, C., Jr., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. P. G. van Gompel (Ed.), *Eye movements: A window on mind and brain* (pp. 341–372). Amsterdam, the Netherlands: Elsevier.

Clifton, C., Jr., Traxler, M. J., Mohamed, M. T., Williams, R. S., Morris, R. K., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language*, *49,* 317–334.

Cohen, A., & Staub, A. (2014). On-line processing of novel noun-noun compounds: Eye movement evidence. *Quarterly Journal of Experimental Psychology*, *67,* 147–165.

Crain, S., & Fodor, J. D. (1985). How can grammars help parsers? In D. R. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing* (pp. 94–128). Cambridge, England: Cambridge University Press.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19,* 450–466.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, *27,* 429–446.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, *20,* 641–655.

Evans, W. S., Caplan, D., & Waters, G. (2011). Effects of concurrent arithmetical and syntactic complexity on self-paced reaction times and eye fixations. *Psychonomic Bulletin & Review*, *18,* 1203–1211.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, *41,* 469–495.

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language*, *54,* 541–553.

Fedorenko, E., Gibson, E., & Rohde, D. (2007). The nature of working memory in linguistic, arithmetic, and spatial integration processes. *Journal of Memory and Language*, *54,* 246–269.

Filik, R. (2008). Contextual override of pragmatic anomalies: Evidence from eye movements. *Cognition*, *106,* 1038–1046.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language: An introduction to psycholinguistics and generative grammar*. New York, NY: McGraw Hill.

Fodor, J. D., & Ferreira, F. (Eds.). (1998). *Reanalysis in sentence processing*. Dordrecht, the Netherlands: Kluwer.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. Doctoral dissertation, University of Connecticut, Storrs.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance: Vol. 12. The psychology of reading* (pp. 559–586). Hillsdale, NJ: Erlbaum.

Frazier, L., & Clifton, C., Jr. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, *28,* 331–344.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14,* 178–210.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6,* 78–84.

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning*, *Memory*, *and Cognition*, *22,* 1219–1248.

Garnsey, S. M., Tanenhaus, M. K., & Chapman, R. M. (1989). Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research*, *18,* 51–60.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68,* 1–76.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). Cambridge, MA: MIT Press.

Gollan, T. H., Slattery, T. J., Goldenberg, D., van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency lag hypothesis. *Journal of Experimental Psychology: General*, *140,* 186–209.

Gordon, P. C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning*, *Memory, and Cognition*, *32,* 1304–1321.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, *13,* 425–430.

Green, M. J., & D. C. Mitchell. 2006. Absence of real evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language, 55,* 1–17.

Grodner, D. J., & Gibson, E. A. F. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science*, *29,* 261–291.

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research, 13,* 339–356.

Hale, J. (2001, June). *A probabilistic Earley parser as a psycholinguistic model*. Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies, pp. 1–8, Association for Computational Linguistics.

Hale, J. T. (2011). What a rational parser would do. *Cognitive Science*, *35,* 399–443.

Hand, C. J., Miellet, S., O'Donnell, P. J., & Sereno, S. C. (2010). The frequency-predictability interaction in reading: It depends on where you're coming from. *Journal of Experimental Psychology: Human Perception and Performance*, *36,* 1294–1313.

Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Malden, MA: Blackwell.

Johnson, R. L. (2009). The quiet clam is quite calm: Transposed-letter neighborhood effects on eye movements during reading. *Journal of Experimental Psychology: Learning*, *Memory*, *and Cognition*, *35,* 943–969.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, *52,* 205–225.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146,* 23–49.

Kuperberg, G. R., Caplan, D., Sitnikova, T., Eddy, M., & Holcomb, P. J. (2006). Neural correlates of processing syntactic, semantic, and thematic relationships in sentences. *Language and Cognitive Processes*, *21,* 489–530.

Kuperberg, G. R., Kreher, D. A., Sitnikova, T., Caplan, D. N., & Holcomb, P. J. (2007). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language*, *100,* 223–237.

Kutas, M., & Federmeier, K. D. (2007). Event-related brain potential (ERP) studies of sentence processing. In G.

Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 385–406). Oxford, England: Oxford University Press.
Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207,* 203–205.
Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307,* 161–163.
Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106,* 1126–1177.
Lewis, R. L. (2000). Falsifying serial and parallel parsing models: Empirical conundrums and an overlooked paradigm. *Journal of Psycholinguistic Research*, *29,* 241–248.
Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29,* 375–419.
Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Science*, *10,* 447–454.
MacDonald, M. C., Perlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101,* 676–703.
Marcus, M. P. (1980). *A theory of syntactic recognition for natural language*. Cambridge, MA: MIT Press.
Martin, A. E., & McElree, B. (2008). A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language*, *58,* 879–906.
McDonald, S. A., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, *14,* 648–652.
McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38,* 283–312.
Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In D. Luce (Ed.), *Handbook of mathematical psychology* (pp. 419–491). Hoboken, NJ: Wiley.
Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *3,* 151–165.
Norris, D. (2006). The Bayesian reader: Exploring word recognition as an optimal Bayesian decision process. *Psychological Review*, *113,* 327–357.
Osterhout, L., & Holcomb, P. J. (1992) Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31,* 785–806.
Pickering, M. J., & Traxler, M. J. (2003). Evidence against the use of subcategorisation frequencies in the processing of unbounded dependencies. *Language and Cognitive Processes*, *18,* 469–503.
Rayner, K., & Duffy, S. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14,* 191–201.
Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, *3,* 504–509.
Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E–Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance*, *30,* 720–732.
Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30,* 1290–1301.
Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher-level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16,* 1–21.
Roland, D., Mauner, G., O'Meara, C., & Yun, H. (2012). Discourse expectations and relative clause processing. *Journal of Memory and Language*, *66,* 479–508.
Slattery, T. J. (2009). Word misperception, the neighbor frequency effect, and the role of sentence context: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *35,* 1969–1975.
Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128,* 302–319.
Snedeker, J. (2013). Children's sentence processing. In R. P. G. van Gompel (Ed.), *Sentence processing* (189–220). Hove, England: Psychology Press.
Sprouse, J., Wagers, M., & Phillips, C. (2012). A test of the relation between working-memory capacity and syntactic island effects. *Language*, *88,* 82–123.
Stanovich, K. E. (1980). Toward and interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, *16,* 32–71.
Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, *116*, 71–86.
Staub, A. (2011). Word recognition and syntactic attachment in reading: Evidence for a staged architecture. *Journal of Experimental Psychology: General*, *140,* 407–433.
Staub, A., & Clifton, C., Jr. (2006). Syntactic prediction in language comprehension: Evidence from *either…or*. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32,* 425–436.
Staub, A., Rayner, K., Pollatsek, A., Hyönä, J., & Majewski, H. (2007). The time course of plausibility effects on eye movements in reading: Evidence from noun-noun compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33,* 1162–1169.
Stowe, L. A. (1986). Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, *1,* 227–245.
Stroud, C., & Phillips, C. (2012). Examining the evidence for an independent semantic analyzer: An ERP study in Spanish. *Brain and Language*, *120,* 108–126.
Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, *36,* 201–216.
Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language*, *50,* 355–370.
Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism Quarterly*, *30,* 415–433.
Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.
Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, *35,* 454–475.
Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role

information in syntactic disambiguation. *Journal of Memory and Language*, *33,* 285–318.
Trueswell, J., Tanenhaus, M. K., & Kello, C. (1993). Verb specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning*, *Memory*, *and Cognition*, *19,* 528–533.
Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning*, *Memory*, *and Cognition*, *31,* 443–467.
van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, *52,* 284–307.
van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2000). Unrestricted race: A new model of syntactic ambiguity resolution. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 621–648). Oxford, England: Elsevier.
van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, *45,* 225–258.
Van Herten, M., Kolk, H. H., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, *22,* 241–255.
Wang, S., Mo, D., Xiang, M., Xu, R., & Chen, H. C. (2013). The time course of semantic and syntactic processing in reading Chinese: Evidence from ERPs. *Language and Cognitive Processes*, *28,* 577–596.
Wanner, E., & Maratsos, M. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 119–161). Cambridge, MA: MIT Press.
Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, *85,* 79–112.
Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulleting & Review*, *14,* 770–775.
Yngve, V. H. (1960). A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, *104,* 444–466.
Zhang, Y., Yu, J., & Boland, J. E. (2010). Semantics does not need a processing license from syntax in reading Chinese. *Journal of Experimental Psychology: Learning*, *Memory*, *and Cognition*, *36,* 765–781.

CHAPTER

15

Models of Discourse Comprehension

Edward J. O'Brien *and* Anne E. Cook

Abstract

Although models of discourse comprehension share many local coherence processing assumptions, the primary distinction that separates these models is how readers gain access to inactive portions of memory. Strategy-based models contain the assumption that readers actively search the long-term memory representation of the text for information relevant to the contents of working memory. In contrast, in memory-based models, inactive information that is related to the current input is activated via a passive resonance mechanism. Three topics within the research on reading comprehension are reviewed in which strategy-based and memory-based models make differing predictions: necessary inferences, elaborative inferences, and the maintenance of global coherence. The chapter demonstrates that the empirical evidence overwhelmingly supports the memory-based view, and offers new questions in an attempt to spur further research in reading comprehension.

Key Words: comprehension, memory-based model, strategy-based model, inference, global coherence

As indicated by the large number of topics covered in this handbook, reading involves multiple complex processes. The ultimate goal of all of these processes is comprehension—understanding the meaning of the information stated in or implied by a text. In order for comprehension to be possible, the reader must first successfully perceive, decode, and recognize words and access their meanings, and then join those words together into syntactic constructions that are both legal and meaningful. In this chapter, we will focus on how the output of those processes feeds and interacts with the mechanisms that drive and influence the reader's primary reason for reading in the first place—comprehension.

Basic Assumptions in Comprehension Research

Within all current models of discourse comprehension there is the assumption that readers generate at least two levels of representation of a text. The first level is commonly referred to as the *textbase*; it is a representation of the actual text itself, the words and relations among the words explicitly stated within the text. The textbase is typically assumed to be in the form of connected propositions, and propositions are merely a formalized and objective representational structure designed to capture individual, basic, idea units. Propositions are connected primarily as function of argument overlap (e.g., Kintsch & van Dijk, 1978; Kintsch & Vipond, 1979) and causal relations (e.g., Keenan, Baillet, & Brown, 1984; O'Brien & Myers, 1987; Trabasso, Secco, & van den Broek, 1984). The second level of representation is the *situation model* (Johnson-Laird, 1983; Johnson-Laird & Garnham, 1980; van Dijk & Kintsch, 1983). The situation model is a representation of what the text is about; it contains information explicitly stated in the text as well as information from the reader's general world knowledge that helps fill in the fully intended meaning of the text. The resulting representation is one that, if accurately constructed by the reader, reflects a coherent understanding of the general situation described in the text as well as the intended meaning of the writer. In what follows, we take the development of a coherent situation model as the minimum requirement for successful comprehension of text, and our discussion will focus on this level of representation.

Because having a representation cannot be considered devoid of process, an important question to consider is What are the processes that underlie comprehension and result in a coherent representation of a text in memory? This question has dominated much of the research in discourse processing for the last two decades. Some fundamental processing assumptions that are common to all models of comprehension were first explicated in the Kintsch and van Dijk (1978; van Dijk & Kintsch, 1983) models. Within their framework, comprehension is a cyclical process that operates within the constraints of a limited-capacity system. Upon each new input cycle, propositions are encoded and connected to other propositions in active memory on the basis of argument overlap or causal relations. A subset of these propositions is maintained in working memory for integration with the next input cycle; the remaining propositions are not, but are connected to the overall discourse representation in long-term memory. As long as each encoded proposition can be connected to the contents of working memory, local coherence is maintained. When there is no ready connection, there is a coherence break; when this occurs, the reader must engage in additional processing (e.g., searching long-term memory for a connection, or drawing an inference in order to create a new connection). However, this raises the question of what is meant by coherence. A simple answer is that coherence is maintained as long as incoming information can be readily integrated with previously presented information. However, coherence can occur at two levels: local and global. Local coherence is maintained as long as incoming information can be readily integrated with information that currently resides within the active portion of memory (i.e., the active portion of the text representation in memory). Global coherence involves establishing connections between incoming information and related information that is no longer within the active portion of the text representation. The processes that are assumed by different theories of discourse comprehension to underlie the maintenance of global coherence will be addressed in the next section.

There are two common conditions under which a reader must gain access to inactive information in memory in order to maintain coherence. The first condition occurs when processes involved in the maintenance of local coherence fail. When the information necessary to maintain local coherence is not readily available in working memory, the reader must gain access to information in long-term memory (either inactive portions of the text representation or general world knowledge) in order to find information that will allow for coherence to be re-established. This typically involves the activation of a necessary inference (e.g., a bridging inference or an anaphoric inference) to re-establish or create a new connection. The second condition under which readers must gain access to inactive information in memory is in the maintenance of global coherence. There is another condition in which readers may gain access to inactive information that does not involve the maintenance of either local or global coherence; related inactive information often becomes available that enables the activation of elaborative inferences that expand on explicitly stated information.

Distinguishing Between Models of Comprehension

Although models of discourse comprehension share many coherence processing assumptions, the primary distinction that separates these models is how readers gain access to inactive portions of memory that are then used to either re-establish local coherence, maintain global coherence, or incorporate inferential information. With respect to how readers gain access to inactive portions of memory, models of comprehension generally fall into one of two categories: strategy-based models and memory-based models. Examples of strategy-based models are Gernsbacher's (1990) structure building framework; Graesser, Singer, and Trabasso's (1994) constructionist framework (see also Singer, Graesser, & Trabasso, 1994); and Zwaan's (1999) event-indexing model (see also Zwaan & Radvansky, 1998). These strategy-based models share—to varying degrees—the common assumption that readers continually and actively search long-term memory for explanations that will guide comprehension and attempt to fully integrate current information with all prior relevant information. Within these models, it is also assumed that in order to ensure full comprehension readers are continually and actively constructing inferences—both necessary and elaborative inferences. In contrast, memory-based models—again, to varying degrees—reject the concept of an active search process; information needed to establish local and global coherence or to create elaborative inferences only becomes available to the reader through passive activation processes. Examples of memory-based models are Sanford and Garrod's (1998, 2005) scenario-mapping focus model; Kintsch's (1988, 1998) construction-integration model; Myers and O'Brien's (1998; O'Brien & Myers, 1999) resonance model; and van den Broek, Risden, Fletcher, and Thurlow's (1996) landscape model. Further, a strong version of a memory-based model (e.g., Myers & O'Brien, 1998; O'Brien & Myers, 1999) also rejects the notion of an active inferential process; inferences are only activated and instantiated into memory to the extent that passive memory activation processes make such information available.

It is certainly the case that any complete model of the comprehension process will have to include both memory-based (i.e., bottom-up) processes and strategic (i.e., top-down) processes. However, no such complete model exists. In what follows, we review much of the evidence regarding the activation of inferential information (both necessary and elaborative) and then review evidence regarding the activation of information necessary for the maintenance of global coherence. We review that evidence from a memory-based perspective in an attempt to establish the boundary conditions in which memory-based processes dominate the comprehension process; at the same time, we will attempt to define the point at which strategic processes become necessary. Finally, we conclude with a framework that outlines the factors necessary for a comprehensive model of reading comprehension that includes both memory-based and strategy-based processes.

The key underlying premise that defines and motivates the memory-based approach to how readers gain access to inactive information in memory is that working memory capacity is limited; a reader cannot maintain every word (or proposition) encoded for a given text in active memory. However, readers often need ready access to information that is not currently in working memory. Constantly searching long-term memory would be a time-consuming and resource-consuming endeavor. Moreover, the longer a text becomes, the more demanding such a search process would become. Thus, readers must have easy access to vast amounts of information stored in long-term memory without having to constantly attend to memory retrieval processes. The simplest solution to this easy-access problem is to assume that memory activation during reading operates via a fast, passive, retrieval mechanism (e.g., Gillund & Shiffrin, 1984; Hintzman, 1988; Murdock, 1982; Ratcliff, 1978).

One instantiation of such a passive memory activation process applied to reading research is the resonance model (Myers & O'Brien, 1998; O'Brien & Myers, 1999), which is an activation mechanism that was critical to the development of the memory-based view (Gerrig & McKoon, 1998; Gerrig & O'Brien, 2005; McKoon & Ratcliff, 1992, 1998). The resonance model is based on the assumption that concepts derived from a sentence currently being processed (along with concepts already residing in active memory as a result of reading earlier portions of text) serve as signals to all of long-term memory; this includes both inactive portions of the text as well as general world knowledge. The intensity of the signal depends on the degree of attention given to information in working memory that is currently in focus, but the signal proceeds autonomously and is unrestricted. Concepts from earlier portions of the discourse representation, as well as from general world knowledge, resonate as a function of the degree of match to the input. The match depends on the overlap of semantic features among concepts. Memory elements that are contacted by this initial signal in turn signal to other memory elements. During this resonance process activation builds, and when the process stabilizes, the most active elements enter working memory. Along with being autonomous and unrestricted, a critical aspect of the resonance process in the context of comprehension is that it is dumb. Information that resonates sufficiently is returned to working memory independent of whether that information will ultimately facilitate or hinder processing (i.e., information resonates as a function of relatedness, not relevance).

Necessary Inferences

Foundational work in the development of the resonance model focused on how previously stated but inactive concepts in a text are (re)activated in response to a referential, or anaphoric, phrase—anaphoric inferences. Anaphoric inferences are considered necessary for comprehension, because failing to connect the referential word or phrase to previously read information would result in a local coherence break. Early memory-based studies on anaphoric inferences addressed the following questions: (1) When an antecedent for a given anaphor is not currently active, what is the process by which that antecedent is reactivated? (2) What are the factors that influence this process?

McKoon and Ratcliff (1980; O'Brien, Duffy, & Myers, 1986) demonstrated that upon encountering an anaphor, antecedents are passively activated from long-term memory, as well as related (but not necessarily relevant) propositional information (e.g., Dell, McKoon, & Ratcliff, 1983). In studies that laid the initial groundwork for the resonance model, O'Brien and colleagues (Albrecht & O'Brien, 1991; O'Brien, 1987; O'Brien, Albrecht, Hakala, & Rizzella, 1995; O'Brien & Myers, 1987; O'Brien, Plewes, & Albrecht, 1990; O'Brien, Raney, Albrecht, & Rayner, 1997) investigated the role of memory-based variables in the activation of anaphoric inferences. In an initial set of experiments, O'Brien (1987) found that when there are two candidate antecedents in a text, reading times on the anaphoric phrase were faster when it required reinstatement of the more recent antecedent (the antecedent presented later in the text) than when it required reinstatement of the more distant antecedent, indicating that the more recent antecedent was reactivated more quickly than the more distant one. Presumably, the signal emanating from the anaphoric phrase would reach the more recent antecedent more quickly. In a subsequent study (O'Brien et al., 1990, 1995), passages were written so that the early and late antecedents were from the same general category contained in the anaphoric phrase. For example, consider the passage presented in Box 15.1. The early antecedent is *train* and the late antecedent is *bus*. The anaphoric phrase is *Mark's neighbor asked him how he traveled to his parents'/brother's.* When the anaphoric phrase contained *parents*, the correct antecedent was the early antecedent (e.g., *train*), whereas when the anaphoric phrase contained *brother's* the correct antecedent was the late antecedent (e.g., *bus*).

Box 15.1 Example Passage From O'Brien, Plewes, and Albrecht (1990)

Introduction

Mark had grown up in the city but he had always wanted to live in the country. The first chance he got, he bought some land and moved there. It made him very happy not having to live in the crowded and noisy city.

Establish and Elaborate Early Antecedent

On holidays he would travel by train into the city to visit his parents. While riding in it he liked to watch the countryside as it raced passed him. Sometimes, the clackety-clack it made on the tracks would put him to sleep. He'd wake up quickly, though, when they came to a crossing and it sounded the horn.

Intervening Sentences

Mark couldn't understand why people like his parents preferred to live in the city. He loved all the open spaces and the clean fresh air. His brother had also moved out of the city and was now living in Colorado.

Establish Late Antecedent

Last summer Mark had traveled by bus to visit him. He had loved looking out of it at the countryside as it passed by. Mark enjoyed seeing the ruggedness of the West, but he really preferred the rolling hills of home.

Background

He thought the people who lived near him were among the nicest he had ever met. On Saturdays, he played golf with his neighbor, and on weekends their families would get together for cookouts. One night while they were talking,

Reinstatement Early / Late

Mark's neighbor asked him how he had traveled to his parents'/brother's.

O'Brien et al. (1990) found that when the anaphoric phrase required reinstatement of either the early or late antecedent, naming times for both the correct and the alternative antecedent were facilitated (relative to a baseline naming time measured immediately prior to the anaphoric phrase) as long as the alternative was from the same category as the antecedent. However, when the text structure was kept the same, but the late antecedent was changed so that it was no longer from the same category as the early antecedent (e.g., *train/shed*), reinstatement of the early antecedent never resulted in any activation of the late antecedent. Further, when the early antecedent was elaborated (thereby increasing the amount of activation it would draw), then the more distant early antecedent was activated more quickly than the more recent late antecedent. In a related study, O'Brien and Myers (1987) showed that the number of causal connections leading to and from a potential antecedent was a stronger predictor of antecedent retrieval time than was distance in the surface structure of a text.

From a strategy-based view, the reader would strategically identify the appropriate antecedent in memory, disregarding alternatives even if they are highly related—causally or otherwise. Readers would certainly not activate a (more distant) early antecedent when a more recently presented antecedent was both available and correct. Thus the consistent finding regarding the factors that influence access to inactive antecedents is that they are all memory-based (e.g., distance and elaboration—causal or otherwise).

Within the memory-based view (and a resonance framework in particular), reactivation of antecedents occurs via a process that is not only passive but also dumb. When an anaphor is encountered, any related concepts in memory may be reactivated, even if they are not the correct antecedent and they may ultimately interfere with anaphor resolution (e.g., Cook, 2014; Corbett & Chang, 1983; O'Brien et al., 1995; see also Klin, Guzmán, Weingartner, & Ralano, 2006; Klin, Weingartner, Guzmán, & Levine, 2004; Levine, Guzmán, & Klin, 2000). Further, because the activation process is unrestricted, these interfering concepts need not even have been part of the text. For example, O'Brien and Albrecht (1991) used passages in which the contexts varied with respect to whether they supported an explicitly mentioned antecedent (e.g., *cat*) or an unmentioned concept (e.g., *skunk*). For example, the contextual phrase *a small black cat with a long furry tail* contains information related to the explicitly mentioned *cat* but not to *skunk*. However, the contextual phrase *a small black cat with a white stripe down its tail* explicitly mentions *cat* but is strongly related to the unmentioned concept, *skunk*. The context was followed by a sentence containing an anaphoric phrase. Immediately following this sentence, O'Brien and Albrecht presented naming probes using either the correct antecedent (e.g., *cat*), or the unmentioned concept (e.g., *skunk*) to assess activation levels of these concepts relative to a control condition (i.e., probes presented immediately prior to the sentence containing the anaphoric phrase). They found that *skunk* was activated in memory even when the text contained an explicit reference to *cat*. Moreover, in a speeded-retrieval task, participants often produced the unmentioned concept (*skunk*); however, when the retrieval task emphasized accuracy and not speed, participants never produced the unmentioned concept. It is difficult to envision a strategy-based view in which antecedents that were explicitly stated in a text would be missed in favor of concepts implied by the context, but never mentioned.

Another type of necessary inference is one in which readers must infer a link between an event in the text and its causal antecedent. This type of inference is commonly called a backward causal bridging inference. Keenan et al. (1984; see also Albrecht & O'Brien, 1995; Myers, Shinjo, & Duffy, 1987) demonstrated that the ease of activating these inferences was predicted by the degree of causal relatedness between a consequence and its causal antecedent that existed in general world knowledge. This research was extended by Singer and colleagues (Singer, 1993; Singer & Ferreira, 1983; Singer & Halldorson, 1996; Singer, Halldorson, Lear, & Andrusiak, 1992), who provided evidence that readers will activate causal bridging inferences even when the events are separated by several sentences of text. They concluded that readers actively seek causal explanations for consequent events. However, Rizzella and O'Brien (1996) found that the activation of distant causal explanations occurred even when there was a sufficient causal explanation readily available in active memory. Consider the passage in Box 15.2. In this example, the consequent event is that Billy would be in trouble when his father came home. The passage contained two potential causal explanations (i.e., bridging inferences) for the consequent event: one early in the passage and one late. The late causal explanation (e.g., Billy had broken a window) immediately preceded the consequent event and provided a sufficient explanation for that event. The early causal explanation (e.g., Billy lost his keys) was also sufficient to explain the consequent event, but the early explanation was not active in memory when the consequent event was read. Even though the late causal explanation was both active in memory and sufficient, reading the consequent event led to the reactivation of the early causal explanation, as measured by speeded naming times to a word reflecting either the early casual explanation (keys) or the late one (window). Naming times for probes representing the early explanation were not as fast as naming times for probes representing the more recent, late causal explanation. However, when the early causal explanation was elaborated (see Box 15.2), naming times for the early causal explanation were actually faster than naming times for the late causal explanation. When filler information was added to reduce the activation of the late causal explanation, the same pattern emerged. This occurred despite the fact that the late causal explanation was sufficient, and in one experiment, clearly active and available in memory when the consequence was read. Thus, Rizzella and O'Brien argued that activation of causal explanations was driven by memory-based factors (i.e., elaboration) rather than sufficiency (see also Albrecht & Myers, 1995, 1998; Myers, Cook, Kambe, Mason, & O'Brien, 2000). Within a strategy-based view, there is no motivation for a reader to seek (and activate) a more distant causal explanation when a sufficient causal explanation is readily available.

The studies described in this section focused on necessary inferences—those that, when left unresolved, would presumably result in a break in local coherence. Whether they are activated, and how quickly they are activated, can be explained based primarily on memory-based factors (e.g., featural overlap, referential distance, causal relations, elaboration) without an appeal to any sort of strategic search process. We now consider the role of these same variables in the activation of elaborative inferences—those that are not required in order for coherence to be maintained.

Elaborative Inferences

In contrast to necessary inferences, *elaborative* inferences provide additional information beyond that which is explicitly stated in the text that is not required for the maintenance of coherence; elaborative inferences are derived from the activation of general world knowledge (Cook & Guéraud, 2005). Based on the assumption that readers have access to a limited amount of information in memory, reading researchers in the 1990s (e.g., Graesser et al., 1994; Singer et al., 1994) believed that elaborative inferences would not be activated unless the reader engaged in strategic processing. However, as Gerrig and O'Brien (2005) noted, readers have easy access to vast amounts of information via a passive resonance process; moreover, this information

Box 15.2 Example passage from Rizzella and O'Brien (1996)

Early Causal Explanation (Experiments 1 and 2)

Billy was walking home from school after playing a game of basketball. Billy looked for his keys to unlock the front door of his house. He searched everywhere but couldn't find the keys.

Early Causal Explanation Elaborated (Experiments 1 and 2)

He realized there was a big hole in his pocket. Now, he had no idea where to look. Billy shuddered when he recalled the warning his father gave him about being more responsible. His father told him that if he was not more responsible, he would ground Billy for an entire month.

Late Causal Explanation (Experiments 1 and 2)

Billy needed to find another way to unlock the door. In order to unlock the door, Billy broke a small window. The window fell in pieces on the ground.

Filler (Experiment 2)

Billy walked into the house and cleaned up the mess. Then he went into the kitchen for something to eat. Then he went into the living room to watch TV.

Consequent Event

He knew that once his father came home he would be in trouble.

can originate in either the episodic representation of the text or general world knowledge regardless of whether or not it is required to maintain local coherence.

O'Brien, Shank, Myers, and Rayner (1988; see also Garrod, O'Brien, Morris, & Rayner, 1990) provided one of the first demonstrations of passively activated elaborative inferences during reading. They monitored participants' eye movements as they read passages that provided either high or low contextual support for a specific category exemplar; in addition, this exemplar was either explicitly mentioned in the passage or merely implied (see Box 15.3 for an example). In the high-context condition, the two characters are making their way through a haunted house, whereas in the low-context condition, they are exploring a house in a new development when they spot a spider or an insect. Thus in both the high- and low-context conditions, the target concept is mentioned either explicitly or implicitly. Most important, in the implicit conditions it was not necessary for the maintenance of either local or global coherence for the reader to infer that the insect was a spider. A subsequent sentence in the text then referenced the spider. O'Brien et al. found that readers' gaze durations on *spider* were equivalent in both the high-context explicit and high-context implicit conditions, suggesting that readers had inferred and encoded *spider* in the high context condition, regardless of whether it had been explicitly mentioned or not.

The strongest test that elaborative inferences are activated through a passive resonance process comes from work examining the activation of predictive inferences. Because prediction is considered optional for comprehension, many researchers have argued that these inferences are less likely to occur online. However, McKoon and Ratcliff (1986) argued that predictive inferences are no different from necessary inferences: If contextual support is sufficient, related information in memory should be activated and lead to the activation of a predictive inference. Consider their classic example:

The director and the cameraman were ready to shoot close-ups when suddenly the actress fell from the fourteenth story.

Immediately following this text, participants were asked to provide a speeded yes/no recognition response to whether a probe word reflecting the predictive inference concept (e.g., *dead*) had appeared in the passage. In the case of this passage, the correct response to *dead* was "no," but McKoon and Ratcliff argued that participants would have difficulty rejecting this probe because the concept *dead* would be active in memory. Consistent with these predictions, recognition times were slower and less accurate than in a control condition, but only when tested immediately after the sentence ended. McKoon and Ratcliff argued that the predictive context primed the inference concept (*dead*) but that this inference was only minimally encoded. That is, when tested immediately after the supportive context, enough features of the concept *dead* were available to allow activation to be detected. With delay, these features decayed, and the reader was left with something broader and more general (e.g., something bad happened).

Cook, Limber, and O'Brien (2001) showed that predictive inferences are not simply primed by the immediately preceding context (i.e., the inference

Box 15.3 Example from O'Brien, Shank, Myers, and Rayner (1988)

High Context (Explicit/Implicit)

Chris and Randy were sneaking through a haunted house, brushing away cobwebs as they went. Suddenly, a fat, hairy eight-legged (spider/insect) dropped on Randy's shoulder.

Low Context (Explicit/Implicit)

Chris and Randy were exploring a house in a new development. They were taking notice of all the room sizes when Randy spotted a strange (spider/insect) in the corner.

Filler

He was so startled that he jumped in the air.

Reinstatement Sentence

He thought that the spider looked like a black widow.

evoking sentence) but by a combination of activation derived from the overall context, which in turn converges on inferential information in general world knowledge (see Cook & Guéraud, 2005). Because this convergence happens via passive memory activation processes, predictive inferences may be activated regardless of whether the supporting context is near or distant in the surface structure of the text. Consistent with McKoon and Ratcliff's (1986) idea of minimal encoding, Cook et al. found that, in most cases, the reader does not infer or encode predictive inferences in the form of a specific lexical item. For example, when reading a passage in which a young boy throws a rock that hits the side of a car, the reader does not encode the specific lexical item (e.g., *dent*), but rather shows evidence of encoding something much more general (e.g., some sort of damage happened). Lassonde and O'Brien (2009) extended this claim by demonstrating that the degree of inferential specificity was directly tied to the degree of contextual support; as contextual support increased, the degree of inferential specificity increased.

The demonstration that predictive inferences result from the combination of the overall context and not just the inference evoking sentence led to several investigations of how these two factors interact to influence inference activation. For example, Peracchi and O'Brien (2004) created conditions in which a protagonist's described characteristics were either consistent, inconsistent, or neutral with respect to a predicted event. In the sample passage in Box 15.4, Carol is described as either ill-tempered and quick to act (consistent condition), or as a peaceful individual opposed to physical violence (inconsistent condition). The text then continues with a description of an event (e.g., a rude customer complains about his food), and the reader is presented with either an inference-evoking sentence or a baseline sentence. Peracchi and O'Brien found that the inference-evoking sentence *Carol lifted the spaghetti above his head* facilitated naming times for the inference concept *dump*, indicating that this concept was activated in memory, but this only occurred when

Box 15.4 Example from Peracchi and O'Brien (2004) and Guéraud, Tapiero, and O'Brien (2008)

Consistent Condition

Carol was known for her short temper and her tendency to act without thinking. She never thought about the consequences of her actions, so she often suffered negative repercussions. She refused to let people walk all over her. In fact, she had just gotten a ticket for road rage. She decided she would never put up with anyone who was not nice to her.

Inconsistent Condition

Carol was known for her ability to peacefully settle any confrontation. She would never even think to solve her problems with physical violence. She taught her students and her own children how to solve problems through conversation. She believed this was an effective way to stop the increasing violence in the schools. Carol also helped other parents learn to deal with their anger.

Alternative Trait Condition

Carol had just come back to work after having had shoulder surgery. She needed to be careful whenever raising anything from a customer's table. Every time she did it, it would hurt so much that she thought she might faint. If she raised something too high, she was extremely uncomfortable all night. But usually, she asked for help when she needed to clear a table.

Filler

One particular night, Carol had an extremely rude customer. He complained about his spaghetti and yelled at Carol as if it was her fault.

Inference-Evoking Sentence

Carol lifted the spaghetti above his head.

Baseline Sentence

She lifted the spaghetti and walked away.

the preceding context also supported that inference (see also Rapp, Gerrig, & Prentice, 2001). In an extension of that work, Guéraud, Tapiero, and O'Brien (2008) demonstrated that the same inference-evoking sentence, when embedded in different contexts, could activate completely different inferences. For example, if Carol was described as ill-tempered and violence-prone and then *lifted the plate of spaghetti over a rude customer's head,* the inference *dump* was activated. However, if Carol was described as having shoulder problems (see alternative trait condition in Box 15.4), and then lifted the plate of spaghetti, the inference *pain* was activated instead.

Given the findings just described, one might be tempted to assume that the reader strategically considers the relevance of the preceding context to the information in the inference evoking sentence during the inference activation process. Cook, et al. (2014) showed, however, that the reactivation process that drives inference activation is dumb. They used passages similar to those used by Peracchi and O'Brien (2004), but added a condition in which the protagonist characteristics in the consistent condition were clearly outdated (see also O'Brien, Cook, & Guéraud, 2010; O'Brien, Cook, & Peracchi, 2004). For example, Carol is described as once having been ill-tempered and prone to act rashly, but not anymore; now she is peaceful. Cook et al. found that the inference *dump* was still activated (i.e., named more quickly following inference-evoking sentence than the baseline sentence) even when the supporting characteristics of the protagonist were described as outdated. These findings corroborate earlier demonstrations that predictive inferences result from the convergence of activation from the overall passage context and the the inference-evoking sentence via a dumb reactivation mechanism. We can conceive of no strategy-based model that would require a reader to actively generate inferences based on incorrect or outdated information, leading to inferences that are clearly incorrect. To date, the most simple and most parsimonious explanation of inference activation is the memory-based view. Limitations and challenges for this view will be addressed after discussing the role of strategy-based and memory-based processes in the role of maintaining global coherence.

Maintaining Global Coherence

As noted earlier, the maintenance of global coherence involves the establishment of relevant and important connections between currently processed text and portions of the text that were read earlier but that are no longer active in memory. Strategy-based models include the assumption that readers actively and strategically search memory for relevant connections with earlier read text. In contrast, memory-based models reject the notion of an active search; the reactivation of global information can only occur through a resonance process that is passive, dumb, and unrestricted.

Important evidence in support of a passive activation process guiding the maintenance of global coherence would be a demonstration that comprehension is sensitive to distant information even when a text is locally coherent and searches of long-term memory are therefore not necessary. The results of several series of experiments (e.g., Albrecht & O'Brien, 1993; Cook, Halleran, & O'Brien, 1998; Cook & O'Brien, 2014; Hakala & O'Brien, 1995; Kendeou, Smith, & O'Brien, 2013; O'Brien & Albrecht, 1992; O'Brien et al., 2010; O'Brien et al., 2004; O'Brien, Rizzella, Albrecht, & Halleran, 1998) have provided evidence that a passive resonance process results in the activation of related backgrounded information even when local coherence has been maintained. Consider the example in Box 15.5. Passages described a particular characteristic of a protagonist *(e.g., Mary was a strict vegetarian or Mary enjoyed junk food and ate at McDonald's).* Although the elaborated characteristic was not active in memory immediately prior to the target sentence (Myers, O'Brien, Albrecht, & Mason, 1994) and there was no local coherence break, reading times on the target sentence (*Mary ordered a cheeseburger and fries*) were longer in the inconsistent condition than in the consistent condition. Consistent with the assumptions of a passive resonance process, encoding the target sentence resulted in a signal being sent to all of memory, and memory traces related to Mary's eating habits resonated in response. When information indicating that Mary was a vegetarian was reactivated by this process, a global coherence break occurred; this disrupted reading as the reader attempted to re-establish coherence. Cook and O'Brien (2014) found that the reactivation process is mediated by the degree of featural overlap between the target sentence and the protagonist characteristics in general world knowledge. They found that the protagonist characteristics (e.g., *vegetarian*) were reactivated faster and had a stronger influence on comprehension when there was a high degree of overlap between those characteristics and the contents of the target sentence (*Mary ordered a* *cheeseburger) than when there was a low degree of overlap (Mary ordered a tuna salad).*

Evidence that the memory activation process is dumb can be found in O'Brien et al. (1998; see also O'Brien et al., 2010). Across several experiments, they included conditions in which the inconsistent elaboration was qualified to make it clear to the reader that Mary was no longer a vegetarian or that Mary had never been a vegetarian (see the qualified condition in Box 15.5). If readers actively search earlier portions of the discourse for relevant information, they should not reactivate Mary's vegetarianism because the overall content of the elaboration makes clear that Mary is no longer (or never was) a vegetarian; comprehension of the target sentence should not be disrupted. In contrast, if the memory activation process is dumb, then disconfirmed or false information should resonate and become

Box 15.5 Sample Passage Used by Albrecht and O'Brien (1993); Kendeou, Smith, and O'Brien (2013); and O'Brien, Rizzella, Albrecht, and Halleran (1998)

Introduction
Today, Mary was meeting a friend Joan for lunch. She arrived early at the restaurant and decided to get a table. After she sat down, she started looking at the menu.

Consistent Elaboration
This was Mary's favorite restaurant because it had fantastic junk food. Mary enjoyed eating anything that was quick and easy to fix. In fact, she ate at McDonalds at least three times a week. Mary never worried about her diet and saw no reason to eat nutritious foods.

Inconsistent Elaboration
This was Mary's favorite restaurant because it had fantastic health food. Mary, a health nut, had been a strict vegetarian for ten years. Her favorite food was cauliflower. Mary was so serious about her diet that she refused to eat anything that was fried or cooked in grease.

Qualified Elaboration
Mary remembered that at a recent party, Joan played a joke by telling people that Mary had been a strict vegetarian for ten years. Joan told everyone that Mary's favorite restaurant had fantastic health food. She said that Mary was a health nut and wouldn't eat anything that was fried or cooked in grease. She also claimed that Mary's favorite food was cauliflower.

One- (and Three-) Sentence Causal Explanation
This was Mary's favorite restaurant because it had fantastic health food. Mary, a health nut, has been a strict vegetarian for ten years. Her favorite food was cauliflower. Mary was so serious about her diet that she refused to eat anything which was fried or cooked in grease. She wasn't getting enough vitamins because of her diet so her doctor said she had to start eating meat. (Mary recently had blood work done. Her lack of iron was causing her to become anemic.)

Filler
After about ten minutes, Mary's friend Joan arrived. It had been a few months since they had seen each other. Because of this Mary and Joan had a lot to talk about and chatted for over a half hour. Finally, they signaled the waiter to come take their orders. They checked the menu one more time. Mary and Joan had a hard time deciding what to have for lunch.

Critical Sentences
Mary ordered a cheeseburger and fries.
She handed the menu back to the waiter.

Closing
Her friend didn't have as much trouble deciding what she wanted. She ordered and they began to chat again. They didn't realize there was so much for them to catch up on.

active simply because it is related to information in the target sentence. Consistent with this latter view, reading times on the target sentence continued to be slow in these qualified conditions.

Evidence that the activation process is unrestricted can be found in Cook et al. (1998). They rewrote the elaboration sections so that they described a secondary character (e.g., Joan is described as a vegetarian instead of Mary). Reading times on the target sentence (*Mary ordered a cheeseburger and fries*) were not slowed when Joan was the vegetarian; however, a subsequent probe study revealed that the target sentence led to reactivation of the characteristics now ascribed to Joan simply because they shared features in common with actions taken by the primary character. We can imagine no strategic search process that would include the assumption that the reader would actively access information about a character not in focus, especially when that information was not relevant.

Finally, the combination of the findings we have described might suggest that once particular information is encoded (e.g., *Mary is a vegetarian*), it is impossible to provide additional information that will eliminate the disrupting impact of that information on subsequent comprehension. However, the conditions under which this will occur must be limited; otherwise, readers would always be disrupted by initially encoded information that turned out to be incorrect. Further, the conditions under which the impact of the inconsistent information can be eliminated should be consonant with a passive activation mechanism. Kendeou et al. (2013) provided just such a test. They modified the inconsistent elaboration section so that it contained a causal explanation (either one or three causal sentences) for why the particular characteristic was no longer true (see the causal explanation conditions in Box 15.5). Causal information inherently provides a rich elaborated network of information that would compete with the inconsistent information for activation. Consistent with a passive activation process, not all related information is equal, and only a subset of related information will resonate sufficiently to return to active memory. Information that is more highly related or more richly interconnected will draw more activation. This, in turn, will reduce activation of other related information (i.e., interfere with the activation of other related information). Kendeou et al. found that when they added a one-sentence causal explanation, the impact of the inconsistent information was reduced: The inconsistent information was reactivated but did not disrupt reading. In contrast, when they added a three-sentence causal explanation, the impact of the inconsistent information was eliminated: the inconsistent information was not even activated. Thus, as the amount of qualifying information was systematically increased through causal explanations that created rich, interconnected networks of competing information, the impact and activation of inconsistent information was systematically reduced.

Summary and Challenges

As noted earlier, strategy-based and memory-based models of text comprehension make the same basic processing assumptions regarding the maintenance of local coherence. The critical factors that separate these models are the processes and conditions under which the reader gains access to information not currently active in memory but that, once available, contributes to the comprehension process. Most of the comparisons of these models can be narrowed down to two major issues: the activation of inferences (necessary and elaborative) and the process by which readers access earlier-presented portions of a text that are relevant to the comprehension of currently processed text (i.e., the maintenance of global coherence).

The brief review of the evidence regarding the activation of necessary and elaborative inferences strongly supports the memory-based view. Both general categories of inferences become available to the reader as a function of a signal emanating from currently read text that makes contact with earlier portions of the text, as well as general world knowledge. In fact, as Gerrig and O'Brien (2005) noted, because the activation process is the same, the need to define categories of inferences becomes less important; inferences are activated to the extent that information in active memory makes contact with related information from inactive portions of the text representation and general world knowledge. Our view is that this simple, passive, activation mechanism (i.e., resonance) forms the fundamental basis for the formation of any inference; and this process does not require any sort of strategic processing on the part of the reader.

However, there are important limitations of the memory-based view with respect to inferential processes. First, the findings that support that view have typically demonstrated the activation of simple one-concept inferences, or one causal idea. Inferences are often far more complex. Second, within the memory-based view, many higher order reading processes are held constant (e.g., the goals of the reader, the reader's life experiences, the need for the reader to explain and understand, the desire for the reader to be entertained), all of which surely influence the level and degree to which a reader will activate an inference. Graesser, Li, and Feng (2015) have argued that these higher order reader characteristics require strategic inferential processing (including strategic memory searches) to be adequately addressed. Indeed, Graesser et al. have suggested that whenever readers encounter cognitive disequilibrium, they will engage in inference generation, problem solving, reasoning, and other effortful cognitive activities in an attempt to resolve the impasse and restore cognitive equilibrium. We suspect that in principle this claim is correct. However, clear support for this position requires the development of models that contain sets of active procedures and productions rules. To date, no such model exists, and this remains the most serious challenge to proponents of strategic inference processing.

Finally, Goldman, McCarthy, and Burkett (2015) noted that much of the evidence regarding inferencing—especially from the memory-based perspective—has implicitly assumed that the reader adopts a literal stance in which the goal is to understand what the text is about. They note that what has been missing are comprehension situations in which the reader adopts an interpretive stance, in which the reader is oriented to what the text means beyond the situation of the specific text. A good example would be narratives with a moral, in which comprehension requires integrating what the text says with prior knowledge of a variety of ways that include knowledge of motivated human action, text genres, plot structures, moral and philosophical systems, and pragmatic aspects of the communicative event. Research designed to gain an understanding of the inferential processes involved when a reader adopts an interpretative stance would be both highly interesting and difficult to conduct. It would, though, also likely lead to a clear understanding of instances in which memory-based inferential processes would prove inadequate.

The contention surrounding the role of memory-based and strategy-based processes in the activation of inferences has produced a wealth of information regarding the conditions under which readers generate inferences and the role of inferences in the overall comprehension process (see Cook & O'Brien, 2015, for a review). However none of these findings truly discriminates between memory-based and strategy-based models of comprehension. Appeals to a memory-based account of inferential processing are made primarily on the basis of parsimony, not on evidence that directly refutes the possibility of strategic processing. Similarly, accounts of inferential processing that appeal to strategic processing can often be interpreted within a memory-based view. We contend that the cutting edge that separates memory-based models from strategy-based models involves the extent to which readers gain access to information from memory. Within memory-based models, information can only become available through a fast-acting, passive, and dumb activation mechanism. In contrast, within strategy-based models, readers continually and actively search memory for explanations that will guide comprehension and attempt to fully integrate current information with all prior relevant information. The evidence to date strongly supports the memory-based view.

One criticism concerning memory access as a strictly passive process is that it does not take into account the goals of the reader. Memory-based models (resonance, in particular) provide a mechanism that can be sensitized to the goals of the reader without the need for an active search process (e.g., search after meaning). As information becomes available, the reader has the ability to attend to, or focus on, any subset of that information that is most in line with his or her goals. The strength of the signal emanating from active memory will be greater for information that is in focus; this in turn, increases the likelihood that information related to what is in focus will eventually become available. This cyclical process continues until the reader has obtained sufficient information to achieve a level of coherence consonant with their standards of coherence. That is, the search after meaning principle is not actually a search of memory; it is an attentional process that operates on information that is made available through passive activation processes. Indeed, a strong version of a memory-based model (one to which we subscribe) rejects any form of active memory search. Active searches for information would involve such activities as consciously and physically looking back in a text, or getting up from one's desk to find another physical source (e.g., another book from one's bookshelf).

The challenge for strategy-based models of comprehension is to demonstrate the need for a strategic search that is triggered by the reader actively tracking specific text characteristics (e.g., time, space) independent of whether the text is locally coherent. Even with strong evidence for such a strategic search, it would still be necessary to describe the characteristics and constraints of a strategic search process. For example, would such a search be unrestricted or directed? If the search is unrestricted, then it takes on many characteristics of a passive activation mechanism. On the other hand, if the search is directed, it is difficult to envision the cognitive machinery necessary to explain such a search without invoking an executive function that then directs the search. Without any process or mechanism to guide the concept of a strategic search, such models have no predictive or explanatory power.

There have been some attempts to develop hybrid views of memory search processes during reading, in which the initial memory search process is assumed to be passive. If that passive process fails to produce sufficient information, the reader switches to an active search of memory (e.g., Long & Lea, 2005; van den Broek et al., 1996). We see little value in such two-stage passive + active search mechanisms because there is no way of knowing when a reader might switch from a passive to a strategic search. And once that switch occurs, the models must rely on the same executive search function that governs memory access in strategy-based models. Although these models were proposed in the spirit of a compromise between memory-based and strategic-based models, unfortunately the addition of a second strategic search stage weakens these models to the point of having no predictive value over models that assume only a strategic search.

Conclusion

The models described in the previous sections range in the extent to which they view the processes that occur during comprehension as passive or strategic, and attempts to integrate passive and strategic processes into a single model have largely fallen short. Undoubtedly, a complete model of comprehension will involve both passive and strategic processes. To date, however, the passive components of the comprehension process have been more clearly identified, specified, and tested (primarily inference activation and memory access). This is probably because it is much easier to design experiments that isolate passive components, especially passive memory components that are based on well-established findings within basic memory research. It has been when research findings could not be easily explained by passive processes that researchers have appealed to strategic processing (e.g., Long & Lea, 2005; Singer et al. 1994). We contend that such appeals are akin to admitting that there is no existing explanation for what processes or factors are driving comprehension beyond a certain point.

We do not deny that there has been some systematic experimental work conducted with the goal of clarifying readers' strategic processes (e.g., Magliano & Radvansky, 2001; Rapp & Gerrig, 2002). Nevertheless, in the current state of the field, we know a great deal about the passive processes involved in comprehension and very little about the strategic processes. Future research on passive processes is likely to only refine much of what we already know. Instead, we believe that the most interesting questions for future research are those that advance our understanding of the mechanisms and limitations of strategic processing (e.g., see Goldman et al., 2015). Unfortunately, however, these questions are also likely to be the most challenging to address.

References

Albrecht, J. E., & Myers, J. L. (1995). Role of context in accessing distant information during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21,* 1459–1468.

Albrecht, J. E., & Myers, J. L. (1998). Accessing distant text information during reading: Effects of contextual cues. *Discourse Processes*, *26,* 87–107.

Albrecht, J. E., & O'Brien, E. J. (1991). Effects of centrality on retrieval of text-based concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17,* 932–939.

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19,* 1061–1070.

Albrecht, J. E., & O'Brien, E. J. (1995). Goal processing and the maintenance of global coherence. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 263–278). Hillsdale, NJ: Erlbaum.

Cook, A. E. (2014). Processing anomalous anaphors. *Memory & Cognition*, *42,* 1171–1185.

Cook, A. E., & Guéraud, S. (2005). What have we been missing? The role of general world knowledge in discourse processing. *Discourse Processes*, *39,* 265–278.

Cook, A. E., Halleran, J. G., & O'Brien, E. J. (1998). What is readily available during reading? A memory-based view of text processing. *Discourse Processes*, *26,* 109–129.

Cook, A. E., Lassonde, K. A., Splinter, A. F., Guéraud, S., Stiegler-Balfour, J. J., & O'Brien, E. J. (2014). The role of relevance in activation and instantiation of predictive inferences. *Language, Cognition, and Neuroscience, 29,* 244–257.

Cook, A. E., Limber, J. E., & O'Brien, E. J. (2001). Situation-based context and the availability of predictive inferences. *Journal of Memory and Language*, *44,* 220–234.

Cook, A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes*, *51,* 26–49.

Cook, A. E., & O'Brien, E. J. (2015). Passive activation and instantiation of inferences during reading. In E. J. O'Brien,

A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading*. Cambridge, England: Cambridge University Press.
Corbett, A. T., & Chang, F. R. (1983). Pronoun disambiguation: Accessing potential antecedents. *Memory & Cognition*, *11,* 283–294.
Dell, G. S., McKoon, G., & Ratcliff, R. (1983). The activation of antecedent information during the processing of anaphoric reference in reading. *Journal of Verbal Learning and Verbal Behavior*, *22,* 121–132.
Garrod, S., O'Brien, E. J., Morris, R. K., & Rayner, K. (1990). Elaborative inferencing as an active or passive process. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16,* 250–257.
Gernsbacher, M. A. (Ed.). (1990). *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.
Gerrig, R. J., & McKoon, G. (1998). The readiness is all: The functionality of memory-based text processing. *Discourse Processes*, *26,* 67–86.
Gerrig, R. J., & O'Brien, E. J. (2005). The scope of memory-based processing. *Discourse Processes*, *39,* 225–242.
Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91,* 1–67.
Goldman, S. R., McCarthy, K., & Burkett, C. (2015). Interpretive inferences in literature. In E. J. O'Brien, A. E. Cook, & R. F. Lorch, Jr. (Eds.). *Inferences during reading*. Cambridge, England: Cambridge University Press.
Graesser, A. C., Li, H., & Feng, S. (2015). Constructing inferences in naturalistic reading contexts. In E. J. O'Brien, A. E. Cook, & R. F. Lorch, Jr. (Eds.), *Inferences during reading*. Cambridge, England: Cambridge University Press.
Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101,* 371–395.
Guéraud, S., Tapiero, I., & O'Brien, E. J. (2008). Context and the activation of predictive inferences. *Psychonomic Bulletin & Review*, *15,* 351–356.
Hakala, C. M., & O'Brien, E. J. (1995). Strategies for resolving coherence breaks in reading. *Discourse Processes*, *20,* 167–185.
Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95,* 528–551.
Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness* (Vol. 6). Cambridge, MA: Harvard University Press.
Johnson-Laird, P. N., & Garnham, A. (1980). Descriptions and discourse models. *Linguistics and Philosophy*, *3,* 371–393.
Keenan, J. M., Baillet, S. D., & Brown, P. (1984). The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior*, *23,* 115–126.
Kendeou, P., Smith, E. R., & O'Brien, E. J. (2013). Updating during reading comprehension: Why causality matters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39,* 854–865.
Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95,* 163–182.
Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, England: Cambridge University Press.
Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85,* 363–394.
Kintsch, W., & Vipond, D. (1979). Reading comprehension and readability in educational practice and psychological theory. In L. G. Nilsson (Ed.), *Perspectives on memory research* (pp. 329–365). Hillsdale, NJ: Erlbaum.
Klin, C. M., Guzmán, A. E., Weingartner, K. M., & Ralano, A. S. (2006). When anaphor resolution fails: Partial encoding of anaphoric inferences. *Journal of Memory and Language*, *54,* 131–143.
Klin, C. M., Weingartner, K. M., Guzmán, A. E., & Levine, W. H. (2004). Readers' sensitivity to linguistic cues in narratives: How salience influences anaphor resolution. *Memory & Cognition*, *32,* 511–522.
Lassonde, K. A., & O'Brien, E. J. (2009). Contextual specificity in the activation of predictive inferences. *Discourse Processes*, *46*, 426–438.
Levine, W. H., Guzmán, A. E., & Klin, C. M. (2000). When anaphor resolution fails. *Journal of Memory and Language*, *43,* 594–617.
Long, D. L., & Lea, R. B. (2005). Have we been searching for meaning in all the wrong places? Defining the "search after meaning" principle in comprehension. *Discourse Processes*, *39,* 279–298.
Magliano, J. P., & Radvansky, G. A. (2001). Goal coordination in narrative comprehension. *Psychonomic Bulletin & Review*, *8,* 372–376.
McKoon, G., & Ratcliff, R. (1980). The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning and Verbal Behavior*, *19,* 668–682.
McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12,* 82–91.
McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99,* 440–466.
McKoon, G., & Ratcliff, R. (1998). Memory-based language processing: Psycholinguistic research in the 1990s. *Annual Review of Psychology*, *49,* 25–42.
Murdock, B. B., Jr. (1982). A theory for the storage and retrieval of items and associative information. *Psychological Review*, *89,* 187–203.
Myers, J. L., Cook, A. E., Kambe, G., Mason, R. A., & O'Brien, E. J. (2000). Semantic and episodic effects on bridging inferences. *Discourse Processes*, *29,* 179–199.
Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, *26,* 131–157.
Myers, J. L., O'Brien, E. J., Albrecht, J. E., & Mason, R. A. (1994). Maintaining global coherence during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20,* 876–886.
Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language*, *26,* 453–465.
O'Brien, E. J. (1987). Antecedent search processes and the structure of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13,* 278–290.
O'Brien, E. J., & Albrecht, J. E. (1991). The role of context in accessing antecedents in text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17,* 94–102.
O'Brien, E. J., & Albrecht, J. E. (1992). Comprehension strategies in the development of a mental model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18,* 777–784.
O'Brien, E. J., Albrecht, J. E., Hakala, C. M., & Rizzella, M. L. (1995). Activation and suppression of antecedents during

reinstatement. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21,* 626–634.
O'Brien, E. J., Cook, A. E., & Guéraud, S. (2010). Accessibility of outdated information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36,* 979–991.
O'Brien, E. J., Cook, A. E., & Peracchi, K. A. (2004). Updating situation models: A reply to Zwaan and Madden (2004). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30,* 289–291.
O'Brien, E. J., Duffy, S. A., & Myers, J. L. (1986). Anaphoric inference during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12,* 346–352.
O'Brien, E. J., & Myers, J. L. (1987). The role of causal connections in the retrieval of text. *Memory & Cognition*, *15,* 419–427.
O'Brien, E. J., & Myers, J. L. (1999). Text comprehension: A view from the bottom up. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 35–53). Hillsdale, NJ: Erlbaum.
O'Brien, E. J., Plewes, P. S., & Albrecht, J. E. (1990). Antecedent retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16,* 241–249.
O'Brien, E. J., Raney, G. E., Albrecht, J. E., & Rayner, K. (1997). Processes involved in the resolution of explicit anaphors. *Discourse Processes*, *23,* 1–24.
O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: A memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1200–1210.
O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur online? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14,* 410–420.
Peracchi, K. A., & O'Brien, E. J. (2004). Character profiles and the activation of predictive inferences. *Memory & Cognition*, *32,* 1044–1052.
Rapp, D. N., & Gerrig, R. J. (2002). Readers' reality-driven and plot-driven analyses in narrative comprehension. *Memory & Cognition*, *30,* 779–788.
Rapp, D. N., Gerrig, R. J., & Prentice, D. A. (2001). Readers' trait-based models of characters in narrative comprehension. *Journal of Memory and Language*, *45,* 737–750.
Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85,* 59–108.
Rizzella, M. L., & O'Brien, E. J. (1996). Accessing global causes during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22,* 1208–1218.
Sanford, A. J., & Garrod, S. C. (1998). The role of scenario mapping in text comprehension. *Discourse Processes*, *26,* 159–190.
Sanford, A. J., & Garrod, S. C. (2005). Memory-based approaches and beyond. *Discourse Processes*, *39,* 205–224.
Singer, M. (1993). Causal bridging inferences: Validating consistent and inconsistent sequences. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *47,* 340–359.
Singer, M., & Ferreira, F. (1983). Inferring consequences in story comprehension. *Journal of Verbal Learning and Verbal Behavior*, *22,* 437–448.
Singer, M., Graesser, A. C., & Trabasso, T. (1994). Minimal or global inference during reading. *Journal of Memory and Language*, *33,* 421–441.
Singer, M., & Halldorson, M. (1996). Constructing and validating motive bridging inferences. *Cognitive Psychology*, *30,* 1–38.
Singer, M., Halldorson, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. *Journal of Memory and Language*, *31,* 507–524.
Trabasso, T., Secco, T., & van den Broek, P. W. (1984). Causal cohesion and story coherence. In H. Mandl, N. L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 83–111). Hillsdale, NJ: Erlbaum.
van den Broek, P., Risden, K., Fletcher, C. R., & Thurlow, R. (1996). A "landscape" view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 165–187). Hillsdale, NJ: Erlbaum.
van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.
Zwaan, R. A. (1999). Five dimensions of narrative comprehension: The event-indexing model. In S. R. Goldman, A. C. Graesser, & P. van den Broek, (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 93–110). Hillsdale, NJ: Erlbaum.
Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123,* 162–185.

CHAPTER

16 The Role of Words in Chinese Reading

Xingshan Li, Chuanli Zang, Simon P. Liversedge, *and* Alexander Pollatsek

Abstract

The Chinese writing system (and the underlying language) is different from European writing systems (and their underlying languages) in many ways. The most obvious difference is the nonalphabetic form of the Chinese written language, although there are also differences in representations of lexicality, grammaticality, and phonological form. This chapter focuses on issues associated with the nonalphabetic nature of written Chinese and the fact that words are not demarcated by spaces. Despite the surface differences in the orthographies between European languages and Chinese, there is considerable evidence that the word is a salient unit for Chinese readers. Properties of words (such as word frequency) as well as character properties (such as character frequency) affect measures of reading time and also affect where eye movements are targeted when passages of text are being read for meaning.

Key Words: word segmentation, interword spacing, saccade target selection, eye movements, Chinese reading

Unlike in many other writing systems, there are no spaces in Chinese text to separate words. Text written in Chinese is formed by strings of equally spaced box-like symbols called characters (Chinese is now standardly read from left to right). The fact that Chinese is unspaced, as distinct from the majority of languages that are spaced, may initially appear somewhat surprising. However, some alphabetic languages such as English did not use spaces to demark word boundaries until around 1000 A.D. (Boorstin, 1983). It is generally believed that introducing spaces to mark word boundaries facilitates reading by providing the reader with explicit visual markers of word beginnings and endings (see reviews by Rayner, 1998, 2009; Zang, Liversedge, Bai, & Yan, 2011). Word spacing also reduces the extent to which adjacent words in text laterally mask each other (Rayner, 1998, 2009; Zang et al., 2011). For these reasons, word spacing is considered to benefit readers (indeed, if spaces are removed from English text, then reading becomes far less efficient; Rayner, Fischer, & Pollatsek, 1998). In this chapter we begin by briefly describing the nonalphabetic nature of the written form of Chinese and the fact that words are not demarcated by spaces. We then review studies exploring how Chinese readers identify word boundaries and how these properties of Chinese affect eye-movement behavior during reading unspaced Chinese sentences.

An interesting question is why word spacing has not been adopted in the written form of Chinese. We do not have any definite answers to this question, although we can speculate as to potential reasons why this may be the case. There may be pragmatic historical reasons why spaces were avoided. For example, ancient Chinese text was often written on bamboo or carved into stone, and it may have been necessary to avoid spaces to make maximum use of the available space. Another historical reason is that in ancient Chinese text each character conveyed a particular aspect of meaning, and consequently text was read one character at a time. In fact, there was no term for a word in Chinese until the concept was imported from the West at the beginning of the twentieth century (Packard, 1998). Also, ancient written Chinese did not contain punctuation to facilitate segmentation at a level even coarser than the word. There is a final factor that we believe may have also contributed to why spacing has not been incorporated into Chinese. Words in Chinese are quite short, as measured by their number of constituent characters. Also, the variance in word length is reduced relative to word length variability in alphabetic languages (as measured by the number of constituent letters). Approximately 97% of words in Chinese are one or two characters in length (token frequency; Lexicon of Common Words in Contemporary Chinese Research Team, 2008). To this extent, the number of potential sites within a character string at which word segmentation might occur is significantly reduced in Chinese, and therefore decisions about where to segment text to form word boundaries might be less of a challenge to Chinese readers than is the case in languages such as English. Thus word spacing may have been less of a necessity for efficient reading in Chinese.

Even without explicit word boundary cues, Chinese readers appear to have little difficulty reading Chinese. If the word is a significant linguistic unit in Chinese—and we will present evidence that it is—Chinese readers have to depend on other mechanisms to segment words in reading. How the lack of interword spaces affects Chinese reading and how Chinese readers segment sentences into words is still far from fully understood. Before we review studies exploring how Chinese readers identify word boundaries during reading within strings of characters forming sentences, we should note that Chinese is not the only written language that doesn't have a space or some other demarcation symbol to mark the boundary between words. For example, Japanese and Thai do not have spaces to mark word boundaries. Thus the studies we review might suggest some phenomena observed in studies of processing in those written systems. Moreover, some alphabetic writing systems such as German and Finnish have long compound words that are complex multimorphemic units written with no spaces separating the component words, and readers of those languages might use similar mechanisms to process such words when they read.

What Is a Word in Chinese?

Chinese linguists define a word as the minimal linguistic unit with a specific meaning and pronunciation that could be used alone to constitute a sentence or as a grammatical component on its own (Hoosain, 1992). However, a Chinese word can be composed of one or more morphemes. For reading, a word is composed of characters, with each character corresponding to a morpheme. Among the 56,008 listed words that are included in one published source (Lexicon of Common Words in Contemporary Chinese Research Team, 2008), 6% are one-character words, 72% are two-character words, 12% are three-character words, and 10% are four-character words. Less than 0.3% of Chinese words are longer than four characters (based on type frequency). When word tokens are taken into account, 70.1% of words are one-character words, 27.1% are two-character words, 1.9% are three-character words, 0.8% are four-character words, and 0.1% are longer than four characters. There are more than 5,000 Chinese characters (Hoosain, 1992), and these differ in their complexity (varying from one to more than twenty strokes). A single character can be a part of different words when combined with other characters. Most Chinese characters are pronounced identically when they comprise different words; however, some characters are pronounced differently when they appear in different words.

Because of the lack of explicit markers to indicate word boundaries in the Chinese writing system, Chinese readers do not always agree with each other on the location of the boundaries between the words of a sentence in text (Hoosain, 1992; Liu, Li, Lin, & Li, 2013). For some words readers almost always agree with each other, but for other words they do not. This inconsistency has caused significant difficulty for researchers designing artificial intelligence systems that attempt to understand Chinese text. To overcome difficulties associated with word segmentation, a Chinese national standard has been established that stipulates word segmentation for artificial information processing systems. Word segmentation according to this system is similar to that which would be derived on the basis of standard linguistic definitions of Chinese words. This standard lists some basic rules for segmenting Chinese text into words (National Standard GB/T 13715–92, 1992). Chinese readers, however, do not always follow the national standard when they are required to segment written sentences into words. Liu et al. (2013) asked Chinese readers to put a slash at word boundary positions in sentences. They then calculated the proportion of subjects that put a slash after each character, noting that the proportion should be 1 or 0 if there was complete agreement among subjects regarding word boundaries. The Chinese readers did not always agree with the national standard when they were required to parse text into words. Their segmentation was influenced by the syntactic categories of consecutive words. Specifically, they were more likely to combine function words (e.g., auxiliary words 的 or 地) with content words (e.g., nouns, verbs, adjectives, or pronouns) to form single-word units. Furthermore, most readers agreed that the numerals (一 'one') and quantifiers (种 'type of'), as well as the verb (躺 'lie') and the preposition (在 'down'), should be combined with other characters as single-word units. Finally, readers usually considered consecutive nouns (e.g., the phrase 森林公园 'forest park') as a whole word. Generally, Chinese readers tended to chunk single words into larger informational units during word segmentation. Although the task used by Liu et al. (2013) is artificial, in that it may not necessarily reflect how subjects segment text into words when they read normally, it is not necessarily more artificial than the judgment of linguists.

The Psychological Reality of Words in Chinese Reading

Given that the word unit in written Chinese text is not clearly demarcated and that there is some ambiguity concerning word boundaries (Hoosain, 1992), is the word a meaningful linguistic unit of information in processing written language in unspaced Chinese text? Furthermore, does the word unit play as central a role in eye movement control during reading for Chinese readers as it does for English readers? The earliest robust evidence of the importance of words as a visual unit in English text reading came from studies investigating the word superiority effect (Reicher, 1969; Wheeler, 1970), such that letter identification is facilitated when the letter is part of a word as compared with when it is embedded in a series of letters that do not form a word or when it is shown in isolation. Similarly, research on Chinese (e.g., Cheng, 1981) has demonstrated that Chinese characters were identified more accurately in a briefly presented word than in a string of characters that did not constitute a word. Such a phenomenon indicates that Chinese characters belonging to a word can be effectively perceived as a unit.

Li, Rayner, and Cave (2009) further investigated how word boundaries affected character perception in Chinese reading and found word boundary effects. Participants were briefly shown four Chinese characters in a horizontal row and were asked to report as many characters as possible. These four characters constituted a four-character word in the one-word condition or two two-character words in the two-word condition. Li et al. found that participants usually reported the four-character word in the one-word condition, but could usually only report the first two-character word in the two-word condition even though there were four syllables to be reported in both conditions. This result demonstrates that word segmentation influences character recognition: The word boundary in the two-word condition induced serial processing, whereas the lack of it in the one-word condition induced parallel processing of the entire string. In sum, the evidence indicates that word segmentation is a necessary and important procedure in Chinese reading.

Lower-Level Word Segmentation Cues Benefit Reading

Recently, there has been a great deal of interest in investigating how readers use lower-level word segmentation cues like spaces to segment and identify words when reading Chinese text (see Zang et al., 2011 for a review). Given the disagreements among Chinese readers about word boundaries, researchers usually prepare their experimental stimuli very carefully. They only use dictionary-defined words. Any character strings for which there is ambiguity concerning their word status are usually avoided or discarded. In addition, after stimulus construction, to confirm that there is general agreement among Chinese readers as to word boundaries, a prescreening test is usually conducted. Bai, Yan, Liversedge, Zang, and Rayner (2008) found that when Chinese adult readers read sentences with spaces inserted between words (or when highlighting was used to demarcate words), they read them as easily as normal unspaced Chinese text. However, when spaces were inserted (or highlighting was used) between characters of a word (in a character segmentation condition) or randomly within words (in a nonword condition), reading was slowed. The results suggest that inserting spaces between the characters of a word in Chinese text slows reading and suggests that inserting spaces between words facilitates, but that the facilitative effect is negated by the fact that the spaces are novel.

Later studies showed that inserting spaces between words could help beginning readers of Chinese to read more efficiently and to learn new words (Blythe et al., 2012; Zang, Liang, Bai, Yan, & Liversedge, 2013). Blythe et al. recorded adults' and second-grade children's (mean age 8.3 years, range 7 to 10 years) eye movements as they read novel two-character words (where both characters were known but their combination formed a new word whose meaning could not be derived from the meanings of the constituent characters). During the learning session of the experiment, subjects read these words in explanatory sentences. Importantly, half of the subjects learned the new words in sentences with word spacing, while the other half learned the new words in unspaced sentences. Subjects returned for a test session on another day where they read the new words again in a different set of sentences. In the test session, all subjects read unspaced text. In the learning session, participants in the spaced groups read the new words more quickly than the matched control participants in the unspaced groups. More importantly, children, but not the adults in the spaced group, maintained this benefit in the test session while reading unspaced text. Blythe et al. argued that the spacing manipulation allowed the children either to form stronger connections between the two characters' lexical representations and the corresponding novel orthographic lexical representation of the word or to form a more fully specified novel lexical representation of the word itself (i.e., form a representation for each new word that is specified semantically with novel connections between that semantic unit and phonological and orthographic representations; see Perfetti, Liu, & Tan, 2005, for a review). Follow-up research also showed that word spacing can be useful for beginning readers of Chinese as a second language (Bai et al., 2013; Shen et al., 2012).

Other studies (Li & Shen, 2013; Liu & Li, 2014) explored whether inserting a space before or after a word facilitates the processing of that word during Chinese reading. When a Chinese word (word *n*) is recognized, its boundaries on both sides are known. Thus, inserting a space before the word to the right (word *n* + 1) does not provide additional word boundary information given that its left boundary has been determined when word *n* is recognized. However, inserting a space after word *n* + 1 provides information about its right boundary, which helps readers segment it from the text before recognizing it. Consistent with these assumptions, Li and his colleagues found that inserting a space after a word facilitated its processing but that inserting a space before a word did not facilitate processing and in fact may even interfere with its integration into sentential meaning as indicated by total reading times. Therefore, the position of a space may affect the ease of word identification differentially.

Word Properties Influence Reading

So far, our descriptions of studies have just included global measures of reading such as comprehension scores or total reading time. However, most of the studies that we discuss in this section employed more detailed eye-movement measures to get local measures of online processing while people read text. Several measures of fixation time on a target region of text are commonly employed (target regions may be a character or a word). The three most common are *first-fixation duration*, the duration of the first fixation on a region of text; *first-pass time*, the sum of all fixation durations on a region of text until it is exited to the right or left; and *total fixation time*, the sum of all fixation durations on a region of text (including fixations after regressions back to the region). In all cases, it is assumed that the reader entered the region of text for the first time from the left and that the script being discussed goes from left to right. Other common eye-movement measures are the size and direction of the jump (*saccade*) from fixation to fixation. Backward saccades are called *regressions*.

Eye-movement studies investigating Chinese reading have shown that a word's linguistic properties, such as its frequency and predictability, affect both the number and the duration of the fixations it receives, even when the properties of the characters that constitute the word have been controlled (see Zang et al., 2011, for a review). For example, first-pass reading times on high-frequency words are significantly shorter than on low-frequency words (Liversedge et al., 2014; Yan, Tian, Bai, & Rayner, 2006; Yang & McConkie, 1999), and first-pass reading times on less predictable words are significantly longer than on more predictable words (Rayner, Li, Juhasz, & Yan, 2005; Wang, Pomplun, Chen, Ko, & Rayner, 2010); Furthermore, readers skip more predictable words more often than less predictable words (Rayner et al., 2005) and skip high-frequency words more often than low-frequency words (Yan et al., 2006). Yan et al. also found that the character frequency effect was modulated by word frequency, being evident only when word frequency was low but negligible when it was high. A possible explanation is that, when a word is frequently used, it is accessed as a single entity in the reader's mental lexicon. In contrast, when it is infrequently used, the word needs to be accessed via the individual characters, and, as a consequence, an effect of character frequency is found. Thus, to some extent, the properties of a word can modulate processing of its constituent characters.

Apart from the linguistic properties of Chinese words, a great deal of research demonstrates that low-level visual information associated with a Chinese word, such as its visual complexity (Liversedge et al., 2014; Yang & McConkie, 1999) and length (Li, Liu, & Rayner, 2011; Li & Shen, 2013), affects lexical identification and saccade programming during reading. Note that in these studies the properties of the words' constituent characters were controlled. For example, Li et al. (2011) reported that saccades leaving a four-character word were longer than saccades leaving a two-character word. This result indicates that the length of the fixated word affects subsequent saccade planning in reading. Taken together, these studies suggest that word properties, either at lower or higher levels, affect eye-movement behavior during Chinese reading. They further demonstrate the importance of word-based processing in Chinese reading.

Characters Belonging to a Word Are Processed as a Unit

More recently, a series of studies using a variety of paradigms have provided direct evidence that Chinese characters belonging to a word are processed as a unit (e.g., Li, Bicknell, Liu, Wei, & Rayner, 2014; Li, Gu, Liu, & Rayner, 2013; Li & Pollatsek, 2011; Li, Zhao, & Pollatsek, 2012). Li et al. (2013) employed a novel variation of the moving window paradigm to test whether reading performance was better when characters belonging to a word were presented simultaneously than when they were not. In *the moving window paradigm* (see Schotter & Rayner, this volume), the area of text around fixation is normal and all other text is replaced by some meaningless alternative material. When the eyes move, the display changes so that this statement now applies to the display around the new fixation point. All of the words in the Li et al. sentences were two characters long, and the size of the moving window was also two characters. Thus only two characters were available to be processed on any particular fixation. All the characters outside the window were masked by the symbol ※. In Experiment 1, the two characters in the window constituted a word in the *word-window* condition but did not in the *nonword-window* (or character) condition (see Figure 16.1). Li et al. found that readers made more and longer fixations when they could not see the characters belonging to a word simultaneously compared with when they could. That is, there was a cost when both characters belonging to a word were not available to be processed simultaneously.

In normal Chinese text, when the characters belonging to a word are shown on different lines, readers are not able to process them as being constituent characters of a word simultaneously. Li et al. (2012) examined whether dividing a word across two lines interferes with Chinese reading. In the divided-word condition of the experiment, the last word in a line was shown with one of the characters at the end of one line and the other character at the beginning of the next. In the word boundary condition, the target word was always shown at the end of a line, and no word was shown crossing two lines. Li et al. found that reading time was longer in the divided-word condition than the word boundary condition. The data thus indicated that characters belonging to a word were easier to process when they were presented on a single line than when they

Sentence	观众正在台下耐心等待演员出场
Experiment 1	
	Word-window condition
Example 1	※※※※※※耐心※※※※※※
	*
Example 2	※※※※※※耐心※※※※※※
	*
	Nonword-window condition
Example 1	※※※※※下耐※※※※※※※
	*
Example 2	※※※※※下耐※※※※※※※
	*

Fig. 16.1 An example of the stimuli used in the study by Li et al. (2013). The English translation of the sentence is 'The audiences are patiently waiting for actors to come on the stage.' The symbol * indicates the fixation point.

were presented on adjacent lines. Again, these findings provide evidence that a word is normally processed as a unit in Chinese reading.

Finally, Li et al. (2014) evaluated the effects of various word properties on eye movements during Chinese reading to determine whether these word properties have effects above and beyond what could be predicted by the properties of their component characters. These word properties included the length, frequency, and predictability of the current, previous, and following word, and the character properties included the frequency and complexity of a range of characters around the point of fixation. Participants' eye movements were recorded when they read sentences. Li et al. found that the effects of the properties of the current, prior, and following words were strikingly similar in Chinese to those observed for word-based alphabetic languages on a range of eye-movement measures. In addition, Li et al. revealed a rich pattern of effects of character properties. Crucially, the effects of word frequency, word length, and predictability were highly reliable with and without character properties included in the same model. However, when the word properties were removed from this model, its prediction for the data became significantly worse. These findings indicate an underlying word-based core to reading that appears to be shared between Chinese and alphabetic language scripts.

The preceding discussion should not be taken to imply that Chinese readers process only the fixated word. Instead, there is extensive evidence that Chinese readers also process word(s) in the parafovea (see Zang et al., 2011, for a review). More generally, readers in all languages extract parafoveal information beyond the fixated word that facilitates processing on subsequent fixations. Yen, Tsai, Tzeng, and Hung (2008) used *the boundary paradigm* (Rayner, 1975; see Schotter & Rayner, this volume) to investigate whether parafoveal word recognition occurs during Chinese reading. In this paradigm, an invisible boundary is positioned just to the left of a target word. Before the reader crosses the boundary, there is typically an initial display stimulus (preview) that is different from the target word. When the eyes cross the boundary, the preview is replaced by the target word. Reading times on the target word are significantly shorter when the target is identical to the preview than when it is different. This is usually referred to as *parafoveal preview benefit* (Liversedge & Findlay, 2000; Rayner, 1998, 2009). By manipulating the characteristics of the preview in relation to those of the target word, one may observe differences in readers' oculomotor behavior and infer which characteristics of a parafoveal word are processed before it is fixated (see Figure 16.2). Yen et al. manipulated whether the preview was a real word or a pseudoword. They found that targets with word previews, even those that were contextually inappropriate and semantically unrelated, were more likely to be skipped than those with pseudoword previews. This result implies that the word preview was processed and identified as a word (as opposed to a pseudoword) in the parafovea.

Cui, Drieghe, et al. (2013) further investigated parafoveal processing across different lexical constituents in the reading of Chinese sentences. The experiment included three types of two-character Chinese target strings: a monomorphemic word, a compound word, or an adjective-noun word pair. The preview of the second character of that string (e.g., 瑰 in the string 玫瑰) was either identical to that character (i.e., 瑰) or was a dissimilar pseudocharacter (e.g., 柆). The pseudocharacters very closely resembled real characters but were

Sentence	老师教导我们永远不要**忘记**这段历史
Before the boundary	老师教导我们永远不要**界料**这段历史
	* * * * * |
Cross the boundary	老师教导我们永远不要**忘记**这段历史
	| *
After the boundary	老师教导我们永远不要**忘记**这段历史
	| * *

Fig. 16.2 An example of a boundary paradigm. The symbol * indicates the fixation point. The invisible boundary that triggers the display change is marked with a vertical line. When the reader's eyes cross the boundary, the preview word (e.g., a pseudoword 界料 in this example) changes to the target word (忘记 'forget'). The English translation of the sentence is 'The teacher taught us that we should never *forget* this period of history'.

meaningless. The analyses of Cui, Drieghe, et al. on the first constituent (but not on the second constituent or the whole target string) showed that a pseudocharacter preview of the second character of the string increased fixation durations on the first character of that string for monomorphemic words (but not for compound words or phrases). This result indicates that the two constituents of monomorphemic words can be processed in parallel and that the morphological structure of a Chinese word, or how predictable the second character is given the first character, modulates how the word is processed in reading (see also Cui, Yan, et al., 2013).

To summarize, the studies we have discussed show that the word plays an important role during Chinese reading and that preventing Chinese readers from processing the component characters of words simultaneously hinders reading efficiency. The findings also provide evidence that words have a psychological reality during Chinese reading. That is, word representations are important and play a functional role in the process of written language comprehension. To this extent, there is fundamental similarity between Chinese reading processes and processes that underlie reading of alphabetic language scripts. Indeed, the word-based E-Z Reader model of eye movement control (Pollatsek, Reichle, & Rayner, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998), which was developed to model eye-movement behavior for skilled readers of alphabetic languages, was extended to Chinese readers by Rayner, Li, and Pollatsek (2007). They showed that the model accounted for fixation durations and word-skipping rates during Chinese reading quite well.

Mechanisms of Word Segmentation

We have summarized results showing that words are important during Chinese reading. However, there are no spaces to mark word boundaries in Chinese text. Without spaces, how do Chinese readers segment words? This question seems like a chicken-and-egg problem. On the one hand, in order for word segmentation to occur, knowledge of the word is needed. On the other hand, to activate word knowledge, readers have to segment the words in order to recognize them.

One approach to this issue has been put forward by Perfetti and Tan (1999), who proposed that Chinese readers prefer to segment two characters into a single word, since most words in Chinese are two characters long. To test this idea, they investigated how Chinese readers segmented overlapping ambiguous strings. In the crucial condition, they embedded the overlapping ambiguous strings into sentences where the string should be segmented as A-BC based on sentence context (e.g., 照顾客 in the experimental sentence frame 经理同意照顾客的想法来设计产品, 'the manager agreed to design products according to the customer's requirements'). The middle character 顾 in the critical region could constitute a word with the first character (照顾 'take care of') and constitute another word with the third character (顾客 'custom'). In the control condition, the first character of the ambiguous string was substituted by a character whose meaning was similar. Thus, it did not constitute an overlapping ambiguity with the other characters in the sentence (e.g., 经理同意按顾客的想法来设计产品, which has the identical meaning to the experimental sentence). For the control condition, in the critical region 按顾客, the first two characters (按顾) are not a word, but the last two characters (顾客 'custom') are a word. Perfetti and Tan found that reading times on the target region were longer for the overlapping ambiguous strings than for the control condition. Hence, they concluded that Chinese readers prefer to initially segment the first two characters in an ambiguous string as a word. If readers did decide it was a word and subsequently found that this was incorrect, they then would need to correct the initial erroneous segmentation. This would take additional time, resulting in increased reading times relative to the control condition. Thus, Perfetti and Tan argued that these results supported the preferred processing strategy.

Evidence against the strictly serial parsing hypothesis, which assumes that characters are grouped into words in a strictly sequential order from left to right, was provided by Inhoff and Wu (2005). They monitored readers' eye movements while they read sentences with a critical four-character sequence (e.g., 专科学生 'college student') consisting of two two-character words (专科 'college' and 学生 'student'). In the ambiguous condition, the central two characters (e.g., 科学 'science') also constituted a two-character word, while in the control condition the central two characters did not constitute a word. Inhoff and Wu found that readers spent more time viewing the critical four-character sequence and its two center characters (科学) in the ambiguous condition than in the unambiguous condition. They concluded that the assignment of characters to words is not a strictly serial left-to-right process. Instead, all of the possible words that can be combined by the characters falling into the perceptual span are activated during the reading of Chinese text. When more words are activated, it takes longer to make the decision regarding how the words should be segmented, resulting in longer reading times in the ambiguous condition than in the unambiguous condition. It should be noted that word frequency might also play an important role in this kind of segmentation. We will discuss this later.

Li et al. (2009) proposed a computational model of Chinese word segmentation based on an interactive-activation perspective (McClelland & Rumelhart, 1981). According to that model, characters in the perceptual span are processed in parallel and the processing of these characters is constrained by how far they are from the point of fixation and by visual attention. The activation of each unit containing a visible character feeds forward to the word recognition level, activating the word unit. When the activation of a word unit reaches a certain level, it feeds activation back to the characters belonging to the activated word. Hence the characters belonging to the activated word will be activated faster than the other characters. In this way, the word-level representations compete with each other until a single word unit wins the competition. At that point, the word is recognized and segmentation occurs. Thus, according to this model, word segmentation and word recognition happen simultaneously.

Some of the assumptions of the word segmentation model were supported by subsequent evidence. Li and Pollatsek (2011) showed that word recognition in Chinese reading is an interactive process such that word knowledge affects lower-level processing during reading. In their study, Chinese readers viewed two Chinese characters. One character was intact. The other, the target, was embedded in a rectangle of visual noise, but it increased in visibility over time (see Figure 16.3). The two characters constituted a word in one condition but not in the other condition. The task was to press a button to indicate whether the character in the noise was at the top or bottom of the rectangle (participants did not have to identify the character). Response times were faster in the word condition than in the nonword condition. As the wordness of the stimulus was logically irrelevant to judging the location of the target character, the data indicate that processing at the word level can feed back to fairly low-level judgments such as where a character is. Thus, these results supported the interactive structure adopted by Li et al. (2009).

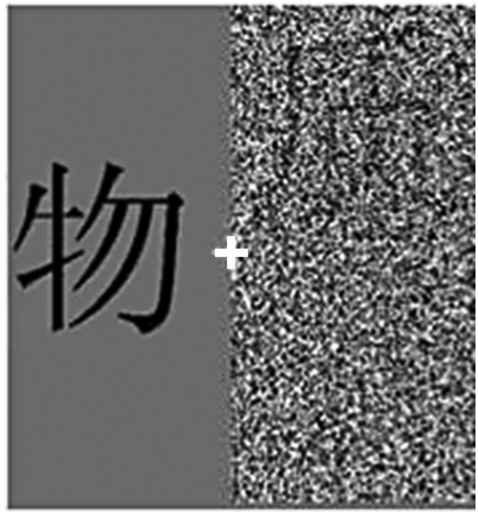


Fig. 16.3 An example of the stimuli: (target is on right). The contrast shown in this figure was the maximum contrast in the experimental display.

Segmentation of Spatially Ambiguous Words

As discussed earlier, there are some complex situations in which word boundaries are ambiguous. The first kind of ambiguity has been called *progressive ambiguity* (Li et al., 2009), where the first one or two characters of a multiple-character word sometimes also constitute another word. For example, in the string 老板娘 'the wife of the boss,' 老板 'boss' is a word, but the three-character string is a word as well. In this example, the word unit 老板娘 receives feed-forward activation from all three characters, while the word 老板 only receives feed-forward activation from two characters. Thus the model proposed by Li et al. (2009) predicts that the word with more characters is always more likely to be initially recognized and selected. For the three characters 老板娘, the model will make an initial commitment to parse them as a three-character word (老板娘) rather than a two-character word (老板), since the word 老板娘 is activated by all three constituent characters.

The second kind of ambiguity occurs (Ma, Li, & Rayner, 2014) when a central character within a string may either be the final character of an earlier word in the string or the first character in a later word in the string, a so-called *overlapping ambiguous string*. For example, in the overlapping ambiguous string 照顾客, both 照顾 'take care of' and 顾客 'customer' are words, but the whole three-character string does not form a word. Ma et al. (2014) explored how Chinese readers segment overlapping ambiguous strings. In Experiment 1, participants were shown three-character ambiguous strings and were simply instructed to name the middle character of the string. The middle character constituted a two-character word with the first character and constituted another two-character word with the third character and was pronounced differently when it paired with each. For each ambiguous string, the frequency of one word was higher than the other. Participants tended to pronounce it as if it belonged to the higher-frequency word, regardless of that word's position (left or right). These results showed that Chinese readers do not always assign the middle character of an overlapping ambiguous string to the left word. Instead, they assigned it to the word that wins the frequency-mediated competition, at least when there is little time pressure.

In Experiment 2 of Ma et al. (2014), two sets of overlapping ambiguous strings with identical first words (AB) but different second words (BC or BD) were embedded in the same sentence frames. The second word in these two strings was either a high-frequency word or a low-frequency word. Eye movements were monitored as these sentences were read. Fixation times on the region AB were longer when the second word was high in frequency than when it was low in frequency. These results showed that the second word in the ambiguous string competes for processing time with the first word when the string is processed. A third experiment investigated how the segmentation of an ambiguous string is constrained by local information such as the frequencies of the two words and global information such as sentential context. Second-pass reading times (the sum of all fixations in a region following the first-pass time, including zero times when a region was not refixated) were shorter and regressions into the ambiguous region were reduced when the segmentation that was based on frequency fit the sentential context. The results support a competition account, such that the characters in the perceptual span activate all of the words they may potentially constitute and any of those candidates can win the competition for identification if its activation is sufficiently high. One way to interpret these results is that word segmentation is at least a two-stage process. During the first stage, word segmentation is determined mainly by local segmentation cues such as relative word frequencies. At a later stage, readers may adjust their initial segmentation commitments if they conflict with sentence context.

Saccade Target Selection in Chinese Reading

In the preceding sections, the fact that the visual and linguistic properties of words in the fovea and parafovea influence eye movement control in reading has been taken as evidence that the word is a basic unit of information associated with ongoing processing in reading. Further evidence that this is the case comes from the observation of a *preferred viewing location* (PVL) on a word in Chinese reading. The PVL (Rayner, 1979) refers to a position on a word that the eyes tend to initially fixate when making a first-pass saccade onto a word (see Schotter & Rayner, this volume). More technically, it is usually reported as a histogram with letter or character position on the *x*-axis and probability of fixating on the *y*-axis, which forms the PVL curve. Rayner (1979) reported that for scripts that are printed from left to right, such as English and French, the PVL is slightly to the left of the center of a word. However, for scripts that are printed from right to left, such as Hebrew (see Deutsch & Rayner, 1999), the PVL on a word is between the middle of the word and the right-most letter (which is the beginning of the word) rather than the left-most letter (as in English). It is generally assumed that readers aim their eyes to the center of a word but for various reasons tend to initially land short of that location on the PVL (see Engbert & Krügel, 2010; McConkie, Kerr, Reddix, & Zola, 1988). These studies suggest that words may be not only the basic units of perceptual encoding but also the functional targets of saccades.

In contrast to the consistency of evidence and views regarding word-based saccadic targeting during reading of alphabetic language scripts, there has been disagreement about whether Chinese readers adopt such a strategy. If Chinese words can be segmented parafoveally, and then saccades targeted on the basis of that parafoveally encoded unit, there should be a tendency for initial fixations on a word to land toward a specific location within words. Yang and McConkie (1999) recorded readers' eye movements while reading Chinese sentences and computed the frequency with which the initial fixation on all two-character words in the sentences was located at different positions in the word. They did not find any differences in terms of the probability of initial fixations on each character; initial fixations landed randomly over the whole word. They thus claimed that there was no preferred viewing location in two-character words. Furthermore, Tsai and McConkie (2003), making an assumption of spatial parity between a two-character Chinese word and a seven-letter English word, found patterns similar to those of Yang and McConkie, such that the PVL curves for both Chinese words and characters were flatter than for English words in reading normally presented text. They concluded that their results provided no evidence for a word-based saccadic targeting strategy in Chinese reading.

In contrast, Yan, Kliegl, Richter, Nuthmann, and Shu (2010) did report that there were more fixations near the beginnings of Chinese words. Their findings were based on corpus analyses of two-, three-, and four-character words in a Chinese text. They further divided the data into single-fixation cases, where readers made only one first-pass fixation on a word, and cases where more than one first-pass fixation was made. The PVL peaked at the word center for words that received single fixations, but peaked at the word beginning when more than one fixation was made on a word (for similar results see Shu, Zhou, Yan, & Kliegl, 2011). Yan et al. argued that Chinese readers target their saccades to the word center if they are able to segment the word in parafoveal vision. If they are not able to do this, they adopt a more cautious targeting strategy: They aim saccades at the word beginning and engage in extra processing on the word after the initial fixation in order to decide where the currently fixated word ends. Thus, Yan et al. proposed that Chinese readers use a word-based strategy to select their saccade target.

Yan et al.'s arguments seem reasonable; however the situation may be more complicated than this. Li et al. (2011) reported experimental data that argue against this model. They embedded either a two-character word or a four-character word in identical sentence frames and compared the fixation distributions on a four-character region of interest. It contained either a two-character word and then another two characters in the two-character word condition, or the whole four-character target word in the four-character word condition. The size of the two regions was identical in the two conditions. Li et al. assumed that if Chinese readers selected the word center as their saccadic target, the mean and mode of the PVL curve in the four-character word condition should be further to the right than in the two-character word condition. However, the PVL curves were almost identical in the two conditions. Additional Bayesian analyses (Rouder, Speckman, Sun, Morey, & Iverson, 2009) showed that the null hypothesis was highly preferred over the alternative hypothesis. This result argues against a saccade target selection strategy based on the length of the upcoming word in Chinese reading.

Li et al. (2011) also considered landing distributions of single fixations and the first of multiple fixations separately, as Yan et al. (2010) did, and their data replicated Yan et al.'s findings for both word lengths. The PVL peaked at the word center in single fixation cases but at the beginning in multiple fixation cases. However Li et al. (2011) argued that these kinds of PVL curves did not necessarily support the word-based targeting strategy. The eyes might fixate toward the word center by chance, and because word processing is more efficient when the eyes fixate at this position (O'Regan, 1981; O'Regan & Lévy-Schoen, 1987), a refixation on the same word may be not necessary. Reinforcing this point, simulations showed that a model in which saccadic targeting was not based on words (e.g., a constant saccade length model) produced very similar patterns of effects. Both the experiment and the simulation of Li et al. indicated that there is no convincing evidence that Chinese readers target any specific position within a word.

A recent study by Zang, Liang, et al. (2013) provided converging evidence against Yan et al.'s (2010) claim that Chinese readers move their eyes to a word's center when they are able to segment words in the parafovea but at a word's beginning when they could not. Yan et al.'s claim would predict that Chinese readers should always move their eyes to a word's center when spaces are inserted between words, since readers should easily perceive word boundary information in the parafovea under these circumstances. Zang, Liang, et al. (2013) examined whether the addition of interword spaces to Chinese text would alter patterns of saccadic targeting during reading. They found that word spacing effects occurred to a similar degree for both children and adults, with differential landing position effects for single and multiple fixation situations. As with Yan et al., for single fixations, readers initially targeted their saccades to a word center. For multiple fixations, initial landing positions were closer to word beginnings (for similar results see Zang, Meng, Liang, Bai, & Yan, 2013). Note again that under interword spaced conditions the beginnings and ends of words are clearly demarcated, and therefore higher order parafoveal word segmentation is no longer necessary. Thus Zang et al.'s results run counter to the prediction of Yan et al.

If Chinese readers do not simply use a word-based strategy or a constant length strategy when selecting a saccade target, what strategy do they adopt? Wei, Li, and Pollatsek (2013) proposed that Chinese readers might estimate how many characters they are processing efficiently on any particular fixation and then send their eyes somewhere to the right of those characters. They termed this possibility a processing-based strategy. Using this strategy, the processing difficulty of the fixated words should affect the length of the saccade from that fixation: The easier the current processing, the longer should be the outgoing saccade. Wei et al. manipulated word length and word frequency separately in two experiments. In the first experiment, the target region was a four-character string that was either a word (one-word condition) or a phrase comprised of two two-character words (two-word condition), where the former has been shown to be easier to process than the latter. In the second experiment, the target region was either a high-frequency two-character word or a low-frequency two-character word. Each pair of the target words was fit into each sentence frame. Wei et al. found that the outgoing saccade length from the last fixation on the target region was longer in the one-word condition than the two-word condition and was longer in the high-frequency two-character word condition than in the low-frequency two-character word condition. These results indicate that the properties of words that are being fixated affect the length of the outgoing saccade from them. Similar findings were reported in Li, Bicknell, et al. (2014). They analyzed a corpus of eye-movement data during Chinese reading and found that outgoing saccade length was affected by the predictability, frequency, and length of the currently fixated word. This finding is consistent with the processing-based view of eye movement control in Chinese reading, and it confirmed the previous finding that outgoing saccade length was affected by the properties of the fixated word. Moreover, Li, Bicknell, et al. found that character fixation probability did not differ as a function of within-word position, confirming the findings by Li et al. (2011) of no PVL in Chinese reading.

In summary, saccade targeting may operate in a different and more complicated manner in Chinese than in most alphabetic languages. As suggested by Zang, Liang, et al., "information such as a word's predictability, parafoveal familiarity, within-word character positional probability, between-word character transitional probability, as well as other sources of information could all contribute to saccadic targeting decisions in Chinese" (2013, p. 731). Much more work is needed to clarify this issue in the future.

Concluding Remarks

One important difference between Chinese and many other writing systems is that there are no spaces to mark word boundaries between words. Because the characters in Chinese reading are salient units, character processing might play an important role in Chinese reading. This does not mean that words are not important in Chinese reading, however. As we have described, numerous studies have shown that words have psychological reality and play an important role in Chinese reading. Considerable progress has been made recently to develop our understanding of the factors affecting eye movements during reading in Chinese. A substantial proportion of this work has focused on issues related to the role of the word in Chinese as well as how word segmentation occurs during normal Chinese reading. Recent progress has improved our understanding of the mechanisms of Chinese reading, both generally, in relation to how processing occurs compared with other languages, and more specifically, in relation to the unique properties of the Chinese writing system itself. It is likely that the findings reviewed in this chapter will also generalize to other writing systems that do not have explicit word boundaries.

Acknowledgments

The work was supported by a Grant from the Natural Science Foundation of China (31100729), the Recruitment Program of Global Experts (1000 Talents Award from Tianjin), and Leverhulme Trust Grants (F/00 180/AN and RPG-2013-205).

References

Bai, X., Liang, F., Blythe, H. I., Zang, C., Yan, G., & Liversedge, S. P. (2013). Interword spacing effects on the acquisition of new vocabulary for readers of Chinese as a second language. *Journal of Research in Reading, 36*(S1), S4–S17.

Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1277–1287.

Blythe, H. I., Liang, F., Zang, C., Wang, J., Yan, G., Bai, X., & Liversedge, S. P. (2012). Inserting spaces into Chinese text helps readers to learn new words: An eye movement study. *Journal of Memory and Language, 67,* 241–254.

Boorstin, D. J. (1983). *The discoverers*. New York, NY: Random House.

Cheng, C. (1981). Perception of Chinese characters. *Act Psychological Taiwanica, 23,* 137–153.

Cui, L., Drieghe, D., Yan, G., Bai, X., Chi, H., & Liversedge, S. P. (2013). Parafoveal processing across different lexical constituents in Chinese reading. *Quarterly Journal of Experimental Psychology, 66,* 403–416.

Cui, L., Yan, G., Bai, X., Hyönä, J., Wang, S., & Liversedge, S. P. (2013). Processing of compound-word characters in reading Chinese: An eye-movement–contingent display change study. *Quarterly Journal of Experimental Psychology, 66,* 527–547.

Deutsch, A., & Rayner, K. (1999). Initial fixation location effects in reading Hebrew words. *Language and Cognitive Processes, 14,* 39–421.

Engbert, R., & Krügel, A. (2010). Readers use Bayesian estimation for eye-movement control. *Psychological Science, 21,* 366–371.

Hoosain, R. (1992). Psychological reality of the word in Chinese. In H. C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 111–130). Amsterdam, the Netherlands: North-Holland.

Inhoff, A., & Wu, C. (2005). Eye movements and the identification of spatially ambiguous words during Chinese sentence reading. *Memory & Cognition, 33,* 1345–1356.

Lexicon of Common Words in Contemporary Chinese Research Team. (2008). *Lexicon of common words in contemporary Chinese*. Beijing, China: Commercial Press.

Li, X., Bicknell, K., Liu, P., Wei, W., & Rayner, K. (2014). Reading is fundamentally similar across disparate writing systems: A systematic characterization of how words and characters influence eye movements in Chinese reading. *Journal of Experimental Psychology: General, 143,* 895–913.

Li, X., Gu, J., Liu, P., & Rayner, K. (2013). The advantage of word-based processing in Chinese reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 879–889.

Li, X., Liu, P., & Rayner, K. (2011). Eye movement guidance in Chinese reading: Is there a preferred viewing location? *Vision Research*, *51,* 1146–1156.

Li, X., & Pollatsek, A. (2011). Word knowledge influences character perception. *Psychonomic Bulletin & Review, 18,* 833–839.

Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology, 58,* 525–552.

Li, X., & Shen, W. (2013). Joint effect of insertion of spaces and word length in saccade target selection in Chinese reading. *Journal of Research in Reading, 36*(S1), S64–S77.

Li, X., Zhao, W., & Pollatsek, A. (2012). Dividing lines at the word boundary position helps reading in Chinese. *Psychonomic Bulletin & Review, 19,* 929–934.

Liu, P., Li, W., Lin, N., & Li, X. (2013). Do Chinese readers follow the national standard rules for word segmentation during reading? *PLoS ONE, 8,* e55440.

Liu, P., & Li, X. (2014). Inserting spaces before and after words affects word processing differently: Evidence from eye movements. *British Journal of Psychology, 105,* 57–68.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Science, 4,* 6–14.

Liversedge, S. P., Zang, C., Zhang, M., Bai, X., Yan, G., & Drieghe, D. (2014). The effect of visual complexity and word frequency on eye movements during Chinese reading. *Visual Cognition, 22,* 441–457.

Ma, G., Li, X., & Rayner, K. (2014). Word segmentation of overlapping ambiguous strings during Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 1046–1059.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial fixations in words. *Vision Research, 28,* 1107–1118.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

National Standard GB/T 13715–92. (1992). *Contemporary Chinese language word segmentation specification for information processing*. Beijing, China: China Standards Press.

O'Regan, J. K. (1981). The "convenient viewing position" hypothesis. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 289–298). Hillsdale, NJ: Erlbaum.

O'Regan, J. K., & Lévy-Schoen, A. (1987). Eye movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and performance: Vol. 12. The psychology of reading* (pp. 363–383). Hillsdale, NJ: Erlbaum.

Packard, J. L. (1998). Introduction. In J. L. Packard (Ed.), *New approaches to Chinese word formation: Morphology, phonology and the lexicon in modern and ancient Chinese* (pp. 1–34). Berlin, Germany: Mouton de Gruyter.

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The lexical constituency model: Some implications of research on Chinese for general theories of reading. *Psychological Review, 112,* 43–59.

Perfetti, C. A., & Tan, L. H. (1999). The constituency model of Chinese word identification. In W. J & I. A. W (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 115–134). Hillsdale, NJ: Erlbaum.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52,* 1–52.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8,* 21–30.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422.

Rayner, K. (2009). The thirty-fifth Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62,* 1457–1506.

Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38,* 1129–1144.

Rayner, K., Li, X., Juhasz, B. J., & Yan, G. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review, 12,* 1089–1093.

Rayner, K., Li, X., & Pollatsek, A. (2007). Extending the E-Z Reader model of eye movement control to Chinese readers. *Cognitive Science, 31,* 1021–1033.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81,* 275–280.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Towards a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16,* 225–237.

Shen, D., Liversedge, S. P., Tian, J., Zang, C., Cui, L., Bai, X., . . . Rayner, K. (2012). Eye movements of second language learners when reading spaced and unspaced Chinese text. *Journal of Experimental Psychology: Applied, 18,* 192–202.

Shu, H., Zhou, W., Yan, M., & Kliegl, R. (2011). Font size modulates saccade-target selection in Chinese reading. *Attention, Perception & Psychophysics, 73,* 482–490.

Tsai, J. L., & McConkie, G. W. (2003). Where do Chinese readers send their eyes? In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 159–176). Oxford, England: Elsevier.

Wang, H., Pomplun, M., Chen, M., Ko, H., & Rayner, K. (2010). Estimating the effect of word predictability on eye movements in Chinese reading using latent semantic analysis and transitional probability. *Quarterly Journal of Experimental Psychology, 63,* 1374–1386.

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology, 1,* 59–85.

Wei, W., Li, X., & Pollatsek, A. (2013). Word properties of a fixated region affect outgoing saccade length in Chinese reading. *Vision Research, 80,* 1–6.

Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology, 97,* 259–268.

Yan, M., Kliegl, R., Richter, E. M., Nuthmann, A., & Shu, H. (2010). Flexible saccade target selection in Chinese reading. *Quarterly Journal of Experimental Psychology, 63,* 705–725.

Yang, H., & McConkie, G. W. (1999). Reading Chinese: Some basic eye-movement characteristics. In J. Wang, A. W. Inhoff, & H.-C. Chen (Eds.), *Reading Chinese script* (pp. 207–222). Hillsdale, NJ: Erlbaum.

Yen, M. H., Tsai, J. L., Tzeng, O. J. L., & Hung, D. L. (2008). Eye movements and parafoveal word processing in reading Chinese. *Memory & Cognition, 36,* 1033–1045.

Zang, C., Liang, F., Bai, X., Yan, G., & Liversedge, S. P. (2013). Interword spacing and landing position effects during Chinese reading in children and adults. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 720–734.

Zang, C., Liversedge, S. P., Bai, X., & Yan, G. (2011). Eye movements during Chinese reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *Oxford handbook of eye movements* (pp. 961–978). New York, NY: Oxford University Press.

Zang, C., Meng, H., Liang, F., Bai, X., & Yan, G. (2013). Is there a vertical component to saccade targeting in Chinese reading? *Journal of Research in Reading, 36*(S1), S78–S93.

CHAPTER

17 How Is Information Integrated Across Fixations in Reading?

Michael G. Cutter, Denis Drieghe, *and* Simon P. Liversedge

Abstract

This chapter explores the integration of information acquired over multiple eye fixations during reading by reviewing studies using the boundary paradigm. This integration process is examined for information extracted from the end of the fixated word, the word to the right of the fixated word, and the word two words to the right of the fixated word. The review shows that the amount of information integrated across fixations varies for these three different types of previewed visual information. A large variety of information extracted from a parafoveal word is integrated across fixations, including orthography, phonology, and meaning. The chapter considers how such integration processes operate across several languages to allow understanding of how the way in which linguistic characteristics are orthographically coded in a particular language constrains parafoveal processing. It is concluded that readers preferentially integrate the information that is most useful for the initiation of lexical access.

Key Words: preview benefit, parafoveal processing, eye movements, reading, orthography, phonology, semantics, morphology, cross-linguistic differences

During reading, eye movements are made in order to fixate a word in high acuity foveal vision so that it is processed as efficiently as possible. The *fovea* is a small region of 2° of visual angle which, depending on factors such as font size and viewing distance, will typically extend over approximately six characters during reading. Beyond the fovea is the *parafovea*, in which visual acuity is considerably reduced. The parafovea extends beyond the fovea—4° to either side (Balota & Rayner, 1991). *Saccades* are rapid ballistic eye movements that move the eye from one point to another, and *fixations* are the periods of stillness between them. It is during the fixations that visual information is extracted from the page, and a large body of research has shown that the amount of time a word is fixated is tightly linked to the processing of that word (see Schotter & Rayner, this volume). While a large proportion of the processing of a word takes place in foveal vision, it is not the case that encoding only begins upon direct fixation of the word. Rather, a word is often partially processed on a fixation on a prior word. The parafoveal information extracted from this fixation is then carried over and integrated with foveal information that is available when the word is fixated. Furthermore, a single word is sometimes fixated more than once, in which case the information extracted during these multiple direct fixations must also be integrated. The process of integrating information extracted across multiple fixations is the focus of the current chapter.

The fact that a word is often processed over multiple fixations is apparent from studies using the *moving window paradigm* (McConkie & Rayner, 1975). In this paradigm, a window of normal text is set by the experimenter around the point of fixation. Within this window the characteristics of the text being read are preserved, whereas outside the window the text is masked. As a saccade is made a display change occurs, so that a new window of unmasked text is set around the new point of fixation (see Figure 17.1). This display change typically completes prior to the end of the saccade, and, as such, participants are usually unaware of the manipulation due to visual information not being encoded during a saccade (e.g., Martin, 1974). The smallest window size for which reading occurs at a rate similar to normal reading is referred to as the *perceptual span*. The moving window paradigm gives an estimate of how much information readers extract during a single fixation, albeit with no indication about the form of the extracted information. For English, this extends 3 to 4 character spaces to the left and 14 to 15 characters to the right of fixation (McConkie & Rayner, 1975, 1976a). While readers are able to extract information 14 to 15 characters into the parafovea, the average saccade tends to move the eyes 7 to 9 characters forward (Rayner, 1998). Therefore, readers usually have overlapping perceptual spans across two fixations, meaning that the same word is often available for processing across multiple fixations. It is clear from this that readers are often integrating information across fixations, since their reading speed decreases when the window is smaller than the size of the perceptual span and a word is thus not available for processing across multiple fixations.

While the moving window paradigm can be used to demonstrate that information is processed across multiple fixations, it does not allow us to infer the nature of this processing, or the type of representation that is integrated across fixations. Several theoretical possibilities exist as to why restricting parafoveal information in moving window studies slows reading. For example, one early theory proposed that purely visual information obtained from the parafovea is stored between fixations and that new visual information obtained on direct fixation is added to this visual representation (McConkie & Rayner, 1976b). Pollatsek, Lesch, Morris, and Rayner (1992) proposed an approach based on the idea that phonological coding serves an important role in silent reading by helping to create a representation of identified words in short-term memory. According to this approach a phonological code is obtained for a word seen in the parafovea, which is used to preserve the memory of that word across fixations. A third possibility is that a parafoveal stimulus activates a set of lexical entries on the basis of several abstract word characteristics (e.g., orthography, phonology, morphology, and semantics) and that this activation is carried across multiple fixations. This lexical activation may then lead to the faster identification of a word once it is directly fixated, explaining the slowdown in reading when parafoveal information is denied in the moving window paradigm.

In order to discriminate between the possibilities we have outlined it is necessary to manipulate specific characteristics of a single word in the parafovea and examine how this affects fixation times on that word. This issue has been investigated using a second eye contingent change technique, the *boundary paradigm* (Rayner, 1975). In the boundary paradigm an invisible boundary is (usually) set at the end of a pretarget word (see Figure 17.2). Prior to the eyes crossing the boundary, a preview string is presented instead of the target word. This preview can be the target word itself (an identity preview), a different word, or a nonword. The preview quickly changes to the target word as a saccade is made that crosses the boundary. As will be seen, this technique has been widely used and has demonstrated that when readers are given an identity preview of a target word, they take less time to process and identify it than when they are given an incorrect preview. This advantage is referred to as the *preview benefit*. The fact that readers gain a preview benefit strongly suggests that they have extracted and processed information about the preview string before fixating it and then have integrated this information with information obtained on the next fixation, which is usually made on the target word itself. By varying the relationship between the preview and the target it is possible to

```
          *
Lbo dhr quickly jumped over lfa pameo ez bg pem tvoq
                   *
Lbo dhr pvlsziv jumped over the feneo ez bg pem tvog
                           *
Lbo dhr pvlsziv fbtyed over the fence az bg pem tvog
```

Fig. 17.1 An illustration of the moving window paradigm. The point of fixation is represented by the asterisk. In the current illustration 4 characters to the left and 14 characters to the right of the point of fixation are available.

```
              *
The man quickly| lvnqod over the fence as he ran away
               |
               |   *
The man quickly| jumped over the fence as he ran away
```

Fig. 17.2 An illustration of the boundary paradigm. The point of fixation is represented by the asterisk. The invisible boundary is represented by the black line. Prior to the point of fixation crossing this boundary there is a nonword preview of the upcoming word. As the point of fixation crosses the boundary, the target word replaces the preview.

discover the types of information that are extracted and integrated across fixations and to discriminate among different explanations regarding the nature of trans-saccadic integration. For example, if preview benefit effects were purely driven by visual overlap between a preview and target word, it would suggest that the integration process depends on the combination of visual information into a single percept. Were it simply the case that a phonological code is used to aid short-term memory of a parafoveal stimulus, then preview benefits should only be observed for previews that share phonological information with the target word. If, however, it were to be found that preview benefits are determined by a wide range of abstract information about a parafoveal word, including orthography, phonology, morphology, and semantics, it would suggest that the integration process works on the basis of lexical entries (representing multiple linguistic characteristics) that have been partially activated by the parafoveal stimulus. As will be seen throughout this chapter, a substantial amount of research supports this third position, with preview effects being documented for types of information that go beyond low-level visual similarity between the preview and target or a phonological representation in short-term memory.

In this chapter we will consider the integration of information from several different types of parafoveal stimuli. We will first examine how information is integrated across multiple fixations on the same word. We will then focus on the different types of information that are extracted and integrated from the word to the right of fixation, demonstrating the range of types of information that are integrated across fixations. We will also explore whether information is integrated from two words to the right of fixation. Finally, we will briefly discuss factors that modulate the degree to which information is integrated across fixations. We will restrict our discussion to research conducted on adult readers whose reading has developed typically. While the majority of our discussion will focus on studies conducted in English and other similarly spaced alphabetic languages, we will also briefly discuss studies conducted in languages in which the phonological, morphological, or semantic characteristics associated with a word are orthographically coded in a manner different from English. These studies will be considered with regard to how these cross-linguistic differences affect integration across fixations.

Before proceeding further, it is necessary to briefly discuss some eye-movement measures that have been widely used to assess preview benefit. There are several fixation time measures that can be calculated that vary regarding the extent to which they take one or multiple fixations on a word into account. Early measures include *first-fixation duration* (the mean duration of the initial fixation on a word) and *gaze duration* (the mean amount of time between first fixating a word and making a saccade away from it). Later measures take further refixations on a word into account. For example, *total fixation time* includes all fixations made on a word (including later fixations made when a word is reread).

The Integration of Information During Refixations

Not all saccades move the eyes onto a new word. Rather, a reader will refixate approximately 15% of words (Rayner, 1998). Refixations are more likely for longer words, and the most common pattern is for the initial fixation to be made toward the beginning of a word and the second toward the end (Rayner, Sereno, & Raney, 1996). The fact that refixations often follow this pattern suggests they are made partly due to a need to process characters from the end of a word more centrally within foveal vision. As such, it is worth considering the extent to which information from the end of a long word is integrated with that from the beginning of the same word across fixations. Drieghe, Pollatsek, Juhasz, and Rayner (2010) examined this for monomorphemic words (e.g., *fountain*) and unspaced compounds, which are made of two smaller lexemes (e.g., *bathroom*). A variation of the boundary paradigm that was first implemented by Hyönä et al. (2004) was used, in which the boundary was placed in the middle of the word instead of between words. In Drieghe et al.'s study, the boundary was between the first and second constituent in the unspaced compounds, or between the corresponding letters of a monomorphemic word that was matched on length, overall frequency, and initial bigram and trigram frequency. In the incorrect preview conditions, the final letters of the target word after the boundary were replaced (e.g., *fountaom, bathroan*). There was a large effect of this preview manipulation, with readers having gaze durations of 151 ms and 146 ms longer in the incorrect preview conditions than in the identity conditions as assessed by the gaze duration on the postboundary portion of the target word for the monomorphemic words and the unspaced compounds, respectively. This demonstrates that when a word is fixated twice, the information from the end of the word has already been processed to a considerable degree during the initial fixation on the word and this information is integrated with information gained on refixation.

Although preview benefits from the second portion of the target words were similar for the monomorphemic and unspaced compound words, there were differences in the degree to which the preview affected fixations prior to crossing the boundary. That is, fixations on the first half of a monomorphemic word were lengthened by the incorrect information in the second half of the preview; however, this was not the case for the unspaced compound words. The characteristics of a parafoveal letter string affecting fixation durations on the prior foveal text is referred to as a parafoveal-on-foveal effect (for a review, see Drieghe, 2011). These parafoveal-on-foveal effects are generally viewed as evidence that information in the parafovea is being processed in parallel with information in the fovea. Otherwise, this parafoveal information would be processed too late to affect the fixation duration on the foveal word. The fact that these preview effects were observed for monomorphemic words but not for unspaced compounds indicates that, while the former are processed as single units, the latter are processed, at least to some degree, as two independent subunits. This in turn affects the rate at which parafoveal information is integrated.

Häikiö, Bertram, and Hyönä (2010) also investigated the integration of information from the second lexeme of an unspaced compound word (in Finnish, all compound words are unspaced), using the within-word version of the boundary paradigm. They varied whether readers received a correct preview of the second constituent or a preview in which all but the initial two letters were replaced. There was a main effect of the preview manipulation on fixation time on the second constituent and an interaction between the frequency of the compounds and the preview manipulation during fixations prior to crossing the boundary (i.e., for the high-frequency compounds, an incorrect preview resulted in longer fixations on the first constituent before the boundary was crossed, whereas this was not the case for the low-frequency compounds). Häikiö et al. proposed that the high-frequency compounds were identified via a single lexical entry, whereas the low-frequency compounds were identified as two separate lexemes. Thus, the incorrect preview was only processed in parallel with the first constituent in the high-frequency compounds.

Hyönä, Bertram, and Pollatsek (2004) demonstrated that the frequency of an unspaced compound's first constituent also affects the integration of information from the second constituent across fixations. They manipulated the frequency of the first constituent of the compound word and all but the two initial letters of the second constituent. These initial letters were either incorrect or correct until the saccade crossed from the first constituent to the second constituent. The preview benefit observed during fixations on the second constituent was larger when the first constituent was a low-frequency word than when it was a high-frequency word. However, there was no evidence of an effect of the letters of the second constituent being incorrect during fixations on the first constituent—even for the compounds with low-frequency first constituents. This suggests that the difference in the preview effect on the second constituent was not caused by the second constituent being processed as part of the whole compound. Rather, Hyönä et al. proposed that these effects were driven by the fact that the low-frequency first constituents potentially combined with fewer second constituents to form a compound word than the high-frequency first constituents. As such, that set of potential second constituents was more constrained given a low-frequency first constituent, and thus those actual second constituents were processed more efficiently in the parafovea, as separate lexemes. (See Cui et al., 2013 for a similar argument in processing Chinese compound words.) Taken together, the findings of Hyönä et al. (2004), Drieghe et al. (2010), and Häikiö et al. (2010) suggest that a number of factors influence the amount of information that is integrated from the second constituent of an unspaced compound across fixations and the time course of this processing. How highly constrained the second constituent is influences the amount of information that is integrated across fixations. Furthermore, the frequency of the whole compound influences the time when this information first has an observable effect on processing. Further research is required in order to extend current understanding of the factors that determine how the second constituents of unspaced compound words are processed in the parafovea.

White, Bertram, and Hyönä (2008) undertook an experiment in Finnish that investigated whether semantic information from an unspaced compound's second constituent is integrated across fixations. In this study, participants were given a preview of the second constituent of an unspaced compound (*vaniljakastike* 'vanilla sauce') while fixated on the first constituent. There were four possible previews: an identity preview (*vaniljakastike* 'vanilla sauce'), a semantically related preview (*vaniljasinappi* 'vanilla mustard'), a semantically unrelated preview (*vaniljarovasti* 'vanilla priest'), and a pronounceable nonword preview (*vaniljaseoklii*). The identity preview led to shorter fixations on both the second constituent and across the whole compound than any of the other preview types. While the semantically related preview provided no benefit relative to the unrelated and nonword previews in either first-fixation duration or gaze duration, on either the second constituent or whole compound, there was a benefit in regression path durations within the compound (this includes all fixations within the compound from first fixating the second constituent, until a rightward saccade was made out of the compound). This fairly late effect of the semantic preview suggests that semantic information was extracted from the second constituent of the compound but was not integrated immediately upon fixating the target constituent. Rather, integration only occurred during the later phases of compound word processing.

In summary, when a word is fixated multiple times, information is indeed integrated across these fixations. This is true for both monomorphemic and compound words, with substantial preview effects being found within both types of word (Drieghe et al., 2010). The extent to which this information is integrated across fixations within a compound word depends on both the frequency of the whole compound word (Häikiö et al., 2010) and the extent to which the first constituent constrains the second (Hyönä et al., 2004).

The Integration of Information From Word *N* + 1

In the following sections we explore the integration of various types of information extracted from the upcoming word, word *n* + 1. We begin our discussion by examining the processing of relatively low-level information (e.g., purely visual information, letter identities) and progress toward increasingly abstract information (e.g., morphological, semantic).

Orthographic Codes

A basic question that we can ask in relation to the integration of information across fixations is whether the information that is integrated is based entirely on the visual form of the words or on abstract linguistic information that is derived from the orthography of the words. Studies have addressed this by examining the effects of changing the visual characteristics of text across saccades while holding letter information constant. McConkie and Zola (1979) had participants read text in which words were written in alternating case, with the case of each letter changing during saccades (e.g., *ReD* -> *rEd*). This manipulation changed the visual information between fixations while keeping the letter identities constant. There was no slowdown in reading when the case changed across fixations relative to a condition with no display changes, indicating that the integration of information is not restricted to visual forms. Similarly, Rayner, McConkie, and Zola (1980) showed that participants were no slower at naming a target word when case changed across fixations as opposed to staying the same.

While changing the case of the letters across fixations does not have a significant effect on reading, other work has demonstrated an effect of the visual similarity between a preview and target word. In a meta-analysis of studies in English using the boundary paradigm, Hyönä et al. (2004) showed that using visually similar replacement letters (e.g., *b* and *d*) results in smaller preview effects relative to an identity condition (15 ms in gaze duration on average) than using visually dissimilar letters (e.g., *p* and *s*; 41-ms effect on average). These findings suggest that the orthographic information that is integrated across fixations is in the form of abstract letter identities, which have been activated by low-level visual features. Since visually similar letters will coactivate each other due to shared features (e.g., the vertical ascender in *d* will activate *d, b,* and *h*), previews with similar letters will activate the target to a greater extent than previews without similar letters. This also explains why case changes across fixations do not affect reading. While low-level features changed across fixations in these studies, letter identities did not. Thus a (case-independent) letter representation would have been activated by the features of its lower-case form on one fixation and the features of its upper-case form on the next (and vice versa). As long as letter activation was carried across fixations, it would not have mattered whether this activation was due to the same low-level features on all fixations or different features on each fixation.

There is evidence that letters from different positions within a word are not equally important when information is integrated across fixations. Inhoff (1989a) gave previews of six-letter words in which the whole word (e.g., *survey*), the initial trigram (e.g., *surxxx*), the final trigram (e.g., *xxxvey*), or nothing (e.g., *xxxxxx*) was available. Furthermore, the reading direction was varied (e.g., *a recent survey* vs. *survey recent a*), with participants reading from right to left in the latter condition in order to ensure that any letter-position effects were not due to visual acuity. The initial trigram led to slightly, though not significantly, greater facilitation than the final trigram (16 ms vs. 12 ms in first-fixation duration) regardless of reading direction. Furthermore, when visually dissimilar replacement letters were used instead of *x*'s, Inhoff found that the final trigram alone no longer provided a significant preview benefit, whereas parafoveal availability of the initial trigram still led to a significant benefit of 6 ms. Briihl and Inhoff (1995) further investigated this issue by varying the number of correctly previewed letters and their position in a word, and found that previewing external and initial letters was significantly more facilitative than previewing internal letters. One probable reason for the greater benefit of external letters is reduced crowding relative to internal letters, due to being located next to a space. Briihl and Inhoff also found that previews of both final and initial letters together did not facilitate processing significantly more than previews including only initial letters, suggesting that final letters do not play a particularly important role in trans-saccadic integration. However, in both studies, whole-word previews were more facilitative than would have been expected had the effect of each extra letter been additive. This suggests that the letters were parafoveally encoded as part of a whole word and mutually reinforced each other's activation.

While word-initial information in English is given preferential treatment in trans-saccadic integration, this does not generalize to Chinese. Rather than consisting of a string of letters representing a phonological code, Chinese characters are made up of a number of strokes that form subunits known as *radicals*. Many characters consist of more than one radical, and the majority of these characters contain a radical that carries phonological information and another that carries semantic information. While these radicals contain this abstract information, the relationship between a character and its radicals is not always strong, with, for example, the pronunciation of only 30% of phonetic radicals corresponding to that of the full character (Zhou & Marslen-Wilson, 1999). As such, two characters with the same phonetic radical may be pronounced differently. Clearly, linguistic information is orthographically coded in Chinese in a vastly different way than in English and so may be integrated differently across fixations. Liu, Inhoff, Ye, and Wu (2002) conducted a boundary study in which the preview and target shared orthographic information via (1) the semantic radical, (2) the phonetic radical, or (3) stroke information while sharing neither radical, or (4) shared no orthographic information. Liu et al. found that participants gained a significant preview benefit given an overlapping phonetic radical but not from the other conditions. This effect was observed regardless of whether the target and preview character were phonologically similar. The phonetic radical typically appears on the right side of a Chinese character. Thus orthographic preview benefit is driven by character-final information in Chinese and word-initial information in English. One possible reason for this is that parafoveal orthographic information is used to initiate lexical access and that the optimal information for this differs across languages. In English the initial letters of a word may be more useful, in part due to their importance in generating a phonological code. However, Liu et al. argued that in Chinese the phonetic radical is more useful for two reasons. First, they claimed that it is the smallest orthographic unit that is represented in the character lexicon, as it can form a character in isolation. They also claimed that the phonetic radical provides more discriminative information with which to select character candidates from the lexicon. Thus, while different orthographic information is integrated to differing extents in each language, the time course of processing appears to be driven by the underlying principle of which information is most optimally used to initiate lexical access.

As well as investigating how letter identity information is integrated across saccades, researchers have also examined letter-position encoding in the parafovea (see Frost, this volume, for a general discussion of letter-position coding). Johnson, Perea, and Rayner (2007) provided readers with parafoveal previews in which two letters had been transposed (e.g., *loewr* as a preview of *lower*) or substituted (e.g., *loanr*), finding that the transposed letter previews were more facilitative than the substituted letter previews. Johnson (2007) found that this effect endured even when the transposition was made between nonadjacent letters (e.g., *flower* to *flewor*). Johnson and Dunne (2012) presented participants with previews that varied in whether letters were transposed or substituted and whether they created a nonword or a word that was orthographically similar to the target (e.g., *besat* and *beats* as transposed letter previews and *berut* and *beach* as substituted letter previews for the target word *beast*). Preview effects were driven exclusively by the extent of orthographic overlap between the previews and the targets, such that the two transposed letter previews produced shorter fixations on the target word than the two substituted letter previews. There was no significant difference between whether the preview was a word or nonword. This study provided further evidence for the transposed letter effect during reading. Furthermore, these findings suggested that processing in the parafovea does not typically proceed to the later stages of lexical processing during which lexical candidates compete by inhibiting the activation of orthographically similar words. If this had occurred, the word previews should have led to smaller preview benefits than the nonword previews. Together, these studies show that the identity of a letter maintains activation across fixations independent of position. However, this is not to say that letter position per se is not important. Clearly it is, since the identity preview always provided reliably more benefit than the transposed letter previews in all of these studies.

The studies we have discussed in this section demonstrate that information about both letter identity and letter position is integrated across fixations. The importance of letter identity information is weighted in relation to a letter's position within a word, and this factor has a differential influence across orthographies (see Frost, this volume, for an in-depth discussion of how orthographic encoding may differ across the orthographies of different languages).

Phonological Codes

One reason for the greater importance of word-initial letters in preview benefit may be their role in generating a phonological code to initiate lexical access. Accordingly, it might be expected that an element of such a code might also be taken from the parafovea and integrated with the phonological codes extracted when the word is fixated. In the following section, we consider a series of studies that examined whether phonological codes are integrated across fixations and the nature of these representations (see Pollatsek, this volume, for a more in-depth discussion of phonological coding during reading).

One way in which phonological processing has been investigated is through the use of *homophones* in preview studies. Homophones are two words that are spelled differently but pronounced the same. Pollatsek et al. (1992) used the boundary paradigm and presented participants with homophone previews (e.g., *beach* as a preview for *beech*) or orthographic control previews (e.g., *bench*). Participants gained a greater preview benefit from the homophones than the controls. These results suggest that the overlapping phonological code was integrated across fixations. Chace, Rayner, and Well (2005) replicated this effect, but only in skilled university aged readers, with less skilled university aged readers showing no preview effects. Bélanger, Mayberry, and Rayner (2013) extended the finding by manipulating the relative frequency of the homophone preview and target (i.e., the higher frequency word of the homophone pairs was the preview in half the trials and the target in the other half). Participants gained a phonological preview benefit from the high-frequency preview but not from the low-frequency preview.

While the preceding studies demonstrate that readers integrate phonological codes across fixations, it is unclear whether this is driven by addressed or assembled phonology. That is, the reader may either gain access to the phonological code via the identification of a complete orthographic representation (a look-up process) or through the use of grapheme-phoneme correspondence rules to assemble a phonological code. Miellet and Sparrow (2004) investigated this issue in French by giving participants nonword homophone previews (e.g., *maizon* as a preview for *maison* 'house') or orthographic controls (e.g., *mailon*). Despite the homophone preview being a nonword, it facilitated reading. The fact that this effect occurred when the preview strings were nonwords suggests that the benefit comes from assembled phonology, since there is no stored lexical representation via which a phonological code might be accessed (for evidence of English readers gaining a phonological preview benefit from nonwords see Ashby, Treiman, Kessler, & Rayner, 2006). However, the fact that Bélanger et al. (2013) observed an influence of word frequency on phonological preview effects suggests that readers do sometimes retrieve, as opposed to assemble, a phonological code, with it being possible to extract this information more rapidly from a high-frequency word than from a low-frequency word. Depending on the circumstances, then, readers may make use of either addressed or assembled phonology. Further research is needed to determine the factors affecting which route a reader takes to obtain the phonological code of a parafoveal word.

The studies discussed in this section up to this point all manipulated phonological overlap at a whole-word level. Other studies have examined the integration of more fine-grained phonological information within a word. Ashby and Rayner (2004) examined the role of syllabic structure by giving participants previews of words with either a consonant-vowel-consonant (e.g., *concave*) or consonant-vowel (e.g., *device*) initial syllable. A space manipulation was also used so that previews either preserved (e.g., *de_pxw* for device) or violated (e.g., *dev_px*) this structure. Participants remained fixated on the target word for less time when the preview maintained the structure. This was true even for words with a consonant-vowel initial syllable, despite the incongruent preview providing more orthographic information. Thus phonological information at the level of syllables is integrated across fixations, and having these syllables clearly visually delimited in the parafovea may facilitate subsequent processing to a greater extent than a larger number of letters that do not maintain syllabic structure. This suggests that word-initial letters may be more facilitative partly because of their role in generating a phonological code. Fitzsimmons and Drieghe (2011) demonstrated that the extraction of a word's syllabic structure in the parafovea must occur rapidly. In this study either a monosyllabic or a disyllabic word matched on word length, frequency, predictability, number of orthographic neighbors, and mean bigram frequency was embedded into a sentence. The monosyllabic word was skipped more regularly than the disyllabic word. On the assumption that the parafoveal word's syllabic structure influenced where the next saccade was targeted, this indicates that this information was extracted early enough during parafoveal processing for it to influence saccadic targeting. Ashby et al. (2006) also investigated whether vowel information is integrated across fixations by contrasting previews that shared a vowel phoneme with the target word to previews that did not (e.g., *cherg* and *chorg*, respectively, as previews for *chirp*). Previews that shared a vowel phoneme with the target word were more facilitative, even when the vowel's pronunciation needed to be modified by subsequent consonants to be concordant (e.g., *raff* as opposed to *rall* as a preview for *rack*). This study demonstrated that individual vowel sounds are also integrated across fixations.

The nature of alphabetic languages means that there is a relatively direct link between orthography and phonology in that letters link reasonably reliably to certain phonemes. This is not true for a character-based language, such as Chinese. In Chinese, similar-looking characters often have different pronunciations, and homophonic characters may be entirely visually distinct (Hoosain, 1991). Furthermore, as mentioned earlier, Chinese characters contain a phonetic radical, which in some cases represents the character's phonology but in other cases contains phonological information that does not match the character's pronunciation. Tsai, Lee, Tzeng, Hung, and Yen (2004) investigated whether Chinese readers integrate phonological information across fixations despite the deeper orthography and whether the relationship between the phonetic radical and whole character influences this process. Participants were presented with homophonic previews and orthographic control previews. Half of the target characters were pronounced in the same way as other characters sharing the same phonetic radical (i.e., high consistency) and the other half were not (i.e., low consistency). For high-consistency targets a phonological preview benefit was observed in both first-fixation and gaze duration measures, whereas for low-consistency targets the effect was only observed in gaze durations. Clearly, readers of Chinese integrate phonological information across fixations, and this information is extracted from both the whole character and the phonetic radical.

We have seen that phonological information is integrated across saccades in English and even in Chinese, where there is a far less clear relationship between orthography and phonology. Furthermore, phonological information is extracted at the whole-word level, the character level, and from subunits such as syllables and radicals. While we have discussed English as having a fairly direct link between orthography and phonology in comparison with a language such as Chinese, this relationship is less consistent in English than in many other alphabetic languages. As such, future work on parafoveal phonological processing should perhaps focus more on these other alphabetic languages with more regular coding schemes. It may be, for example, that in these languages (e.g., Spanish) even less skilled readers would show evidence of integrating phonological codes across fixations, unlike less skilled readers of English (Chace et al., 2005).

Morphological Codes

A further form of information that may be integrated across fixations relates to a word's morphology. Often words can consist of more than one morpheme, and therefore a word's constituent morphemes may be used to guide lexical access to the whole-word form (e.g., *cowboy* may be identified via the lexical entries for *cow* and *boy*). Given this, readers may decompose a parafoveal word into its constituent morphemes and integrate these units across fixations. If this were the case, then a clearly defined parafoveal morphological unit could impact subsequent fixations downstream in reading (Lima, 1987).

Several studies have examined this possibility in English (Inhoff, 1989b; Juhasz, White, Liversedge, & Rayner, 2008; Kambe, 2004; Lima, 1987). Researchers have taken the approach of using the boundary technique to provide parafoveal previews to either multimorphemic words (e.g., *revive, cowboy*) or monomorphemic control words (e.g., *rescue, carpet*) where the previews show a plausible morphemic unit (e.g., *reXXXX, carXXX*). The logic behind this manipulation is that a clearly delimited morphological subunit might allow participants to initiate lexical access of the word on this basis. For true multimorphemic words this should be facilitative, since the subunit would be represented as part of the target word's morphological structure. On the other hand, for the monomorphemic control words there should be no advantage beyond an orthographic effect. The results of these studies generally suggest that morphology is not extracted in the parafovea, as there was no difference between the preview effects for multimorphemic and control words. Both Lima (1987) and Kambe (2004) observed no effect for prefixed words (e.g., *revive, dislike*). Lima found no beneficial effects of providing just the prefix (e.g., *disxxxx* for *dislike*) of a multimorphemic word relative to a control word, and Kambe observed no effect of giving either the prefix or the stem (e.g., *xxxlike* for *dislike*). Thus information about prefixes and affixes does not seem to play a role in trans-saccadic integration during English reading. Inhoff (1989b) found a similar pattern of results for words consisting of two morphemes that can stand alone as words (e.g., *cowboy*). Finally, Juhasz et al. (2008) removed a letter from both compound (e.g., *sawdust*) and monomorphemic (e.g., *lettuce*) words in a position that either preserved (e.g., *saw ust, let uce*) or violated (e.g., *sawd st, lett ce*) a morpheme boundary. The preview that preserved the morpheme boundary did not result in faster processing than the preview that violated this boundary, regardless of the type of word. This suggests that participants did not attempt to process the individual morphemes prior to direct fixation.

These studies provide little evidence that English words are decomposed into their constituent morphemes in the parafovea. Similarly, effects have not been observed in Finnish, a language in which spatially concatenated compounds are very common. Bertram and Hyönä (2007) gave participants previews of Finnish compounds that had a short (3 to 4 letters) or long (8 to 11 letters) first constituent, and were on average 12 letters long. The preview consisted of the whole compound or just the first three or four letters. This constituted all of the short first constituents but not all of the long first constituents. Were morphological subunits being integrated across fixations, then a smaller difference between the two preview conditions for the compounds with short first constituents should have occurred, since participants should have gained a greater morphological benefit from the partial preview for the words with short first constituents. However, no interaction was observed between the preview type and first constituent length, suggesting that parafoveally available morphological information was not being used to initiate lexical access.

While morphological units may not be integrated across fixations in English and Finnish, morphological preview effects have been found in Hebrew (Deutsch, Frost, Pelleg, Pollatsek, & Rayner, 2003; Deutsch, Frost, Pollatsek, & Rayner, 2000, 2005). In Hebrew, all verbs and most nouns and adjectives consist of two morphemes. One morpheme is the root, which represents the semantic nature of the word and consists of a series of three consonants. The other is the word pattern, which modifies the root by giving the word its class (e.g., noun, verb, adjective) and other characteristics. Of these two morphemes the root is more important to word meaning, and thus in Hebrew words there are three letters that provide more useful information than the rest of the letters. The two morphemes are interwoven, rather than concatenated. For example, the root morpheme חבר and the word pattern ת___מ take the form of a word with interwoven constituents like תחברמ rather than a concatenated format like תמחבר. Word patterns' structures are highly constrained, such that they can only begin with certain consonants, and each letter imposes a set of transitional probabilities on subsequent letters. Consequently, it is possible for readers of Hebrew to rapidly determine which letters belong to the word pattern and which belong to the root morpheme. In sum, within Hebrew words there are several letters that carry more useful semantic information than the others, and these letters are more easily located and thus extracted from a word (see Frost, this volume, for a more in-depth discussion of the characteristics of Hebrew). Due to this, readers of Hebrew may be able to rapidly decompose a word into its constituent morphemes in the parafovea and then integrate these morphemes across fixations.

Deutsch et al. (2000) first investigated whether the root morpheme is integrated across fixations in Hebrew using a naming paradigm. In this study, an isolated preview of a target word was presented in the parafovea. This preview was either the target word (e.g., תחברמ), the three letters of the root morpheme (e.g., חבר), an orthographic control (e.g., מבת), or an X-string. Participants gained a benefit from the morphological preview relative to the orthographic control, such that they named the target more quickly upon fixating it. Deutsch et al. (2003) extended this finding by showing that a morphological preview benefit is obtained during sentence reading using the boundary paradigm and when the letters of the root morpheme had to be extracted from the letters of the word pattern rather than being presented as an isolated unit. One preview was morphologically related to the target, in that it included the target word's root morpheme within an alternative word pattern. This provided a preview benefit relative to an orthographic control, which shared the same number of letters with the target but was derived from a different root. Participants had clearly extracted the root morpheme in the parafovea and used this to guide lexical access.

Deutsch et al. (2005) also investigated whether the morphological code of the word pattern is integrated across fixations. This was examined for both verbal patterns (i.e., word patterns that combine with the root to form a verb) and nominal patterns (i.e., word patterns that combine with the root to form a noun). An important difference between these two types of word patterns is that, while the verbal patterns possess properties that may guide lexical access, the nominal patterns do not. Specifically, nominal patterns do not have precise semantic characteristics, and the frequency of most nominal patterns is low in comparison with the frequency of the verbal patterns. Deutsch et al. showed that it is possible to gain a morphological preview benefit from a preview consisting of the word pattern in an alternative word in the case of verbs, but not nouns.

In summary, Hebrew readers decompose words into their constituent morphemes in the parafovea and then integrate this information (usually) on the following fixation on the word in order to aid lexical identification. There is clearly a difference between parafoveal morphological processing for readers of Hebrew and readers of English and Finnish. The cross-linguistic difference that may most plausibly account for this is the speed with which it is possible to extract individual morphemes in the parafovea. In Hebrew there are strict rules governing which letters within a word can belong to each morpheme. This is not the case in English, where there are relatively few constraints on where one morpheme ends and another begins. Indeed, the existence of the monomorphemic control words used in the English studies demonstrates this, with it being possible for *re* to either be a prefix or two letters in a monomorphemic word. Thus, readers of Hebrew have stronger cues with which to reliably morphologically decompose words than readers of English, and these cues may partially account for differences in the parafoveal extraction of morphological units.

Semantic Information

Over the past several decades the predominant view has been that semantic information is not integrated across fixations, due to early findings from studies conducted primarily in English. Rayner, Balota, and Pollatsek (1986) presented participants with previews of a target word (e.g., *father*) that were either semantically related (e.g., *mother*), orthographically similar (e.g., *fatlon*), or unrelated (e.g., *circle*). The semantically related previews provided no benefit, suggesting that semantic information was not carried over to subsequent fixations (see Rayner, Schotter, & Drieghe, 2014, for a replication). A similar pattern of results was found in a gaze-contingent naming study (Rayner et al., 1980). Further evidence against semantic information being integrated across fixations was found by Altarriba, Kambe, Pollatsek, and Rayner (2001). In this study, Spanish-English bilinguals read sentences with previews that were translations of a target word that were either orthographically similar (e.g., *crema* as a preview for *cream*) or dissimilar to the target (e.g., *fuerte* as a preview of *strong*), orthographically similar words in the opposite language that were not translations (e.g., *grasa* as a preview for *grass*), or an unrelated word in the opposite language (e.g., *torre* as a preview for *cream*). Since the translation preview shared a semantic representation with the target word, it was hypothesized that significantly more preview benefit might occur for the translation preview than for the orthographically similar nontranslation if semantic information was integrated across fixations. However, the amount of preview benefit was primarily driven by orthography and not semantics, as previews that were translations of the target word did not result in shorter target word fixation durations than nontranslations with the same level of orthographic overlap. This study offers little support for the view that semantic information is integrated across fixations.

Research conducted on semantic preview benefit in Finnish also suggests that semantic information is not integrated across fixations. Hyönä and Häikiö (2005) gave participants parafoveal previews of the target word (e.g., *pentu* 'cub') that were either identical, emotionally arousing (e.g., *penis*), or neutral (e.g., *penni* 'penny'). They hypothesized that if readers extracted semantic information from these previews, then there would be disruption to reading in the emotional condition due to the possibility that this information would be arousing enough to disrupt processing. However, there was no effect of the emotive content of the preview.

Although these studies suggest that semantic information is not integrated across fixations, recent evidence suggests that this is not necessarily the case. Reliable semantic preview effects have now been observed in several studies of Chinese reading. In Chinese the majority of characters include a semantic radical, and therefore there is a more direct link between the orthography and semantics of a word than in alphabetic languages. This makes it more likely that semantic information can be extracted in the parafovea and then integrated on the next fixation. Yan, Zhou, Shu, and Kliegl (2012) examined whether semantic information from both the radical and character level is integrated across fixations. Participants were given an unrelated preview and two different types of semantically related previews. One of the semantically related previews was semantically transparent, in that the meaning of the character was congruent with the meaning of the semantic radical, whereas the other was opaque. None of the previews contained the same semantic radical as the target character; thus any preview benefit could not be due to orthographic confounds. Yan et al. found that both types of semantic preview led to shorter reading times than an unrelated preview, with the semantically transparent preview leading to a larger benefit in gaze duration than the semantically opaque preview. This pattern of results demonstrates that semantic information from both the whole character and the radical is activated in the parafovea and that both types of semantic information are then integrated with semantic information extracted from the target character on fixation. This can be seen from the fact that semantic overlap between the preview and target character reduced target fixation durations and that there was a greater effect when the preview's semantic radical and the target character also shared semantic information. Furthermore, Yan et al. observed larger semantic preview effects when fixation times on the preboundary word were longer (see Hohenstein & Kliegl, 2014, for a discussion of this effect).

Semantic preview effects have also been observed in German. Hohenstein, Laubrock, and Kliegl (2010) found effects in German using parafoveal *fast priming*. In this technique, a nonword preview of the target word is present until readers make a saccade over an invisible boundary prior to the pretarget word. As a saccade is made onto the pretarget word, a display change is triggered. In the Hohenstein et al. (2010) experiment, this led the target word to change to either a semantically related or an orthographically matched preview for a set amount of time before becoming the target word. The amount of time the parafoveal preview was available was varied. At short prime durations (e.g., 35, 60, and 80-ms) there was no semantic preview benefit. At a longer prime duration (125-ms) there was a significant semantic preview benefit of 24 ms. Furthermore, there was a change in this pattern of effects when the target word was made more salient via being presented in bold. Here a significant semantic preview benefit of 18 ms was found at the 80 ms prime duration but no facilitation was found for the 125 ms prime.

The authors claimed that this was due to semantic information being facilitative only up to a certain moment, beyond which the orthographic mismatch overrides the effect. Some caution may be necessary in interpreting these results, as it is not entirely clear how the visual changes that occur in the fast priming technique influence attentional allocation during reading.

Hohenstein and Kliegl (2014) found further evidence for semantic preview benefit in German using the standard boundary paradigm. They found that a semantically related preview (e.g., *Schädel* 'skulls' as a preview for *Knochen* 'bones') was more facilitative than an unrelated preview that shared the same amount of orthographic information with the target word (e.g., *Stiefel* 'boots'). This effect was reliable across fixation time measures over three experiments and averaged 26 ms in gaze duration. Furthermore, the effect endured regardless of whether the target noun was capitalized (in German, nouns are capitalized). This is important, since it may be easier to extract parafoveal semantic information for nouns in German because the capitalization may give readers a salient cue to the syntactic class of the parafoveal word, allowing for more processing resources to be allocated to that word than might otherwise be the case. Furthermore, there was an effect of pretarget fixation duration that was similar to that reported by Yan et al. (2012), such that there was a greater semantic preview benefit following longer fixations on the pretarget word.

The final study we will consider in this section is that of Schotter (2013). In this study investigating reading of American English, participants were given two different types of semantically related previews. The first type (e.g., *rollers* as a preview for *curlers*) was highly related to the target (7.5 on a 9-point rating scale in a norming study) and maintained the sentence meaning (7.2 on a 9-point rating scale). The second type (e.g., *styling*) was less semantically related (5.6) and maintained the sentence meaning to a lesser extent (4.9). Unrelated previews (e.g., *suffice*; 2.4 and 1.9 on the rating scales) were also included. All three previews shared a similar amount of orthographic information with the target. Relative to unrelated previews, the highly related previews led to shorter fixation durations on the target word (16 and 19 ms in gaze durations across two experiments). There was no benefit from the less semantically related previews. Furthermore, the extent to which the preview changed the meaning of the sentence predicted fixation times on the target word. Schotter argued that this suggests the lack of effects in English in prior studies arose because the semantic relationship between the preview and the target word did not preserve meaning to the same degree that her stimuli did. For example, Rayner et al. (1986) used target-preview pairs such as *father-mother, ocean-river*, and *sick-well* which, while semantically related to each other, did not necessarily share the same meaning.

In sum, the evidence regarding whether semantic information is integrated across fixations is currently mixed. Some studies have failed to show clear effects, while other studies do appear to show effects, often under specific experimental circumstances. It is not possible at present to provide a coherent explanation of the current state of this aspect of processing—in some senses it is quite contradictory. Further research is necessary in order to gain a clearer understanding.

The Integration of Information from Word *N* + 2

The preceding sections have all focused on how various types of information about the upcoming word ($n + 1$) are integrated across fixations. Recently, however, research has begun to investigate whether information from word $n + 2$ is also integrated across fixations (Angele & Rayner, 2011; Angele, Slattery, Yang, Kliegl, & Rayner, 2008; Kliegl, Risse, & Laubrock, 2007; Rayner, Juhasz, & Brown, 2007; Risse & Kliegl, 2012).

To investigate the integration of information from word $n + 2$ across fixations, researchers have manipulated the preview of a word while it is two words to the right of fixation, with the preview changing to the target as a saccade is made onto the pretarget word (word $n + 1$). Any effect of this manipulation would suggest that readers are extracting information from word $n + 2$ when it is in the parafovea and integrating this information during subsequent fixations. Rayner et al. (2007) presented participants with either a correct or incorrect preview of a target word and manipulated whether the boundary was directly before the target word or directly before the pretarget word. As such, the incorrect preview was either visible as word $n + 1$ or word $n + 2$. The preview manipulation only had an effect when the preview was visible as word $n + 1$. Thus, Rayner et al. did not observe evidence for the integration of information from word $n + 2$. Kliegl et al. (2007) further investigated this issue. In their study, word $n + 1$ was always three letters long, thus ensuring that the preview of word $n + 2$ was as close to central vision as was reasonably possible. Furthermore, they tested for effects of the $n + 2$ preview on fixation times on both word $n + 1$ and $n + 2$. While the *n* + 2 preview did not affect fixations on word *n* + 2, it did affect fixations on word *n* + 1, suggesting that information from word *n* + 2 was extracted (see Risse & Kliegl, 2012, for a discussion and test of why this effect appeared on word *n* + 1). Angele et al. (2008) orthogonally manipulated previews of word *n* + 1 and *n* + 2 and ensured that word *n* + 1 was always at least four characters long. They found that, while there were reliable *n* + 1 preview effects, there were no effects of the *n* + 2 preview. The posited reason for the discrepancies across studies is the length and processing difficulty of word *n* + 1. When word *n* + 1 exceeds three characters it is more difficult to process, and therefore word *n* + 2 is less likely to be processed before a saccade is made across the boundary. Furthermore, even when word *n* + 2 is processed, information extraction occurs less efficiently, since it is further into the parafovea.

Angele and Rayner (2011) manipulated whether readers received identity or nonword previews of a three-letter word *n* + 1 and a word *n* + 2 that was on average seven letters long. While *n* + 2 preview effects were found when there was an identity preview of word *n* + 1, there was no effect when it was a nonword. Thus when word *n* + 1 cannot be lexically processed (due to it being a nonword), information from word *n* + 2 does not appear to be integrated.

More recently Cutter, Drieghe, and Liversedge (2014) found an *n* + 2 preview effect even when word *n* + 1 was longer than three letters. In this study, word *n* + 1 (e.g., *teddy*) was on average 5.65 letters long and formed a spaced compound (e.g., *teddy bear*) with word *n* + 2 (e.g., *bear*). Participants were given either a correct preview of both constituents, of only the first constituent, of only the second constituent, or of neither constituent. When a correct preview of the first constituent of the spaced compound was available for processing in the parafovea, there was a significant *n* + 2 preview effect of 27 ms. However, when a correct preview of the first constituent was not available for processing in the parafovea, there was no effect of the preview of word *n* + 2. This demonstrates that while *n* + 2 preview effects are not typically observed given a long word *n* + 1, this can be modulated by the extent to which word *n* + 2 forms a single multiword unit with word *n* + 1. Furthermore, it shows that the absence of *n* + 2 preview effects in prior studies was not due to visual limitations.

In summary, there is evidence for information from word *n* + 2 being extracted in the parafovea. This information is arguably integrated across fixations in English, but only under specific circumstances. That is, the studies reviewed suggest that word *n* + 1 must be short and easy to process for information from word *n* + 2 to be extracted and integrated across fixations. Furthermore, even when such effects are observed they are small (e.g., 7 to 20 ms) when compared with effects of word *n* + 1 (e.g., 20 to 50 ms). The one exception to this is when word *n* + 2 was part of a spaced compound, an issue that we will return to later.

Modulating Factors

So far, we have discussed the extent to which information is integrated across fixations as if this is an invariant process. However, several factors have been shown to modulate this process. The first is foveal load and the second is the extent to which the foveal and parafoveal word can be considered a single unit.

Foveal load refers to the difficulty of processing on any particular fixation. When the currently fixated word is difficult to process, then foveal load is high. It has been argued that increased foveal processing load results in reduced parafoveal processing (Henderson & Ferreira, 1990; White, Rayner, & Liversedge, 2005), reducing the extent to which information may be integrated across saccades. Henderson and Ferreira manipulated foveal load via either a word frequency or a syntactic manipulation and presented participants with a correct or incorrect preview using the boundary paradigm. Significant effects of the preview type were only observed when foveal load was low. The effect of the foveal word's frequency on preview benefit has also been observed by White et al. (2005).

While several studies have shown that foveal load modulates the parafoveal preview benefit, research by Drieghe, Rayner, and Pollatsek (2005) suggests that this is not always the case. In this study, foveal load was varied using the same frequency manipulation as in earlier studies and participants were given a preview of a three-letter target word. However, the size of the preview benefit was the same for the high and low foveal load conditions. Drieghe et al. proposed that the absence of an interaction may have been due to the short parafoveal words being processed differently than the longer parafoveal words used in other investigations of foveal load. However, it is unclear why the length of a parafoveal word would determine the extent to which foveal load influences parafoveal processing. More work is required to explore this effect.

A second factor that influences how far into the parafovea information is extracted from and then integrated across fixations is the degree to which the foveal and parafoveal text is unified spatially and linguistically. This is an issue that has been touched on throughout the current chapter. In terms of spatial unification, a larger preview benefit is observed when the preview is of the end of the fixated word (e.g., 151 ms in gaze durations on the second half of a word in Drieghe et al., 2010) than when the preview is of the word to the right of fixation (e.g., an average of 41 ms for dissimilar letters in Hyönä et al., 2004). Even less of an effect is observed from previews of word $n + 2$, with the literature only finding effects of between 7 and 20 ms. The one exception to this was Cutter et al.'s (2014) study, in which a 27-ms effect was observed in gaze duration on word $n + 1$ when word $n + 2$ formed a spaced compound with word $n + 1$. This effect suggests that whether two physically separated parafoveal words form a single lexical unit influences the amount of information integrated across fixations.

The results of several studies suggest that the lexical unification of information within a fixated word also influences the extent and time course of the integration of information from the end of this word. As discussed in the section on within-word integration, it has been found that people differentially integrate information from the end of unspaced compounds and monomorphemic words (Drieghe et al., 2010). Furthermore, Häikiö et al.'s (2010) study suggested that information from the end of unspaced compounds was integrated differently depending on whether the compound was identified as a single lexical unit or two separate lexemes. Häikiö et al. showed that when an unspaced compound was identified as a single lexical unit, incorrect information at the end of the second constituent was integrated early enough to affect fixations on the first constituent. This was not the case when the unspaced compounds were processed as two separate lexemes. Thus, this research suggests that the time course in which information is integrated across fixations is modulated by whether the information in the fovea and parafovea are processed as part of the same lexical unit.

There is also evidence that a greater amount of information is integrated from word $n + 1$ when it forms part of a larger unit with the fixated word (Inhoff, Starr, & Shindler, 2000; Juhasz, Pollatsek, Hyönä, Drieghe, & Rayner, 2009). Inhoff et al. examined preview benefit for the second word of spaced compounds (e.g., *traffic light, fairy tale, video tape*). This study found a considerably larger preview benefit than is usual between words, such that there was a 91-ms effect of a dissimilar preview in comparison with the average of 41 ms (Hyönä et al., 2004). Furthermore, the manipulation affected fixation times on the first constituent in a similar manner to preview manipulations within monomorphemic words (Drieghe et al., 2010) and frequent unspaced compounds (Häikiö et al., 2010). Juhasz et al. also found a larger than usual preview benefit for the second constituents of spaced compounds (34 ms vs. an average of -7 ms for studies using an equivalent level of disruption), although this study did fail to find significant differences between spaced compounds and adjective-noun pairs, for which there was a 21-ms effect. The findings of both Inhoff et al. and Juhasz et al. suggest that a greater amount of information may be extracted from a parafoveal word if it forms part of a larger unit with the foveal word. Furthermore, Inhoff et al.'s finding indicated that this parafoveal information may have been integrated earlier than is typical for a parafoveal word that does not form a single unit with the fixated word.

To summarize, several factors have been found to influence the extent to which parafoveal information is integrated across fixations. One is foveal load, with preview benefit effects being reduced when the fixated word is difficult to process. The second factor is the extent to which the information in the parafovea forms a single unit with the fixated word.

Conclusion

We have seen that a large variety of information is integrated across fixations, both from the end of a single word and from a parafoveal word. The integration of information from word $n + 1$ operates on the basis of abstract codes for word characteristics such as orthography, phonology, semantics, and, in the case of Hebrew, morphology. There are several interesting cross-linguistic differences that influence the information that is preferentially integrated across fixations. For example, readers of English preferentially integrate word-initial letters, Chinese readers integrate the final radical of the parafoveal character, and Hebrew readers integrate morphological codes. The underlying reason for these differences may well be the extent to which the information allows the reader to initiate lexical access. For readers of English, the phonological code granted by the word-initial letters may be most useful, while in Hebrew the root morpheme may provide more useful information for activating appropriate lexical candidates. Finally, in Chinese the final radical may provide more discriminative information to activate a limited set of character candidates. As such, research suggests that information is integrated across fixations on the basis of the most useful information for identifying words in a particular language. While the research shows that a large amount of information about word *n* + 1 is integrated across fixations, the same is not true for word *n* + 2, with preview manipulations to this word having small effects that only occur under optimal conditions. Finally, the way in which readers integrate information across fixations is influenced both by foveal load and whether the parafoveal text forms a larger unit with either the foveal text or more distal parafoveal text.

References

Altarriba, J., Kambe, G., Pollatsek, A., & Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish-English bilinguals. *Perception & Psychophysics, 63,* 875–890.

Angele, B., & Rayner, K. (2011). Parafoveal processing of word n + 2 during reading: Do the preceding words matter? *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1210–1220.

Angele, B., Slattery, T. J., Yang, J., Kliegl, R., & Rayner, K. (2008). Parafoveal processing in reading: Manipulating *n* + 1 and *n* + 2 previews simultaneously. *Visual Cognition, 16,* 697–707.

Ashby, J., & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements. *Language and Cognitive Processes, 19,* 391–426.

Ashby, J., Treiman, R., Kessler, B., & Rayner, K. (2006). Vowel processing during silent reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 416–424.

Balota, D. A., & Rayner, K. (1991). Word recognition processes in foveal and parafoveal vision: The range of influences of lexical variables. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading* (pp. 198–232). Hillsdale, NJ: Erlbaum.

Bélanger, N. N., Mayberry, R. I., & Rayner, K. (2013). Orthographic and phonological preview benefits: Parafoveal processing in skilled and less-skilled deaf readers. *Quarterly Journal of Experimental Psychology, 66,* 2237–2252.

Bertram, R., & Hyönä, J. (2007). The interplay between parafoveal preview and morphological processing in reading. In R. P. G. van Gompel, M. H. Fisher, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 391–407). Oxford, England: Elsevier.

Briihl, D., & Inhoff, A. W. (1995). Integrating information across fixations during reading: The use of orthographic bodies and exterior letters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 55–67.

Chace, K. H., Rayner, K., & Well, A. D. (2005). Eye movements and phonological preview: Effects of reading skill. *Canadian Journal of Experimental Psychology, 59,* 209–217.

Cui, L., Yan, G., Bai, X., Hyönä, J., Wang, S., & Liversedge, S. P. (2013). Processing of compound-word characters in reading Chinese: An eye-movement-contingent display change study. *Quarterly Journal of Experimental Psychology, 66,* 527–547.

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2014). Preview benefit in English spaced compounds. Journal of Experimental Psychology: Learning, Memory, and Cognition, 40, 1778–1789.

Deutsch, A., Frost, R., Pelleg, S., Pollatsek, A., & Rayner, K. (2003). Early morphological effects in reading: Evidence from parafoveal preview benefit in Hebrew. *Psychonomic Bulletin & Review, 10,* 415–422.

Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2000). Early morphological effects in word recognition in Hebrew: Evidence from parafoveal preview benefit. *Language and Cognitive Processes, 15,* 487–506.

Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2005). Morphological preview benefit effects in reading: Evidence from Hebrew. *Language and Cognitive Processes, 20,* 341–371.

Drieghe, D. (2011). Parafoveal-on-foveal effects on eye movements during reading. In S. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 839–855). Oxford, England: Oxford University Press.

Drieghe, D., Pollatsek, A., Juhasz, B. J., & Rayner, K. (2010). Parafoveal processing during reading is reduced across a morphological boundary. *Cognition, 116,* 136–142.

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 954–969.

Fitzsimmons, G., & Drieghe, D. (2011). The influence of number of syllables on word skipping during reading. *Psychonomic Bulletin & Review, 18,* 736–741.

Häikiö, T., Bertram, R., & Hyönä, J. (2010). Development of parafoveal processing within and across words in reading: Evidence from the boundary paradigm. *Quarterly Journal of Experimental Psychology, 63,* 1982–1998.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 417–429.

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 166–190.

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1150–1170.

Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese.* Hillsdale, NJ: Erlbaum.

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye movement-contingent display change study. *Memory & Cognition, 32,* 523–532.

Hyönä, J., & Häikiö, T. (2005). Is emotional content obtained from parafoveal words during reading? An eye movement analysis. *Scandinavian Journal of Psychology, 46,* 475–483.

Inhoff, A. W. (1989a). Parafoveal processing of words and saccade computation during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 544–555.

Inhoff, A. W. (1989b). Lexical access during eye fixations in reading: Are word access codes used to integrate lexical information across interword fixations? *Journal of Memory and Language, 28,* 441–461.

Inhoff, A. W., Starr, M., & Shindler, K. L. (2000). Is the processing of words during eye fixations in reading strictly serial? *Perception & Psychophysics, 62,* 1474–1484.

Johnson, R. L. (2007). The flexibility of letter coding: Nonadjacent letter transposition effects in the parafovea. In R. P. G. van Gompel, M. H. Fisher, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 425–440). Oxford, England: Elsevier.

Johnson, R. L., & Dunne, M. D. (2012). Parafoveal processing of transposed-letter words and nonwords: Evidence against parafoveal lexical activation. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 191–212.

Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 209–229.

Juhasz, B. J., Pollatsek, A., Hyöna, J., Drieghe, D., & Rayner, K. (2009). Parafoveal processing within and between words. *Quarterly Journal of Experimental Psychology, 62,* 1356–1376.

Juhasz, B. J., White, S. J., Liversedge, S. P., & Rayner, K. (2008). Eye movements and the use of parafoveal word length information in reading. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1560–1579.

Kambe, G. (2004). Parafoveal processing of prefixed words during eye fixations in reading: Evidence against morphological influences on parafoveal preprocessing. *Perception & Psychophysics, 66,* 279–292.

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word n+2. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1250–1255.

Lima, S. D. (1987). Morphological analysis in sentence reading. *Journal of Memory and Language, 26,* 84–99.

Liu, W., Inhoff, A. W., Ye, Y., & Wu, C. (2002). Use of parafoveally visible characters during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1213–1227.

Martin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin, 81,* 899–917.

McConkie, G. W., & Rayner, K. (1975). The span of effective stimulus during a fixation in reading. *Perception & Psychophysics, 17,* 578–586.

McConkie, G. W., & Rayner, K. (1976a). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society, 8,* 365–368.

McConkie, G. W., & Rayner, K. (1976b). Identifying the span of the effective stimulus in reading: Literature review and theories of reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 137–162). Newark, NJ: International Reading Association.

McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixation in reading? *Perception & Psychophysics, 25,* 221–224.

Miellet, S., & Sparrow, L. (2004). Phonological codes are assembled before word fixation: Evidence from boundary paradigm in sentence reading. *Brain and Language, 90,* 299–310.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 148–162.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422.

Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology, 40,* 473–483.

Rayner, K., Juhasz, B. J., & Brown, S. J. (2007). Do readers obtain preview benefit from word *n*+2? A test of serial attention shift versus distributed lexical processing models of eye movement control in reading. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 230–245.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology, 12,* 206–226.

Rayner, K., Schotter, E. R., & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit in reading revisited. *Psychonomic Bulletin & Review, 21,* 1067–1072.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of model. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1188–1200.

Risse, S., & Kliegl, R. (2012). Evidence for delayed parafoveal-on-foveal effects from word *n*+2 in reading. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 1026–1042.

Schotter, E. R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language, 69,* 619–633.

Tsai, J., Lee, C., Tzeng, O. J. L., Hung, D. L., & Yen, N. (2004). Use of phonological codes for Chinese characters: Evidence from processing of parafoveal preview when reading sentences. *Brain and Language, 91,* 235–244.

White, S. J., Bertram, R., & Hyönä, J. (2008). Semantic processing of previews within compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 988–993.

White, S. J., Rayner, K., & Liversedge, S. P. (2005). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A re-examination. *Psychonomic Bulletin & Review, 12,* 891–896.

Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2012). Lexical and sublexical semantic preview benefits in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 1069–1075.

Zhou, X., & Marslen-Wilson, W. (1999). Sublexical processing in reading Chinese. In J. Wang, H.-C. Chen, R. Radach, & A. Inhoff (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 37–64). Hillsdale, NJ: Erlbaum.

CHAPTER

18 Direct Lexical and Nonlexical Control of Fixation Duration in Reading

Eyal M. Reingold, Heather Sheridan, *and* Erik D. Reichle

Abstract

This chapter focuses on the *eye-mind link* in reading, or how perceptual and cognitive processes influence when and where the eyes move when people read. The chapter is organized into four parts. First, early theoretical accounts of the eye-mind link are reviewed, and key findings that are problematic for these accounts are discussed. Timing constraints on the eye-mind link that have been derived from behavioral and neurophysiological studies are examined, along with the implications of these constraints for current models of eye movement control in reading. Next, evidence is provided for the direct control of eye movements during reading from a number of eye-movement experiments that have used distributional analyses and survival analyses to examine the time course over which perceptual and/or lexical variables affect fixation durations during reading. Finally, the findings of the review are summarized, and possible directions for future research on this topic are presented.

Key Words: eye-mind link, lexical processing, distributional analysis, time course, eye movement control, fixation duration

The subjective experience reported by skilled readers often includes descriptions of smooth and continuous gliding of gaze position across the line of text. However, 135 years ago Emile Javal (see Huey, 1908/1968) observed that readers' eyes do not sweep smoothly across the text, but instead make a sequence of high-velocity eye movements during which vision is largely suppressed (referred to as *saccades*). The periods between saccades during which the eye is relatively still and visual information is encoded are referred to as *fixations*. In order to enable the extraction of fine visual features of the text during reading, saccades align the high-acuity *foveal* region of the visual system (the central 2° of vision) with the portion of the text that is being encoded. This is necessary because visual acuity declines rapidly as the location of the retinal image of the text shifts from the fovea to the surrounding *parafoveal* region (1° to 5° on either side of fixation) and *peripheral* region (beyond 5° on either side of fixation) of the visual field.

More than a century of research employing eye-movement measures and techniques reveals that the subjective experience of effortless skilled reading obscures an intricate orchestration of a wide array of oculomotor, perceptual, lexical, linguistic, and cognitive processes (for reviews, see Rayner, 1998, 2009; and Schotter & Rayner, this volume). Indeed, the rapidly growing use of eye-tracking methodology in reading research was motivated in part by the assumption that the record of the locations and durations of fixations is a valid and useful trace measure for inferring the processes underlying reading performance. Consequently, extensive empirical and theoretical efforts have been directed at developing models of eye-movement control during reading. These models attempt to explain the factors that determine when the eyes move (i.e., fixation durations) as well as where the eyes move (i.e., fixation locations). One source of considerable controversy concerns the influence of cognitive processes on the location and duration of fixations (henceforth, the *eye-mind link*). There is little doubt that, for reading to be efficient, some form of coordination between the progression of the eyes over the text and comprehension processes must occur. This fact was acknowledged even when the behaviorist era in psychology discouraged any attempt to infer mental processes from eye-movement data. For example, Tinker (1958), perhaps the most prominent eye movement and reading researcher of that era, stated that "in addition to seeing clearly during a fixation pause, the reader must comprehend the ideas and relationships involved. Actually, therefore, pause duration includes perception time plus thinking time" (p. 218). As an illustration of the importance of the eye-mind link to skilled reading, consider the case in which the synchronization between eye movements and comprehension processes breaks down during *mind-wandering* episodes (sometimes referred to as *mindless reading*). There is now considerable evidence indicating that while the general characteristics of eye-movement parameters are similar between normal reading and mindless reading, there are important quantitative and qualitative differences in the observed pattern of eye movements across these conditions (e.g., Reichle, Reineberg, & Schooler, 2010; Schad, Nuthmann, & Engbert, 2012). For example, whereas the durations of fixations during normal reading are modulated by local properties of the text being read (e.g., the frequency of the word being fixated; Inhoff & Rayner, 1986; Just & Carpenter, 1980), this is less true during mindless reading (e.g., word-frequency effects are absent or much attenuated in size; Reichle et al., 2010; Schad et al., 2012).

Given the critical importance of the eye-mind link in reading, our primary goal is to examine the validity of the *direct lexical-control hypothesis* (e.g., Rayner & Pollatsek, 1981), which argues that lexical and linguistic processes play a major role in determining the durations of individual fixations during reading. Recently, Reingold, Reichle, Glaholt, and Sheridan (2012) proposed a general framework for the investigation of the possible mechanisms that might mediate the control of fixation durations in reading. In the present chapter, we present an extended version of this approach. Accordingly, we begin with a brief overview of the central ideas, terminology, and assumptions concerning eye movement control in reading. Next we outline some of the timing constraints that must be considered in evaluating the feasibility of the direct control of eye movements in reading and their implications for models of eye-movement control in reading. Finally, we review several categories of empirical findings that make a compelling case for a tight link between visual and lexical processing and fixation duration in reading. It is important to note that in the present chapter we exclusively review evidence from reading studies that employed European alphabetic languages.

Eye Movement Control in Reading: Brief History and Basic Concepts

Models of eye movement control in reading attempt to explain the considerable variability in both saccade amplitude (mean 7–9 letter spaces; range 1–20 letter spaces) and fixation duration (mean 200–250 ms; range 50–1,000 ms). Prior to the 1970s, there was a great deal of skepticism over whether cognitive processes could have an impact on eye movements (for reviews, see Rayner, 1978; Rayner & McConkie, 1976). This was in part based on the *cognitive-lag assumption*, or the widely held belief that cognition was simply too slow to have an impact on eye movement parameters (e.g., Kolers, 1976). For that reason, it was commonly believed that skilled reading was characterized by a consistent and nonvariable pattern of forward eye movements, with the exception of occasional *regressions*, or backward eye movements to previously read text. These early conceptualizations of eye movement control in reading (called *minimal-control models* by Rayner, 1978) ascribed variability in saccade amplitude and fixation duration to random variation or physiological factors and postulated no meaningful association between eye movements and properties of the fixated text.

Minimal-control models were unable to explain the finding that difficult text passages produced longer fixation durations and shorter saccade amplitudes. This led to the development of *indirect-control models*, which assumed that cognition can have delayed (i.e., non-real-time) influences on fixation durations and saccade amplitudes based on global processing difficulty and contextual factors (e.g., task instructions). For example, *gain-control models* (e.g., Kolers, 1976) assumed that fixation duration and saccade amplitude varied randomly around a preset rate that (on average) allowed enough time to encode the text, but with cognition being used to adjust the overall reading rate via feedback to the oculomotor system. Similarly, buffer-control models (e.g., Bouma & DeVoogd, 1974) assumed that readers maintain an optimal reading rate by increasing the speed of their eye movements whenever their cognitive buffer is empty and decreasing the speed whenever the buffer is full. Thus, indirect-control models could account for effects of text difficulty in a manner that does not require assuming immediate cognitive influences.

Interestingly, however, more than fifty years prior to the introduction of indirect-control models, Buswell (1922) reported a finding that could not be explained by these models. Specifically, Buswell observed that fixation durations on unfamiliar words were longer than fixations on other words, and given that indirect-control models assume that the properties of the fixated word could have no bearing on the duration of that fixation, such models could not account for this particular finding. Although the theoretical significance of this finding might have been underappreciated at the time, it later spurred the development of another class of eye movement control models—those that posit direct control. In marked contrast to indirect-control models which assume delayed adjustment of fixation duration based on the global processing difficulty of the text, *direct-control models* argue for an immediate fixation-by-fixation adjustment based on the properties of the local stimulus (i.e., the word being fixated). Thus in the context of the eye-movement control literature, the direct/indirect dichotomy often incorporates both the immediate/delayed and local/global distinctions. For example, in the Reader model proposed by Just and Carpenter (1980) two key assumptions were incorporated: (1) the *immediacy assumption,* which hypothesized that readers will attempt to interpret each word as soon it is encountered, even though the initial interpretation might occasionally turn out to be wrong in light of subsequent visual, lexical, semantic, or syntactic processing, and (2) the *eye-mind assumption,* which postulated that the currently fixated word is processed completely before the eyes move to the next word. Thus, according to this model, eye movement control entails both immediate and local control (i.e., the fixation duration on a word exclusively reflects the processing of that word).

Unfortunately, there is considerable vagueness and inconsistency in the literature concerning the use of the terms "immediate" and "local" as part of the definition of direct-control. For example, it is often unclear whether immediate processing of the fixated word (word *n*) implies that such processing begins with the first fixation on word *n*. This simple view is not tenable, because it has long been known (see Rayner, 1978) that information about word *n* can be extracted during a fixation on the previous word (i.e., the *parafoveal preview* of word *n* during the fixation on word $n - 1$). This suggests that the concept of immediacy should be extended to include the preprocessing of word *n* during the parafoveal preview. Similarly, it is often unclear whether local control implies control based entirely on the processing of word *n*, or whether it should be extended to include the processing of the adjacent words, word $n - 1$ and word $n + 1$. To avoid the ambiguities inherent in the terms "immediate" and "local," Reingold et al. (2012) defined direct control as the assumption that the processing of the properties of the fixated word (word *n*) influences the timing of the saccade terminating that fixation, regardless of whether this processing was initiated while word *n* was foveated or when it was parafoveally processed during fixations on the previous word (word $n - 1$).

Early direct-control models made vastly different assumptions concerning the role of ongoing lexical and linguistic processes in controlling eye movements in reading (see Rayner, 1998, 2009, for reviews). *Oculomotor* models (e.g., McConkie, Kerr, Reddix, & Zola, 1988; McConkie, Kerr, Reddix, Zola, & Jacobs, 1989; O'Regan, 1990) assumed that nonlexical, low-level information determines eye movement control in reading. In contrast, *processing* models (e.g., Henderson & Ferreira, 1990; Just & Carpenter, 1980; Morrison, 1984; Pollatsek & Rayner, 1990; Rayner & Pollatsek, 1981, 1989) advocated a critical role for lexical and attentional processes. The nature of this debate has become more complex in recent years with several mixed-control models that incorporate combinations of direct and indirect control mechanisms. To help clarify these issues, Reingold et al. (2012) proposed a taxonomy of eye movement control mechanisms based on two orthogonal dimensions: the type of eye-movement control that is assumed (direct vs. indirect) and the type of information (lexical vs. nonlexical) used in the control. Reingold et al. further proposed two nonmutually exclusive general types of direct-control mechanisms: a *triggering mechanism,* whereby some index of word *n* processing fluency triggers the programming of the saccade terminating a fixation, and (2) an *interference mechanism,* whereby some index of word *n* processing difficulty delays the initiation and/or execution of the saccade terminating a fixation (see Figure 18.1). Based on the preceding taxonomy, the assumption that the completion of some stage of lexical processing initiates saccadic

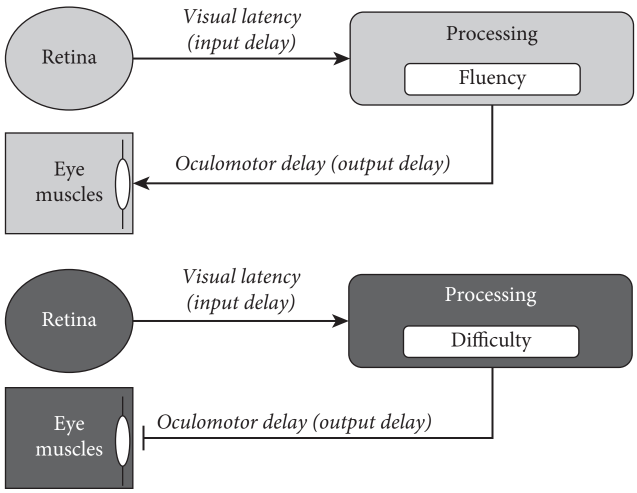


Fig. 18.1 A schematic diagram showing two general types of direct-control mechanisms: a triggering mechanism whereby some index of the fluency of processing of the fixated word triggers the programming of the saccade terminating the fixation (top panel), and an interference mechanism, whereby some index of the difficulty of processing of the fixated word produces delays in the initiation or execution of the saccade terminating the fixation (bottom panel). In either triggering or interference mechanisms, the first discernible influence of the processing of the fixated word on the timing of the saccade terminating that fixation would be expected to be constrained by the minimum input (visual), output (oculomotor), and processing delays.

programming (e.g., Reichle, Pollatsek, Fisher, & Rayner, 1998) is an example of a direct-control triggering mechanism based on lexical information. In contrast, the proposal that lexical processing difficulty inhibits saccade initiation (e.g., Engbert, Nuthmann, Richter, & Kliegl, 2005) represents a direct-control interference mechanism based on lexical information. Examples of direct-control mechanisms based on the processing of prelexical visual information also include both triggering (e.g., McDonald, Carpenter, & Shillcock, 2005) and interference versions (e.g., Reilly & Radach, 2006). Furthermore, examples of indirect lexical mechanisms include a cognitive-buffer assumption (e.g., Bouma & DeVoogd, 1974) and the assumption that saccades are initiated at an average rate that is occasionally adjusted to support the overall pace of lexical processing (e.g., Kolers, 1976). Conversely, saccade triggering by a random timer (e.g., Engbert et al., 2005) constitutes an example of an indirect nonlexical control mechanism.

It is important to emphasize that the taxonomy proposed by Reingold et al. (2012) provides a basis for classifying the mechanisms that control eye movements during reading, and not the models themselves, which often incorporate more than a single mechanism. Interestingly, most current models of eye movement control propose either triggering or interference mechanisms but not both (but see Reingold et al., 2012). Another curious aspect of current models is the fact that they largely ignore (but see Ehrlich & Rayner, 1983; Kennedy, 1998; Kennedy, Pynte & Ducrot, 2002; Reingold et al., 2012; Risse & Kliegl, 2012) an important distinction made by Rayner and McConkie (1976) between eye movement control that is driven by the output of various processes versus eye movement control that is based on a system that monitors the progress of such processes (henceforth called the *process-monitoring hypothesis*). In the final section of this chapter we briefly discuss the potential advantage of incorporating process monitoring into current models of eye movement control in reading. Thus over the past four decades, an intensive investigation of the nature of eye movement control during reading has generated a wealth of findings as well as considerable controversy. Recently, the theoretical focus in this field has shifted away from qualitative models and toward quantitatively implemented models (e.g., Engbert et al., 2005; McDonald et al., 2005; Reichle et al., 1998; Reilly & Radach, 2006; Salvucci, 2001). Importantly, these models require precise (quantitative) architectures that specify the mechanisms that control the eyes during reading, which then raises questions about whether the theoretical assumptions and parameters that are incorporated into the models are feasible given the available evidence about neural delays in the perceptual and oculomotor systems. These timing constraints are explored in the next section.

Timing Constraints and Models of Direct Control in Reading

To be plausible, the direct control of fixation duration in reading must respect fairly severe timing constraints. As shown in Figure 18.1, for either triggering or interference mechanisms the fastest influence of the processing of the properties of word *n* on the timing of the saccade terminating a fixation on word *n* should equal the sum of the duration of the following intervals: (1) *minimum input delay*, or the time required for information about word *n* to reach the cortical systems where processing of word *n* begins; (2) *processing delay*, or the time required for the processing system (e.g., visual, lexical) or the process-monitoring system to establish that the encoding of word *n* is progressing well (i.e., processing fluency) or that progress is stalled (i.e., processing difficulty); and (3) *minimum output delay*, or the time required to transmit a facilitatory (triggering) or inhibitory (interference) signal to the oculomotor system and brainstem circuitry that is ultimately responsible for moving the eyes. In this section we briefly summarize what has been learned about direct-control timing constraints from studies using two different electophysiological measures—*event-related potentials* (ERPs) and *megnetoencephalography* (MEG).

Reichle and Reingold (2013) reviewed studies using ERP and MEG methodology in order to evaluate the minimum latencies with which visual information can be propagated from the eyes to the brain (i.e., the *retina-brain lag*), and the processing delays associated with visual encoding and lexical processing of printed words. As explained earlier, several models of eye movement control in reading posit a tight link between the eye and mind, with lexical processing directly triggering most decisions about when to start programming a saccade to move the eyes from one word to the next. One potential problem with this theoretical assumption, however, is that it may violate neurophysiological constraints imposed by the time required to first encode visual information, then complete some amount of lexical processing, and then program a saccade. Note that this objection is simply a restatement of the cognitive-lag hypothesis that was discussed earlier. The findings from the studies reviewed by Reichle and Reingold (2013) are summarized in Figure 18.2. Given that the estimates derived from these studies are inherently conservative because they correspond to the first statistically reliable effects of experimental variables on neuroimaging markers, Figure 18.2 displays the minimum, mean, and maximum values of each estimate. As can be seen by an inspection of this figure, mean estimates of 60, 92, and 148 ms were obtained for the retina–brain lag, visual encoding time, and lexical processing time, respectively.

If one accepts that the latency for lexical processing is approximately 150 ms, then the remaining time in the average fixation duration of 200 to 250 ms (i.e., 50–100 ms) is insufficient for completing all of the operations that are necessary for initiating and programming the saccade that terminates the fixation. This is because eye-movement experiments (e.g., Becker & Jürgens, 1979; Rayner, Slowiaczek, Clifton, & Bertera, 1983) suggest that saccades require 120 to 200 ms to program. Furthermore, Reingold et al. (2012) found that word-frequency effects (a marker of lexical processing) on target words were discernible more than 100 ms earlier when target words were available for parafoveal processing (i.e., normal reading) than when parafoveal processing was prevented. Consideration of these findings led Reichle and Reingold (2013) to argue that, given that the average fixation duration on word *n* lasts approximately 250 ms, the temporal constraints imposed by visual processing and saccadic programming are too severe to permit direct lexical control of fixation duration without a significant amount of parafoveal processing (i.e., preprocessing of word *n* from word *n* – 1).

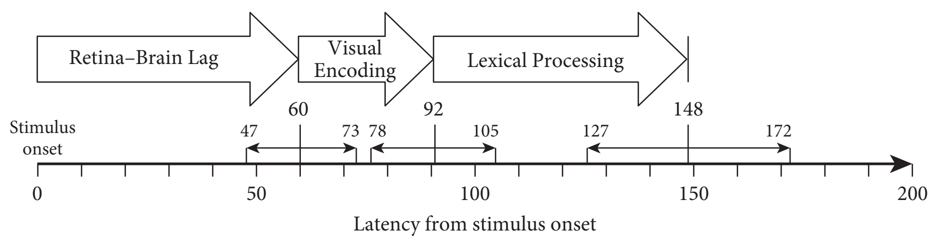


Fig. 18.2 A schematic diagram, based on Reichle and Reingold's (2013) review of studies employing ERP and MEG methodology, displaying mean and range estimates of input latencies involved in the arrival of text information to cortical processing centers (retina–brain lag), visual encoding, and lexical processing.

The preceding estimates derived from ERP and MEG studies have broader implications for direct-control models that incorporate a triggering mechanism. Most importantly, these estimates underscore the fact that triggering must be based on shallow, partial, and rapid processing of word *n*. This insight is consistent with the trend in the literature. Recall that the Reader model proposed by Just and Carpenter (1980) postulated a triggering mechanism based on the complete lexical and postlexical processing of word *n*. More recently, advocates of the triggering mechanism have adopted a much more modest view about how much word processing must be completed to trigger an eye movement. For example, in the EMMA model (Salvucci, 2001), the saccadic program to move the eyes from word *n* to word *n* + 1 is triggered by the encoding of word *n* but not its subsequent processing. Similarly, the E-Z Reader model (Reichle et al., 1998; see also Reichle & Sheridan, this volume) assumes that a superficial stage of lexical processing triggers saccadic programming. This early stage of processing has been described as corresponding to either a rapid recognition response that reflects the familiarity of word *n* (Reichle & Perfetti, 2003) or orthographic processing (Reichle Tokowicz, Liu, & Perfetti, 2011). By either interpretation the completion of this stage indicates that access to the meaning of word *n* is imminent, so that the oculomotor system can begin programming a saccade to move the eyes to word *n* + 1 (see also Reingold & Rayner, 2006; Sheridan & Reingold, 2013). The amount of cognitive processing that, by assumption, is necessary to initiate saccadic programming is even further reduced in the SERIF (McDonald et al., 2005) and Glenmore (Reilly & Radach, 2006) models, where visual word encoding rather than lexical processing provides the trigger that determines when the eyes will move.

Finally, it is important to examine the temporal constraints that are relevant for direct-control models that incorporate an interference mechanism. Although there were several proposals of interference mechanisms in the literature on eye movement control in reading (e.g., Engbert et al., 2005; Nuthmann & Henderson, 2012; Reingold et al., 2012; Yang & McConkie, 2001), the most complete instantiation of such a mechanism is incorporated into the SWIFT model (e.g., Engbert et al., 2005). According to this model, saccades are triggered by an autonomous random timer (which is an exemplar of an indirect-control mechanism) and not by the completion of some cognitive process. Importantly, lexical processing difficulty can modulate fixation durations by actively inhibiting the timer so that it cannot initiate new saccadic programs. The assumption here is that by preventing the initiation of saccadic programming, fixations will be lengthened, allowing additional time for lexical processing. In order to avoid unnecessary interruptions to the reading process, it is likely that establishing processing difficulty would require a more conservative threshold and longer delay than establishing processing fluency. However, experiments using the *saccadic inhibition paradigm* in reading (e.g., Reingold & Stampe, 1999, 2000, 2002, 2003, 2004) suggest that once processing difficulty has been established, the minimum latency for inhibiting saccades via a direct-control interference mechanism should be on the order of 20 to 30 ms (see Reingold & Stampe, 2000, 2002, for reviews of the timing constraints involved in saccadic inhibition based on evidence from both behavioral and neurophysiological studies). In addition, temporal constraints are further relaxed for a reading model such as SWIFT due to the assumption that multiple words are processed in parallel, thereby permitting the parafoveal processing of word *n* during fixations on word *n* – 1 and word *n* – 2.

The Empirical Case for Direct Control

It is now generally accepted that fixation times in reading are influenced by both lexical and nonlexical variables. Importantly, only the analysis of the duration of the very first fixation on word *n* (henceforth, *first-fixation duration*) offers a straightforward opportunity for examining the possible influence of direct control of fixation duration (see Reingold, Yang, & Rayner, 2010, for a related discussion). This is because refixations on the word, either immediate (i.e., on the *first pass* through the text) or delayed (i.e., rereading the word following a regression back to that region of text), might be influenced by a variety of other factors such as the memory for the previous foveal analysis. Nonlexical variables with an influence on first-fixation duration include word length (e.g., Just & Carpenter, 1980; Rayner, Sereno, & Raney, 1996) and initial landing position (e.g., Kliegl, Nuthmann, & Engbert, 2006; McDonald et al., 2005; Nuthmann, Engbert, & Kliegl, 2005, 2007; Vitu et al., 2001, 2007). In addition, lexical variables that were demonstrated to influence first-fixation duration in reading include word frequency (Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Reingold et al., 2012; see White, 2008, for a review), contextual constraint or predictability (Ehrlich & Rayner, 1981; Rayner, Ashby, Pollatsek, & Reichle, 2004; Rayner & Well, 1996; Sheridan & Reingold, 2012a), lexical ambiguity (e.g., Duffy, Morris, & Rayner, 1988; Rayner & Duffy, 1986; Sheridan & Reingold, 2012b; Sheridan, Reingold, & Daneman, 2009; see Duffy, Kambe, & Rayner, 2001, for a review), and age of acquisition (e.g., Juhasz & Rayner, 2006). However, simply showing that a variable affects the mean first-fixation duration is insufficient for demonstrating direct control, because such an effect might be due to a very infrequent but sizable influence of a variable that does not have an impact on the vast majority of fixations.

The analysis of distributions is inherently more suitable than the analysis of mean fixation durations for determining the time course of the influence of variables on fixation duration. Accordingly, in this section of the chapter we primarily focus on reviewing the results from recent studies that employed distributional analysis methods in order to study the time course of direct control of fixation duration in reading. Specifically, we summarize findings from studies using an ex-Gaussian fitting procedure that was introduced by Staub, White, Drieghe, Hollway, and Rayner (2010) and a survival analysis technique that was introduced by Reingold et al. (2012). In addition, we explore how these findings provide convergent evidence for the conclusions from earlier investigations that employed the *text-onset delay* paradigm (e.g., Dambacher, Slattery, Yang, Kliegl, & Rayner, 2013; Hohenstein, Laubrock, & Kliegl, 2010; Inhoff, Eiter, & Radach, 2005; Luke et al., 2013; Morrison, 1984; Nuthmann & Henderson, 2012; Rayner & Pollatsek, 1981) and the *disappearing text* paradigm (Blythe, Liversedge, Joseph, White, & Rayner, 2009; Ishida & Ikeda, 1989; Liversedge et al., 2004; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981; Rayner, Liversedge, & White, 2006; Rayner, Liversedge, White, & Vergilino-Perez, 2003; Rayner, Yang, Castelhano, & Liversedge, 2011).

Direct Control and Evidence From Distributional Analysis Methods

Recently, ex-Gaussian fitting (Staub et al., 2010) and a survival analysis technique (Reingold et al., 2012) were used to provide valuable information about the time course of lexical and nonlexical influences on fixation durations during reading. The characteristic shape of the empirical distribution of fixation durations resembles a Gaussian normal distribution, but the right tail of the distribution is typically skewed to some degree. Consequently, the ex-Gaussian distribution, which is the convolution of the Gaussian normal distribution and an exponential distribution, appears to be a good candidate for modeling the empirical distributions of fixation durations. The ex-Gaussian distribution was previously established to generate excellent fits for reaction time distributions (see Balota & Yap, 2011, for a review), and Staub et al. (2010) first demonstrated its great promise for modeling fixation durations in reading. The ex-Gaussian distribution can be specified with the parameters μ (which corresponds to the mean of the Gaussian distribution), σ (which corresponds to the standard deviation of the Gaussian distribution), and τ (the exponential component that indicates the degree of skew). Importantly, the sum of the μ and τ parameters from the ex-Gaussian distribution equals the mean of the empirically obtained fixation-duration distribution, and a comparison of the best-fitting μ and τ parameters can reveal whether a variable's impact on the mean empirical fixation durations is due to an overall shift in the location of the distribution and/or a change in the degree of skew. Whereas a shift effect (i.e., a difference in μ between conditions) indicates that the variable has an early-acting influence on the majority of fixation durations, a skew effect (i.e., a difference in τ between conditions) indicates that the variable primarily influences the long fixation durations.

Using this logic, Staub et al. (2010) fitted an ex-Gaussian distribution to individual participants' first-fixation duration distributions on both high- and low-frequency target words. These fits indicated that the low-frequency distribution was significantly shifted to the right of the high-frequency distribution, and that the low-frequency distribution also exhibited greater positive skew than the high-frequency distribution. The finding that word frequency caused a shift in the distributions indicates that this lexical variable has an impact on both short and long fixations (see Rayner, 1995); this was predicted by the direct lexical-control hypothesis. However, the more pronounced positive skew for the low-frequency distribution also indicates that long fixations are differentially affected by word frequency.

Another approach for examining the distributions of fixation duration was introduced by Reingold et al. (2012). This approach was aimed at deriving a precise estimate for the first discernible influence of a variable on fixation duration. Specifically, Reingold et al. (2012) explored the onset of the influence of a lexical variable (word frequency: high vs. low frequency) and a nonlexical variable (initial landing position on the fixated word: central vs. outer location) on first-fixation duration using a novel survival analysis technique. In this procedure, for a given time *t*, the percentage of first fixations with a duration greater than *t* is referred to as the percent *survival* at time *t*. Thus, when *t* equals zero survival is at 100%, but then declines as *t* increases. For each variable and condition, Reingold et al. (2012) calculated survival curves and computed confidence intervals using a bootstrap resampling procedure (Efron & Tibshirani, 1994) in order to examine the earliest point at which the survival curves for high- versus low-frequency words and for central- versus outer-location fixations first started to diverge (henceforth referred to as the *divergence point*). Importantly, Reingold et al. (2012) argued that these divergence points provide estimates of the earliest significant influence of word frequency and initial landing position on first-fixation duration, and consequently that this analysis technique is uniquely suited for testing the feasibility of lexical and nonlexical direct-control mechanisms.

In addition, in order to test the role of parafoveal processing in enabling direct control of first-fixation duration, Reingold et al. (2012) employed the *gaze-contingent boundary paradigm* (Rayner, 1975) to contrast invalid and valid preview trials. As shown in Figure 18.3, in invalid preview trials an unrelated letter string occupied the position of the target word and was replaced with the target word during the saccade that crossed an invisible boundary located just to the left of that word. In contrast, in the valid preview condition, which represents normal reading, target words were always displayed throughout the trial. Finally, in order to replicate and extend the findings reported by Staub et al. (2010), in addition to survival analysis, Reingold et al. (2012) also used ex-Gaussian fitting in order to investigate the time course of the influence of word frequency and initial landing position in both the valid and invalid preview conditions.

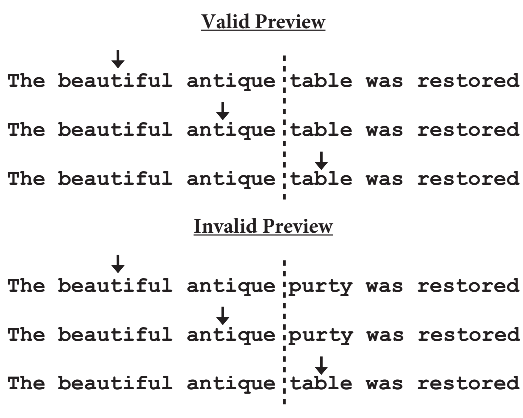


Fig. 18.3 An illustration of the valid and invalid preview conditions that were used by Reingold et al. (2012). Three consecutive fixations are shown in each condition (the arrow above each sentence denotes the location of the fixation). In invalid preview trials, an unrelated letter string (e.g., *purty*) occupied the position of the target word (e.g., *table*). During the saccade that crossed an invisible boundary (illustrated by the dotted vertical line) the target word replaced the unrelated letter string. In contrast, in the valid preview condition, which represents normal reading, target words were always displayed throughout the trial.

Figure 18.4 displays some of the key findings reported by Reingold et al. (2012). As can be seen by an inspection of the divergence points shown in the figure, under normal reading conditions (i.e., the valid preview condition shown in the top two panels), word frequency and initial landing position produced an equally rapid influence on first-fixation duration (divergence point 145 ms). In contrast, preventing parafoveal processing of target words (i.e., the invalid preview condition shown in the bottom two panels) had a dramatic impact on the onset of the word-frequency effect (divergence point 256 ms) but did not influence the timing of the effect of the initial landing position (divergence point 142 ms). Similarly, the results of the ex-Gaussian analysis indicated a strong effect of preview validity on word frequency. In the valid preview condition, the low-frequency distribution exhibited a rightward shift (i.e., μ effect) and a greater skew (i.e., τ effect), whereas in the invalid preview condition only a τ effect was obtained. Furthermore, the pattern of results for the initial fixation position was qualitatively different, with a μ effect but not a τ effect occurring in both the valid and invalid preview conditions. Reingold et al. (2012) argued that, taken together, these findings clearly demonstrated the critical role of parafoveal processing in enabling direct lexical control of fixation duration despite the tight constraints imposed by neural delays in the perceptual and oculomotor system. In addition, Reingold et al. (2012) suggested that the results from the distributional analyses of the influence of initial landing position on fixation duration provide strong support for the existence of a direct nonlexical control mechanism. Such a mechanism might involve the processing of visual cues that are extracted very early during the first fixation on the word (e.g., Vitu et al., 2001, 2007), or reliance on the internal monitoring of the signal sent to the oculomotor muscles (i.e., the efference copy) prior to the execution of the saccade that first moved

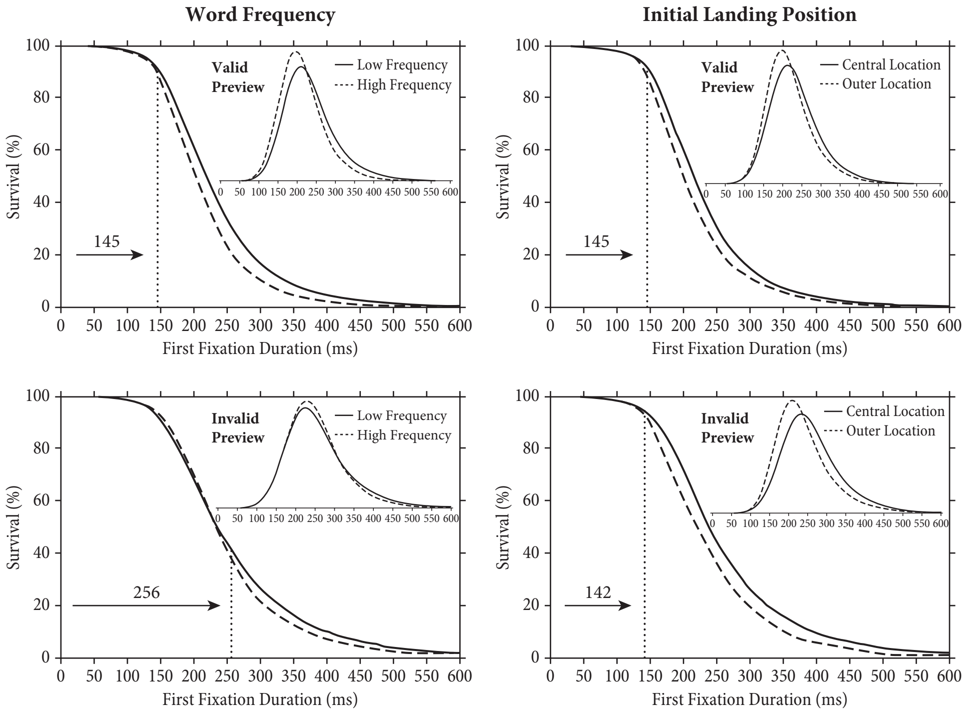


Fig. 18.4 Key findings reported by Reingold et al. (2012) based on the survival analysis of first-fixation durations for both valid preview (top panels) and invalid preview (bottom panels), as a function of word frequency (left panels) or initial landing position (right panels). In each panel, the divergence point is marked by a vertical dashed line and the ex-Gaussian density functions are shown in the top right section of the panel (see text for details).

the eyes to the word (Nuthmann et al., 2005, 2007; see also Engbert et al., 2005; McDonald et al., 2005; Reichle, Warren, & McConnell, 2009).

As summarized in Table 18.1, following Staub et al. (2010) and Reingold et al. (2012), the influence of a variety of lexical and nonlexical variables on first-fixation duration have been examined using ex-Gaussian fitting or survival analysis. During normal reading, all of the variables that have been examined to date have shown rapid influences on first-fixation durations, as indicated by a significant μ effect (i.e., a shift effect) for all of the variables in the table, and by survival analysis results showing that the first discernible influence of these variables occurs 112 to 146 ms after fixation onset. Taken together, the ex-Gaussian and survival analysis results provide strong support for direct control by demonstrating that a variety of variables can have a rapid influence on both short and long fixation durations.

The survival analysis technique can also be used to compare the time course of variables under different reading conditions. As previously discussed, Reingold et al. (2012) showed that word frequency effects were substantially delayed in the absence of a parafoveal preview. Similarly, Sheridan, Rayner, and Reingold (2013) investigated the hypothesis that lexical processing would be delayed by the removal of interword spaces (see Rayner, Fischer, & Pollatsek, 1998). To do this, Sheridan et al. conducted an experiment in which high- and low-frequency words were embedded either in normal (English) text or unsegmented text in which the blank spaces between the words were replaced with random numbers. Sheridan et al. demonstrated that word frequency divergence points were delayed (by 23–40 ms) in the unsegmented condition relative to the normal reading condition. Taken together, these findings indicate that rapid lexical processing is facilitated by the availability of parafoveal preview and by interword spacing, and these findings demonstrate the usefulness of the survival analysis technique for testing predictions about the time course of lexical processing (see also Inhoff & Radach, 2014; Schad, Risse, Slattery, & Rayner, 2014).

In addition to considering μ effects and survival divergence points, it is also informative to consider the relationship between the μ and τ parameters shown in Table 18.1. For example, Staub and Benatar (2013) suggested that τ effects

Table 18.1 Summary of Results From Studies Using Ex-Gaussian Modeling and/or Survival Analysis.

Variable	Article Reference	Fixation Duration	Mu (μ)	Sigma (σ)	Tau (τ)	Divergence Point
Word Frequency (Low-High Frequency)	Staub et al. (2010) Exp.1	25***	16***	8†	10†	—
	Staub et al. (2010) Exp. 2	27***	13*	4	15*	—
	Reingold et al. (2012) Valid Preview	20***	9**	2	11**	145 (9%)
	Reingold et al. (2012) Invalid Preview	9**	–8	–2	16***	256 (60%)
	Sheridan et al. (2013) Exp.1A, Normal	20***	7**	3	13***	112 (5%)
	Sheridan et al. (2013) Exp.1A, Unsegmented	20***	7*	4*	13***	152 (9%)
	Sheridan et al. (2013) Exp.1B, Normal	19***	9**	4	9**	146 (10%)
	Sheridan et al. (2013) Exp.1B, Unsegmented	20***	7*	6**	12***	169 (15%)
Predictability (Low-High Predictability)	Staub (2011)	16***	14***	5	3	—
	Sheridan & Reingold (2012a)	8**	8*	–1	1	140 (10%)
Lexical Ambiguity (Subordinate-Dominant context)	Sheridan & Reingold (2012b)	12***	8**	0	5	139 (8%)
Preview effect (Invalid-Valid)	Reingold et al. (2012) High Frequency	37***	24***	6*	13*	133 (5%)
	Reingold et al. (2012) Low Frequency	25***	10*	3	14**	172 (20%)
Initial Landing Position (Central-Outer Location)	Reingold et al. (2012) High Frequency	19***	14***	5*	3	141 (8%)
	Reingold et al. (2012) Low Frequency	26***	21***	12***	4	144 (8%)
	Reingold et al. (2012) Valid Preview	14***	11***	6**	2	145 (10%)
	Reingold et al. (2012) Invalid Preview	25***	20***	8*	2	142 (6%)
Stimulus Quality (Degraded-Normal)	White & Staub (2012)	20***	19***	4	1	—
	White & Staub (2012)	52***	47***	24***	3	—
	Glaholt et al. (2014)	50***	32***	17***	18**	141 (10%)

Note: For the contrasts shown above, † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.
All divergence points were significant ($p < .001$), and the percentages of fixations with a duration that is shorter than the divergence point are shown in brackets next to each divergence point. Exp. = Experiment. See original studies for details.

might reflect the probability of processing disruption (i.e., episodes when normal processing fails and some type of repair or recovery process is necessary), whereas μ effects might reflect variations in processing difficulty. For example, the word frequency variable might cause a μ effect because low-frequency words are more difficult to process, and might cause a τ effect because low-frequency words are more likely to produce episodes of processing disruption (e.g., failure to access or integrate a word's meaning). In support of the hypothesis that μ and τ effects for fixation-duration distributions reflect functionally distinct processes, Staub and Benatar (2013) performed a correlation analysis and demonstrated that the μ and τ parameters varied independently across subjects. Moreover, as shown in Table 18.1, the two parameters appear to be functionally dissociable because it is possible to find instances in which μ effects occur in the absence of τ effects and vice versa. For example, whereas word frequency produced a τ effect in the absence of a μ effect in the invalid preview condition (Reingold et al., 2012), a variety of other variables produced μ effects in the absence of a τ effect, including the initial fixation position (Reingold et al., 2012), predictability (Sheridan & Reingold, 2012a; Staub, 2011), lexical ambiguity (Sheridan & Reingold, 2012b), and stimulus quality (White & Staub, 2012). Furthermore, word frequency produced μ and τ effects during normal reading, as did the contrast between valid and invalid parafoveal preview (Reingold et al., 2012).

Thus, to summarize, Table 18.1 lists a wide range of ex-Gaussian results, including simultaneous μ and τ effects, selective μ effects, and a selective τ effect. As discussed by Staub and Benatar (2013), the suggestion that μ and τ reflect functionally distinct processes has been controversial (for further discussion of this issue in the reaction-time literature see Matzke & Wagenmakers, 2009), and future work is required to further explore this issue. However, the evidence provided by this method in combination with the evidence of rapid survival divergence points provides support for both lexical and nonlexical direct control by demonstrating a fast-acting influence of several important reading-related variables on fixation durations.

Direct Control and the Text-Onset Delay and Disappearing Text Paradigms

The findings from studies employing distributional analysis methods are also consistent with evidence in support of lexical and nonlexical direct control that was previously obtained using two gaze-contingent techniques, the text-onset delay paradigm (e.g., Dambacher et al., 2013; Hohenstein et al., 2010; Inhoff et al., 2005; Luke et al., 2013; Morrison, 1984; Nuthmann & Henderson, 2012; Rayner & Pollatsek, 1981) and the disappearing text paradigm (Blythe et al., 2009; Ishida & Ikeda, 1989; Liversedge et al., 2004; Rayner et al., 1981; Rayner et al., 2006; Rayner et al., 2003; Rayner et al., 2011). In this section we briefly outline the key findings from these investigations.

It is often argued that the text-onset delay paradigm offers the most straightforward approach for demonstrating that information extracted during a fixation has an impact on the timing of the saccade terminating that fixation (i.e., direct control of fixation duration). As illustrated in Figure 18.5, the basic procedure used in this paradigm involves delaying the availability of a portion of the text during an interval at beginning of each fixation. This dead time at the beginning of each fixation is implemented by replacing the text with a visual mask during the preceding saccade and reinstating the text at a certain delay from the onset of the fixation (the range of delays used across studies was 0–350 ms).

Importantly, direct-control models would predict that the saccade terminating the fixation should be delayed and that the magnitude of this delay should be proportional to length of the interval during which the text information was unavailable. In contrast, indirect-control models assume that the properties of the fixated word do not have an impact on the duration of the fixation and consequently would predict no difference in fixation times as a function of the length of the delay. The key findings that were initially reported by Rayner and Pollatsek (1981) and Morrison (1984) provided support for both direct and indirect influences on fixation duration (i.e., mixed-control). Specifically, the general pattern of an increase in fixation duration that was proportional to the length of the delay provided support for the existence of a direct-control mechanism, which influenced a large population of fixations. However, there was a distinct population of fixations with durations that were shorter than the delay (especially for longer text-onset delays). These fixations, which were terminated prior to the removal of the mask, probably reflected the influence of an indirect-control mechanism and/or saccadic programming that was initiated based on parafoveal processing of the text during the previous fixation (see Dambacher et al., 2013; Hohenstein et al., 2010; Inhoff et al., 2005, for investigations of the role of parafoveal processing using the text-onset delay paradigm). One potential complication in interpreting the findings obtained using the text-onset delay paradigm is due to the fact that the visible display change due to the removal of the mask results in saccadic inhibition (e.g., Reingold & Stampe, 1999, 2000, 2002, 2003, 2004). However, although saccadic inhibition clearly contributes to prolonging fixation duration, it could not fully account for the qualitative pattern of findings that is observed across delays (see also Luke, Nuthmann, & Henderson, 2013; Nuthmann & Henderson, 2012; Slattery, Angele, & Rayner, 2011), and consequently it is safe to conclude that the results from studies employing the text-onset delay paradigm provide strong support for a sizable influence of direct control on fixation duration in reading.

The disappearing text paradigm (see Figure 18.5 for an illustration) is essentially the inverse of the text-onset paradigm in that text is either removed or masked after some amount of time following fixation onset (Blythe et al., 2009; Ishida & Ikeda, 1989; Liversedge et al., 2004; Rayner et al., 1981; Rayner et al., 2006; Rayner et al., 2003; Rayner et al., 2011). The key finding from this paradigm is that reading is relatively unimpaired provided that the text is visible during the first 50 to 60 ms during the fixation (i.e., before it is blanked or masked). Most importantly, despite the disappearance or masking of the text, the fixation duration on the word is strongly influenced by the frequency of the word, indicating that direct lexical influences on fixation duration primarily depend on information that is extracted during the parafoveal preview period and during the first 50–60 ms of foveation on the word. In further support of the importance of parafoveal processing in the disappearing text paradigm, Rayner et al. (2006) demonstrated that simultaneous removal of the fixated word (word *n*) and the word to the right of the fixation (word $n + 1$) substantially disrupts reading performance.

Conclusions and Future Directions

This chapter attempted to outline some of the main contours of a controversy concerning the control of the duration of fixations during reading that was over a century in the making. We would strongly argue that at the present juncture, due to the growing emphasis on computational modeling and the emerging convergence across multiple lines of empirical evidence, a consensus position can be reached that would represent true progress in this field. Specifically, given the decisive accumulated evidence for direct lexical and nonlexical influences on individual fixation durations, the current literature on the topic is beginning to transcend the existence proof stage and has instead become more

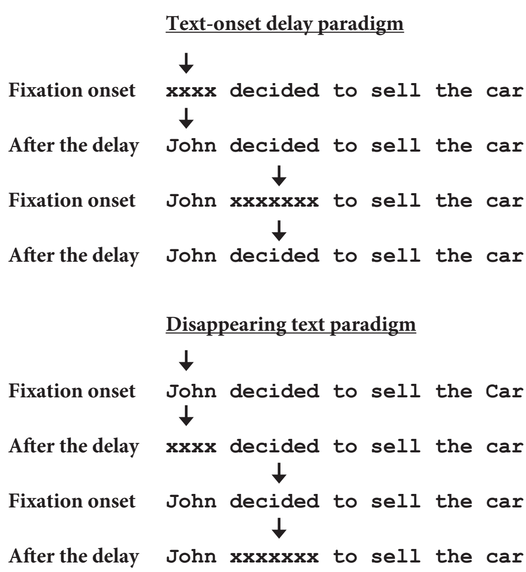


Fig. 18.5 An illustration of the text-onset delay paradigm and disappearing text paradigm. The arrow above each sentence denotes the location of the fixation. A display change occurs in both paradigms following a certain interval from the onset of each fixation. In the text-onset delay paradigm, a mask that is shown at the onset of the fixation is replaced by the fixated word after a delay. In contrast, in the disappearing text paradigm, the fixated word that is displayed normally at the onset of the fixation is replaced by a mask or blanked after a delay (see text for details).

focused on deriving quantitative estimates for the timing and magnitude of these influences. Such estimates constitute critical benchmarks for the development and testing of computational models of eye movement control in reading.

Another point of emerging consensus calls into question the widely held assumption that lexical and linguistic processing are simply too sluggish to produce real-time adjustment of fixation durations (i.e., the cognitive-lag hypothesis). Specifically, when the essential role of parafoveal lexical processing is taken into account, the tight timing constraints imposed by neural delays are no longer inconsistent with the hypothesis of direct lexical control of fixation duration in reading. Based on our review, we would suggest that mechanisms of eye movement control that are incorporated into current computational models are likely to be somewhat oversimplified. Processing of the fixated word could potentially result in both facilitatory and inhibitory influences on the timing of the saccade terminating the fixation. Furthermore, models might need to include both lexical and nonlexical direct control mechanisms (possibly combining both triggering and interference versions of direct control), as well as a variety of indirect control influences. The fact that models of eye movement control require more than one mechanism to explain how the processing of a word influences the time spent looking at that word suggests that the term "eye-mind link" is probably a gross misnomer—to fully explain the patterns of looking times that are observed in reading, eye-movement models have to posit multiple links that interact in various ways.

One immediate advantage that comes from adopting this more complex perspective about how the mind interacts with the eyes is that it suggests that the models might be improved by considering how the various basic mechanisms that we have been discussing are configured in reading. Because existing models only occupy a small part of the space of possible configurations, we suspect that our understanding of eye movement control in reading might really benefit from examining those parts of the space that have been ignored. In particular, as we argued earlier, we believe that the process-monitoring hypothesis that was first suggested by Rayner and McConkie almost 40 years ago deserves more consideration. Given the complexity just described, it is not parsimonious to assume that multiple processes and mechanisms have dedicated interfaces to the oculomotor system. It seems to us far more intuitive that at least some of the influences on fixation duration might be mediated by a system that could monitor the pattern of activation across a variety of prelexical, lexical, and postlexical processes in order to monitor their progress and infer fluency and/or difficulty without accessing the content of the various input and output representations that are involved. Such a process monitoring system could also provide a common structure for the summation of multiple influences and for interfacing with the oculomotor system. Clearly, further exploration is required in order to enhance our understanding of the nature of eye movement control in reading.

Acknowlegments

The writing of this chapter was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) operating grant to Eyal Reingold, an NSERC postdoctoral fellowship awarded to Heather Sheridan, and support from the Center for Vision and Cognition, University of Southampton.

References

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science, 20,* 160–166.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research, 19,* 967–983.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research, 49,* 1583–1591.

Bouma, H., & DeVoogd, A. H. (1974). On the control of eye saccades in reading. *Vision Research, 14,* 273–284.

Buswell, G. T. (1922). *Fundamental reading habits: A study of their development*. Chicago, IL: University of Chicago Press.

Dambacher, M., Slattery, T. J., Yang, J., Kliegl, R., & Rayner, K. (2013). Evidence for direct control of eye movements during reading. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1468–1484.

Duffy, S. A., Kambe, G., & Rayner, K. (2001). The effect of prior disambiguating context on the comprehension of ambiguous words: Evidence from eye movements. In D. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 27–43). Washington, DC: American Psychological Association.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27,* 429–446.

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. Boca Raton, FL: Chapman & Hall.

Ehrlich, S. E, & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20,* 641–655.

Ehrlich, K. & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior, 22,* 75–87.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112,* 777–813.

Glaholt, M. G., Rayner, K., & Reingold, E. M. (2014). A rapid effect of stimulus quality on the durations of individual fixations during reading. *Visual Cognition, 22,* 377–389.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 417–429.

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1150–1170.

Huey, E. B. (1968). *The psychology and pedagogy of reading.* Cambridge, MA: MIT Press (Original work published 1908).

Inhoff, A. W., Eiter, B. M., & Radach, R. (2005). Time course of linguistic information extraction from consecutive words during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 979–995.

Inhoff, A. W., & Radach, R. (2014). Parafoveal preview benefits during silent and oral reading: Testing the parafoveal information extraction hypothesis. *Visual Cognition, 22,* 354–376.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40,* 431–439.

Ishida, T., & Ikeda, M. (1989). Temporal properties of information extraction during reading studied by a text-mask replacement technique. *Journal of Optical Society A: Optics and Image Science, 6,* 1624–1632.

Juhasz, B. J., & Rayner, K. (2006). The role of age of acquisition and word frequency in reading: Evidence from eye fixation durations. *Visual Cognition, 13,* 846–863.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329–354.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade-off. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 149–179). Oxford, England: Elsevier.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology, 55A,* 1307–1337.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General, 135,* 12–35.

Kolers, P. A. (1976). Buswell's discoveries. In R. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes* (pp. 373–395). Hillsdale, NJ: Erlbaum.

Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., & Kentridge, R. W. (2004). Eye movements when reading disappearing text: Is there a gap effect in reading? *Vision Research, 44,* 1013–1024.

Luke, S. G., Nuthmann, A., & Henderson, J. M. (2013). Eye movement control in scene viewing and reading: Evidence from the stimulus onset delay paradigm. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 10–15.

Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review, 16,* 798–817.

McConkie, G. W, Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations in words. *Vision Research, 28,* 1107–1118.

McConkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D., & Jacobs, A. M. (1989). Eye movement control during reading: II. Frequency of refixating a word. *Perception & Psychophysics, 46,* 245–253.

McDonald, S. A., Carpenter, R. H. S., & Shillcock, R. C. (2005). An anatomically-constrained, stochastic model of eye movement control in reading. *Psychological Review, 112,* 814–840.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 667–682.

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research, 45,* 2201–2217.

Nuthmann, A., Engbert, R., & Kliegl, R. (2007). The IOVP effect in mindless reading: Experiment and modelling. *Vision Research, 47,* 990–1002.

Nuthmann, A., & Henderson, J. M. (2012). Using CRISP to model global characteristics of fixation durations in scene viewing and reading with a common mechanism. *Visual Cognition, 20,* 457–494.

O'Regan, J. K. (1990). Eye movements and reading. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 395–453). Amsterdam, the Netherlands: Elsevier.

Pollatsek, A., & Rayner, K. (1990). Eye movements and lexical access in reading. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 143–164). Hillsdale, NJ: Erlbaum.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85,* 618–660.

Rayner, K. (1995). Eye movements and cognitive processes in reading, visual search, and scene perception. In J. M. Findlay, R. Walker, & R. W. Kentridge (Eds.), *Eye movement research: Mechanisms, processes and applications* (pp. 3–22). Amsterdam, the Netherlands: North Holland.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422.

Rayner, K. (2009). The Thirty-Fifth Sir Frederick Bartlett lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62,* 1457–1506.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 720–732.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14,* 191–201.

Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38,* 1129–1144.

Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 167–179.

Rayner, K., Liversedge, S. P., & White, S. J. (2006). Eye movements when reading disappearing text: The importance of the word to the right of fixation. *Vision Research, 46,* 310–323.

Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading disappearing text: Cognitive control of eye movements. *Psychological Science, 14,* 385–389.

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research, 16,* 829–837.

Rayner, K., & Pollatsek, A. (1981). Eye movement control during reading: Evidence for direct control. *Quarterly Journal of Experimental Psychology, 33A,* 351–373.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1188–1200.

Rayner, K., Slowiaczek, M. L., Clifton, C., & Bertera, J. H. (1983). Latency of sequential eye movements: Implications for reading. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 912–922.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3,* 504–509.

Rayner, K., Yang, J., Castelhano, M. S., & Liversedge, S. P. (2011). Eye movements of older and younger readers when reading disappearing text. *Psychology and Aging, 26,* 214–223.

Reichle, E. D., & Perfetti, C. A. (2003). Morphology in word identification: A word experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading, 7,* 219–237.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science, 21,* 1300–1310.

Reichle, E. D., & Reingold, E. M. (2013). Neurophysiological constraints on the eye-mind link. *Frontiers in Human Neuroscience, 7,* 361.

Reichle, E. D., Tokowicz, N., Liu, Y., & Perfetti, C. A. (2011). Testing an assumption of the E-Z Reader model of eye-movement control during reading: Using event-related potentials to examine the familiarity check. *Psychophysiology, 48,* 993–1003.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16,* 1–21.

Reilly, R., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research, 7,* 34–55.

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z Reader model. *Psychological Science, 17,* 742–746.

Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive Psychology, 65,* 177–206.

Reingold, E. M., & Stampe, D. M. (1999). Saccadic inhibition in complex visual tasks. In W. Becker, H. Deubel, & T. Mergner (Eds.), *Current oculomotor research: Physiological and psychological aspects* (pp. 249–255). London, England: Plenum.

Reingold, E. M., & Stampe, D. M. (2000). Saccadic inhibition and gaze contingent research paradigms. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 119–145). Amsterdam, the Netherlands: Elsevier.

Reingold, E. M., & Stampe D. M. (2002). Saccadic inhibition in voluntary and reflexive saccades. *Journal of Cognitive Neuroscience, 14,* 371–388.

Reingold, E. M., & Stampe, D. M. (2003). Using the saccadic inhibition paradigm to investigate saccadic control in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movements* (pp. 347–360). Amsterdam, the Netherlands: Elsevier.

Reingold, E. M., & Stampe D. M. (2004). Saccadic inhibition in reading. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 194–211.

Reingold, E. M., Yang, J., & Rayner, K. (2010). The time course of word frequency and case alternation effects on fixation times in reading: Evidence for lexical control of eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1677–1683.

Risse, S., & Kliegl, R. (2012). Evidence for delayed parafoveal-on-foveal effects from word $n + 2$ in reading. *Journal of Experimental Psychology: Human Perception & Performance, 38,* 1026–1042.

Salvucci, D. D. (2001). An integrated model of eye movements and visual encoding. *Cognitive Systems Research, 1,* 201–220.

Schad, D. J., Nuthmann A., & Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels. *Cognition, 125,* 179–194.

Schad, D. J., Risse, S., Slattery, T., & Rayner, K. (2014). Word frequency in fast priming: Evidence for immediate cognitive control of eye movements during reading. *Visual Cognition, 22,* 390–414.

Sheridan, H., Rayner, K., & Reingold, E. M. (2013). Unsegmented text delays word identification: Evidence from a survival analysis of fixation durations. *Visual Cognition, 21,* 38–60.

Sheridan, H., & Reingold, E. M. (2012a). The time course of predictability effects in reading: Evidence from a survival analysis of fixation durations. *Visual Cognition, 20,* 733–745.

Sheridan, H., & Reingold, E. M. (2012b). The time course of contextual influences during lexical ambiguity resolution: Evidence from distributional analyses of fixation durations. *Memory & Cognition, 40,* 1122–31.

Sheridan, H., & Reingold, E. M. (2013). A further examination of the lexical processing stages hypothesized by the E-Z Reader model. *Attention, Perception, & Psychophysics, 75,* 407–414.

Sheridan, H., Reingold, E. M., & Daneman, M. (2009). Using puns to study contextual influences on lexical ambiguity resolution: Evidence from eye movements. *Psychonomic Bulletin & Review, 16,* 875–881.

Slattery, T. J., Angele, B., & Rayner, K. (2011). Eye movements and display change detection during reading. *Journal of*

Experimental Psychology: Human Perception and Performance, 37, 1924–1938.

Staub, A. (2011). The effect of lexical predictability on distributions of eye fixation durations. *Psychonomic Bulletin & Review, 18,* 371–376.

Staub, A., & Benatar, A. (2013). Individual differences in fixation duration distributions in reading. *Psychonomic Bulletin & Review, 20,* 1304–1311.

Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1280–1293.

Tinker, M. A. (1958). Recent studies of eye movements in reading. *Psychological Bulletin, 55,* 215–231.

Vitu, F., Lancelin, D., & d'Unienville, V. M., (2007). A perceptual-economy account for the inverted-optimal viewing position effect. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1220–1249.

Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision Research, 41,* 3511–3531.

White, S. J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 205–223.

White, S. J., & Staub, A. (2012). The distribution of fixation durations during reading: Effects of stimulus quality. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 603–617.

Yang, S.-N., & McConkie, G. W. (2001). Eye movements during reading: A theory of saccade initiation times. *Vision Research, 41,* 3567–3585.

CHAPTER

19

E-Z Reader: An Overview of the Model and Two Recent Applications

Erik D. Reichle *and* Heather Sheridan

Abstract

This chapter reviews what is known about eye movements during reading and describes a computational model that simulates many of the perceptual, cognitive, and motor processes that guide readers' eye movements—the *E-Z Reader* model. The chapter discusses how the model is being used to examine two fundamental questions related to reading: (1) What mediates the development of reading skill? (2) What is the time course of lexical processing? Simulations using the model suggest that very rapid lexical processing is necessary for skilled reading and that this processing must be highly coordinated with other ongoing perceptual, cognitive, and motor processes. Thus a significant portion of the lexical processing of a word is completed while it is still in the parafovea (prior to the word being fixated). The implications of these conclusions are discussed, as are future directions in modeling the cognitive processes that control eye movements during reading.

Key Words: attention, computational model, development, eye movements, E-Z Reader, lexical processing, time course

> On the outside, the reader has rotated his eyes only a few millimeters…But on the inside, there has been a rapid succession of intricate events. Clearly, the succession could only be the product of a complex information processing system…It contains components that are asked to perform amazing feats with amazing rapidity, and precisely in concert.
>
> —Gough (1972, p. 341)

E-Z Reader is a computational model of eye movement control in reading. As such, it provides a formal description of how the perceptual and cognitive processes that are involved with reading interact with each other and the systems that program and execute saccades to produce the patterns of eye movements that are observed during reading (Pollatsek, Reichle, & Rayner, 2006; Rayner, Ashby, Pollatsek, & Reichle, 2004; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 1999, 2003; Reichle, Warren, & McConnell, 2009). In this capacity it has already proven highly successful, simulating the various benchmark phenomena that are related to readers' eye movements and that have been used to evaluate such models (for a review, see Reichle, 2011), and having been used as an analytical framework to examine a variety of theoretical issues related to reading (e.g., how eye movement control during reading may differ from eye movement control in other visual-cognitive tasks; Reichle, Pollatsek, & Rayner, 2012). Perhaps more important, however, is that the model is an existence proof showing how the serial lexical processing of words can be the engine moving the eyes forward during reading, thus allowing one to make sense of behavior that—because of its inherent complexity—would otherwise be difficult to interpret.

In the remainder of this chapter, we will first very briefly review what is known about the basic characteristics of eye movements during reading and explain why the endeavor of understanding and modeling readers' eye movements is a worthwhile enterprise. We will then describe the E-Z Reader model and how it has recently been used to understand two key areas of reading research—the developmental changes that occur as beginning readers become skilled adult readers, and the time course over which lexical processing occurs during natural reading of text. Finally, we will discuss a limitation of the E-Z Reader model and how future research might improve both it and other models of eye movement control during reading.

Eye Movements During Reading

The eye movements that occur during reading are largely composed of two basic components—the *saccades* or ballistic movements of the eyes from one viewing location to the next, and the *fixations* or intervals during which the eyes are relatively stationary. The majority of saccades move the eyes forward through the text, but approximately 10% to 15% are *regressions* that move the eyes back to previously fixated locations in the text. Because our eyes can only see fine detail in a small region of the retina called the *fovea*, readers must move their eyes from word to word, typically fixating 70% to 80% of words in a text at least once, and many words are fixated two or more times. Although fixations vary considerably in duration, ranging from 50 to 1,000 ms, most fixations are 200 to 250 ms in duration. Importantly, there is overwhelming evidence that a variety of variables related to both lexical and higher-level linguistic processing affect the durations and locations of fixations, thereby making the measurement of eye movements an ecologically valid method to examine the online cognitive processes that mediate text comprehension (for reviews, see Rayner, 1998; Schotter & Rayner, this volume).

Because eye movements reflect ongoing lexical and higher-order linguistic processing during reading, it is important to understand the precise manner in which both types of processing relate to visual processing, on the one hand, and oculomotor control, on the other. Efforts to understand this eye-mind link have resulted in a small number of computational models of eye-movement control during reading, of which E-Z Reader—the model that will be the focus of this chapter—is just one example (for a review, see Reichle et al., 2003). Although these models are often been described as models of eye-movement control rather than of reading per se (e.g., see Rayner & Reichle, 2010), they attempt to specify how several basic perceptual, cognitive, and motor processes dynamically interact across time to generate the moment-to-moment patterns of eye movements that are observed during reading. For that reason, the models provide theoretical frameworks for thinking about how the patterns of eye movements that are observed during reading are generated by the various components that support reading comprehension. Two examples illustrating this claim will be provided later in this chapter, but first we will provide a detailed description of the E-Z Reader model.

The E-Z Reader Model

The model has two core assumptions. The first is that lexical processing is completed in a strictly serial manner, on one word at any given time. Within the framework of the model, this assumption effectively means that the type of attention that is required for lexical processing (e.g., by binding together the features that make up a word) is allocated in a strictly serial manner, to exactly one word at a time. Because of this assumption, the model is an instance of a more general class of *serial-attention models* (see Reichle, 2011), which can be contrasted to models in which attention is allocated as a gradient to support the concurrent processing of multiple words (e.g., Glenmore: Reilly & Radach, 2006; SWIFT: Engbert, Nuthman, Richter, & Kliegl, 2005).

The second core assumption of the E-Z Reader model is that the completion of a preliminary stage of lexical processing called the *familiarity check* on a word normally initiates the programming of a saccade to move the eyes to the next word. This assumption can be conceptualized as a heuristic that skilled readers acquire to afford maximal reading efficiency. That is, by initiating saccadic programming prior to the completion of lexical access, the eyes will not remain fixated on a word during the time that is required to program a saccade, thereby increasing fixation durations unnecessarily (e.g., for a discussion of this heuristic, see Reichle & Laurent, 2006). However, the fact that the subsequent completion of lexical access on a word then causes attention to shift to the next word means that there is a decoupling between the movements of overt and covert attention. As will be explained, this decoupling allows the model to explain a certain amount of slippage that seems to occur between where the eyes are located and what the mind is processing—as evidenced, for example, by the finding that words are fixated for shorter durations if they are previewed in the parafovea prior to being fixated (Rayner, 1975; Reingold, Reichle, Glaholt, & Sheridan, 2012; for a review, see Schotter, Angele, & Rayner, 2012).

The two core model assumptions, in conjunction with several more specific assumptions about how the various processes involved in reading interact with each other to move readers' eyes through text, form the framework of the model that is schematically illustrated in Figure 19.1A. That figure shows the various components of the model (represented by the gray boxes) and how both information and the control of processing are propagated through those components (represented by the arrows). These assumptions will now be explained in detail.

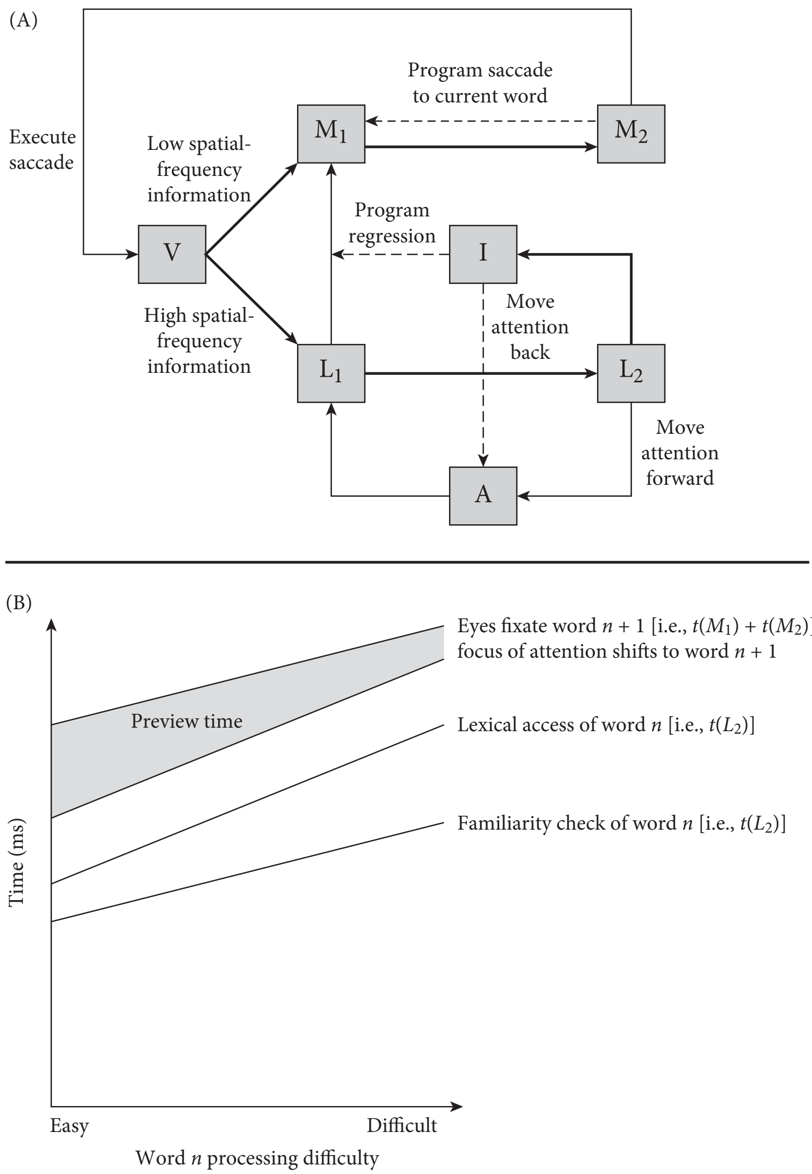


Fig. 19.1 Panel A is a schematic diagram of the E-Z Reader model of eye movement control during reading. The labeled components are: *V* preattentive stage of visual processing; L_1 familiarity check; L_2 lexical access; *A* attention shift; *I* postlexical integration; M_1 labile stage of saccadic programming; and M_2 nonlabile stage of saccadic programming. The thick arrows represent the flow of information, the thin solid arrows show how the control of processing is obligatorily passed between model components, and the thin dashed arrows show how the control of processing can be probabilistically passed between model components. Panel B is a schematic diagram showing how parafoveal processing of word *n* +1 is modulated by the processing difficulty of word *n* (i.e., the fixated word). The *x*-axis shows the relative processing difficulty of word *n*, and the *y*-axis shows the mean time courses of key processes in the model, and how the slippage between when attention shifts to word *n* +1 and when the eyes move to word *n* +1 gives affords some amount of preview of that word. (The eye-mind lag would be expected to lengthen the preview time, but this is ignored in the figure for convenience and because its duration is a constant.)

In the model, 50 ms is required for this propagation of information from the eyes to the mind, based on estimates of the eye-mind lag (for a review, see Reichle & Reingold, 2013). This stage of visual processing (labeled *V* in the model) is assumed to be preattentive because information is acquired from across the entire visual field, independent of where the focus of attention is located. But as Figure 19.1A shows, the two types of information are used for different purposes. Whereas the low spatial-frequency information (which is available in peripheral vision, where visual acuity is reduced) is used to segment upcoming words for the purposes of saccadic targeting, a smaller portion of the high spatial-frequency information (which is only available in central vision) is selected via attention for the purposes of lexical processing.

As already indicated, lexical processing is completed in two successive stages—the familiarity check (labeled L_1 in the model) and lexical access (labeled L_2). This distinction was originally motivated by dual-process theories of memory, in which the recognition of an item (e.g., word) can be based on two sources of information—a rapidly available sense of familiarity and a slower retrieval of information representing the item and the context in which it was encoded (e.g., see Yonelinas, 2002). Alternatively, L_1 and L_2 may be conceptualized as respectively corresponding to orthographic and semantic processing (Reingold & Rayner, 2006; Reingold, Yang, & Rayner, 2010). These two accounts are not mutually exclusive, so that, for example, word familiarity may be largely based on orthographic information.

In the model, the time (in ms) to complete the first stage of lexical processing for word *n,* $t(L_1)$, is given by Equation 1:

$$(1)\ t(L_1) = \begin{cases} 0 & \text{if } p < predictability_n \\ \alpha_1 - \alpha_2 \ln(frequency_n) - \alpha_3\, predictability_n & \text{if } p \geq predictability_n \end{cases}$$

In Equation 1, the top branch represents instances in which a word is guessed from its preceding sentence context, allowing the familiarity check to be completed in 0 ms. This happens with a probability, *p*, equal to a word's cloze predictability, which is the mean proportion of time that the word is guessed from its preceding sentence context by a group of independent subjects (Taylor, 1953). The assumption that words can be guessed in this manner was motivated by the finding that in eye-movement experiments during which only the word being fixated is visible (e.g., the letters in the nonfixated words are replaced by random letters), readers sometimes completely skip highly predictable words (e.g., high-frequency function words like the article *the*; Rayner, Well, Pollatsek, & Bertera, 1982). Such words are presumably skipped because the semantic or syntactic constraints on the words are sufficient for them to be identified using only minimal visual information about the word (e.g., its length).

However, except for these predictable words, the time required to complete the familiarity check is assumed to be a linear function of the logarithm of word *n*'s frequency of occurrence in printed text (as tabulated in various text corpora; e.g., Francis & Kučera, 1982) and its cloze predictability, as modulated by three free parameters shown in the bottom branch of Equation 1: $\alpha_1 = 104$, $\alpha_2 = 3.4$, and $\alpha_3 = 39$. (These and other parameter values were selected to optimize the model's goodness-of-fit to empirical data.) Thus, on average, the familiarity check will require less time to complete for frequent or predictable words.

Because there is considerable inherent variability in the time required to process a word, the time that is specified by Equation 1 is only the mean time that is required to complete the familiarity check for a word of a given frequency and predictability; the actual time to complete the familiarity check on such a word during any given Monte-Carlo simulation run of the model is a random deviate that is sampled from a gamma distribution. (The times required to complete several of the processes in the model are random deviates that are sampled from gamma distributions having a specified mean and a standard deviation equal to $\sigma\gamma = 0.22$ of the mean.) The time that is required to complete the familiarity check is then adjusted as a function of the mean *eccentricity* (i.e., the distance in character spaces) between the point of fixation and each of the letters of the word being processed, as specified by Equation 2:

$$(2)\quad t(L_1) \leftarrow t(L_1) \bullet \varepsilon^{\sum_{i=1}^{N} |fixation - letter_i| / N}$$

In Equation 2, the free parameter $\varepsilon = 1.15$ determines the absolute amount by which eccentricity modulates the slowing effect of limited visual acuity, with *i* in the exponent indexing each of the *N* letters in the attended words. Thus, according to Equations 1 and 2, with all else being equal, words that are frequent, predictable, short, or close to fixation will be the recipients of fewer, shorter fixations than words that are infrequent, unpredictable, long, or far from fixation, consistent with what is typically observed (Rayner, 1998; Schotter & Rayner, this volume).

Turning now to lexical access, the time (in ms) required to complete lexical access on word *n*, $t(L_2)$, is a fixed proportion ($\Delta = 0.34$) of the time required to complete the familiarity check, as specified by Equation 3:

$$(3)\quad t(L_2) = \Delta\left[\begin{array}{l}\alpha_1 - \alpha_2 \ln(frequency_n) \\ -\alpha_3 predictability_n\end{array}\right]$$

In contrast to the familiarity check, lexical access always requires some nonzero amount of time to complete based on the assumption that it involves the activation of a word's meaning, irrespective of whether that meaning is activated from its prior sentence context, from visual input, or some combination of the two. As was true of the familiarity check, the actual time that is required to complete lexical access during any Monte-Carlo simulation run of the model is a random deviate that is sampled from a gamma distribution.

As Figure 19.1A shows, the completion of lexical access simultaneously causes two things to happen. The first is that attention shifts from the word that was just identified to the next word, so that lexical processing (i.e., the familiarity check) of the next word can begin. The shifting of attention is not instantaneous, however; the time required to shift attention, $t(A)$, is a random deviate sampled from a gamma distribution having a mean determined by the free parameter $A = 25$ ms.

The second thing that happens when a word is identified is that postlexical *integration* of that word's meaning begins. This integration (labeled *I* in the model) is the minimal time required for the reader to know that the meaning of the identified word fits into the semantic and syntactic framework of the sentence representation that is being constructed. Because postlexical processing of a word is normally completed in the background of ongoing lexical processing, thus having no discernable affect on the progression of the eyes through the text, and because postlexical processing is not important for the two model applications to be discussed, it will not be described in detail here. However, it is important to note that integration failure can cause the eyes and attention to move back to the location of integration failure (see Reichle et al., 2009), allowing the model to simulate the regressions observed with sentences that are syntactically ambiguous (e.g., Frazier & Rayner, 1982) or semantically implausible (Warren & McConnell, 2007) that are discussed more fully in Staub (this volume).

The remaining assumptions of the E-Z Reader model are all related to saccadic programming and execution. The first of these assumptions is that saccadic programming is completed in two successive stages—a labile stage (labeled M_1 in the model) that is subject to cancellation by the initiation of a subsequent saccadic program, following by a nonlabile stage (labeled M_2) that is not subject to cancellation. The motivation for this assumption was based on seminal experiments in which subjects were instructed to move their eyes as rapidly as possible from one cued location to another. These experiments which involved simple stimuli but complex situations in which saccades should be made or suppressed (e.g., Becker & Jürgens, 1979) showed that saccades are programmed in two successive stages.

In E-Z Reader, the times required to complete both the labile and nonlabile stages of saccadic programming are random deviates sampled from gamma distributions having means of $t(M_1) = 125$ ms and $t(M_2) = 25$ ms, respectively. This allows the model to explain word skipping. To understand how, imagine that the eyes and attention are on word *n*. At some point, the familiarity check on that word will complete, causing the initiation of a saccade program to move the eyes to word *n* + 1. Lexical processing of word *n* will continue, however, until word *n* has been identified (i.e., the completion of L_2), causing attention to shift to word *n* + 1 so that parafoveal processing of that word can begin. At this point, two things can happen. The first is that the labile saccadic program to move the eyes to word *n* + 1 completes before the familiarity check on word *n* + 1 completes; in this situation, the saccadic program has reached a point of no return and upon completion of the nonlabile stage the eyes will obligatorily be directed toward word *n* + 1. The second possible situation is that the familiarity check on word *n* + 1 completes before the labile saccadic program to move the eyes to word *n* + 1; in the second situation, a second labile saccadic program (to move the eyes to word *n* + 2) will be initiated, thereby canceling the first and resulting in the eyes eventually being directed to word *n* + 2 and causing word *n* + 1 to be skipped. This account of word skipping via the replacement of one saccadic program by another gives rise to the prediction of skipping cost, or inflated fixations immediately before skipped words—a prediction that has been partially confirmed (e.g., Kliegl & Engbert, 2005; see also Reichle & Drieghe, 2013).

The model also assumes that saccades are always directed toward the centers of words, but that the length of any given saccade will be a linear combination of three components, as indicated in Equation 4:

$$(4) \quad \textit{saccade length} = \textit{intended saccade length} + \textit{systematic error} + \textit{random error}$$

The intended saccade length is the actual distance (in character spaces) between the current fixation location and the saccade target (which is the center of whatever word the eyes are being directed toward). As Equation 5 shows, the systematic error (in character spaces) is a function of the disparity between the intended saccade length and an optimal saccade length (Ψ = 7), and the fixation duration on the launch-site word, $\textit{fixation}_{LS}$. Thus saccades that are longer/shorter than seven character spaces will tend to undershoot/overshoot their intended targets by approximately half a character space of deviation, with the amount also modulated by the fixation duration on the launch site. (Ω_1 = 6 and Ω_2 = 3 are free parameters that control the degree to which the launch-site fixation duration modulates the systematic error.

$$(5) \quad \textit{systematic error} = (\psi - \textit{intended saccade length}) \bullet \left\{\left[\Omega_1 \ln\left(\textit{fixation}_{LS}\right)\right]/\Omega_2\right\}$$

And finally, the random error component in Equation 4 is a random deviate that is sampled from a Gaussian distribution with μ = 0 character spaces and σ specified by Equation 6. In that equation, the free parameters η_1 = 0.5 and η_2 = 0.15 control the degree to which the variability of the random error component increases with the intended saccade length, so that long saccades are more prone to error than short saccades.

$$(6) \quad \sigma = \eta_1 + \eta_2 \bullet \textit{intended saccade length}$$

In combination, Equations 4–6 cause the distributions of fixation landing sites to resemble those reported in the literature—the distributions are approximately Gaussian in shape, centered near the middles of words but with missing tails that reflect instances when a saccade undershot/overshot its intended target and that increase in magnitude as the launch-site fixation duration decreases (e.g., McConkie, Kerr, Reddix, & Zola, 1988).

For the sake of simplicity, the model assumes that saccades require a constant S = 25 ms to execute. Although visual processing halts during the actual saccades (Matin, 1974), lexical processing continues at a rate determined by the intrinsic properties of the word being processed (i.e., its frequency and predictability; see Equations 1 and 3) and the eccentricity of saccade launch-site location (see Equation 2). Lexical processing then continues at its presaccade rate for an additional V = 50 ms after the eyes fixate their new viewing location (i.e., the duration of the eye-mind lag in the model). Because the time required to complete lexical access is some fixed proportion of the time required to complete the familiarity check (see Equation 3), and because the times required to complete saccadic programming, execute the saccade, shift attention, and propagate visual information from the eyes to the mind are (on average) constants, there is often a considerable amount of time available for parafoveal processing of word n + 1 from a fixation on word n, but this time varies as a function of the processing difficulty of word n. The duration of the processes determines the time that is available for parafoveal preview, as shown in Figure 19.1B. In this figure, the processing difficulty of word n is indicated along the x-axis, the process durations are indicated along the y-axis, and the amount of time available for previewing word n + 1 from word n is indicated by the gray shading. As can be seen, preview of word n + 1 is modulated by word n's processing difficulty. This allows the model to explain the finding that, as the processing difficulty of word n increases, the time available for parafoveal processing of word n + 1 decreases (e.g., Henderson & Ferreira, 1990).

The final assumption of E-Z Reader is related to automatic refixations, or rapid eye movements to a new viewing location following an initial fixation near the edge of a word. The motivation for this assumption is that an initial fixation near the beginning and ending of a word affords a poor viewing location from which to process the word, and as such might be expected to result in a rapid movement of the eyes toward the center of the word, a location that affords more rapid and accurate lexical processing (O'Regan & Lévy-Schoen, 1987). According to the model, this propensity is based on efference copies of the saccadic programs that aim to move the eyes from one word to the middle of the next (Carpenter, 2000); to the extent that the intended saccade is prone to error and deviates from its target (see Equations 4–6), the probability of rapidly initiating a second, corrective saccade increases, as specified by Equation 7. The probability of initiating a corrective saccade (i.e., refixating) increases with the absolute distance (in character spaces) between the initial fixation position and the original saccade target (i.e., the center of the word being targeted), but is modulated by the free parameter λ = 0.16.

$$(7)\quad p(\textit{refixation}) = \max\left(\lambda\left|\textit{landing position} - \textit{saccade target}\right|, 1\right)$$

The model as described is able to simulate all of the benchmark findings that have been used to evaluate models of eye movement control in reading (see Reichle et al., 2012). To give a specific example, Figure 19.2 shows the mean values of six commonly used word-based dependent measures for five frequency classes of words in a corpus of sentences used by Schilling, Rayner, and Chumbley (1998): (1) *first-fixation duration*, or duration of the first of one or more fixations on a word; (2) *single-fixation duration*, or duration of a fixation on a word that is fixated exactly once; (3) *gaze duration*, or the sum of all first-pass fixations on a word; (4) the probability of fixating a word exactly once; (5) the probability of fixating a word two or more times; and (6) the probability of skipping a word. All of these measures are *first-pass* measures. That is, they are calculated using only fixations that occurred during the first pass through the sentences excluding any fixations that occurred after interword regressions. As Figure 19.2 indicates, as a word's frequency of occurrence in printed text increases, the mean fixation duration measures on those words decrease, as do the mean probabilities of the words being fixated once as opposed to more than once.

The finding that both the propensity to fixate a word and the durations of those fixations are modulated by the word's frequency is extremely robust (Inhoff & Rayner, 1986; Just & Carpenter, 1980; Rayner et al., 2004; Schilling et al., 1998) and provides compelling evidence that the decisions about when to move the eyes from a word are sensitive to the local processing difficulty of that word (i.e., its frequency). As Figure 19.2 shows, the E-Z Reader model does a fairly good job of simulating the effects of word frequency in reading, thereby demonstrating how something that may seem as slow as lexical processing in a serial word-by-word model can nonetheless be the engine that controls the progression of the eyes through text.

Two Recent Model Applications

As already indicated, the E-Z Reader model has been used to examine a large number of theoretical issues related to reading (see Reichle, 2011). In this section, we will review two areas of research that have recently been examined using the model—the question of how reading skill develops and an attempt to better understand the time course over which lexical processing occurs.

The Development of Reading Skill

A number of studies have examined the eye movements of beginning readers—8 to 10-year-old children with two to four years of formal reading education and who are proficient at decoding words and who can silently read complete sentences, but at slower rates than adults, even when reading age-appropriate texts. The key results of these comparative studies are remarkably consistent: Relative to skilled adult readers, children typically read fewer words per minute, making more fixations that are longer in duration, shorter saccades, with a larger proportion of those saccades being regressions (for a review, see Blythe & Joseph, 2011).

There have also been other documented differences between the eye movements of child versus adult readers. For example, relative to adults, children have a smaller *perceptual span*, or region of effective vision, being less able than adults to use parafoveal vision to identify letters, the features of letters, and the blank spaces between words (e.g., Häikiö, Bertram, Hyönä, & Niemi, 2009). Children's fixation durations are also modulated by word frequency (e.g., Blythe, Liversedge, Joseph, White, & Rayner, 2009) and word length (e.g., Joseph, Liversedge, Blythe, White, & Rayner, 2009) to a greater degree than are adults'. And children are slower at detecting violations of semantic plausibility (e.g., *Robert used a hook to catch the horrible mouse*, where *mouse* is implausible) than adults, typically detecting such violations only after their eyes have moved from the implausible word (Joseph et al., 2008). Interestingly, however, children's fixation landing-site distributions on words are very similar to those of adults, suggesting that even beginning readers are targeting their saccades on a word in a manner similar to that of skilled readers (Joseph et al., 2008).

Two general accounts have been proposed to explain the observed differences between eye movements of children versus adult readers. According to the *oculomotor-tuning hypothesis*, these differences reflect the fact that children are less skilled at moving

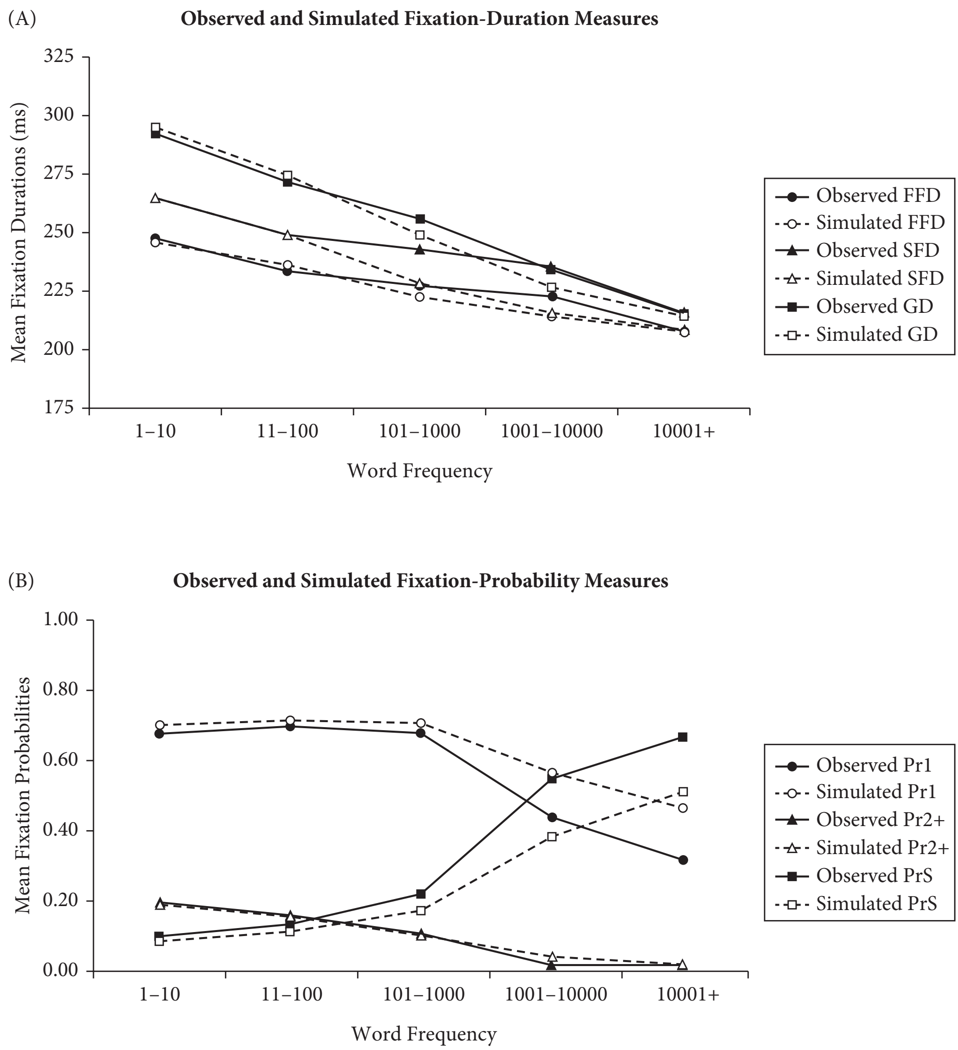


Fig. 19.2 Panel A shows the mean observed (solid lines) and simulated (dashed lines) first-fixation (*FFD*), single-fixation (*SFD*), and gaze durations (*GD*), for five frequency classes of words. Panel B shows the mean observed (solid lines) and simulated (dashed lines) probabilities of fixating once (*Pr1*), two or more times (*Pr2+*), and skipping (*PrS*) for five frequency classes of words.

their eyes, possibly because they are slower at programming saccades or more prone to saccadic error (e.g., see Klein & Foerster, 2001). According to the alternative, *linguistic-proficiency hypothesis*, these differences reflect the fact that children are simply less proficient than adults at identifying printed words and integrating their meanings into linguistic representations (e.g., see Perfetti, 2007).

To evaluate the plausibility of the oculomotor-tuning and linguistic-proficiency hypotheses, a series of simulations were completed using the E-Z Reader model. In these simulations, the parameters that modulate the rate and manner of both saccadic programming and execution, on one hand, and lexical and postlexical processing, on the other, were systematically manipulated (Reichle et al., 2013). The goal in doing this was to first determine which parameters could be adjusted to produce the global pattern of eye movements observed with children (i.e., slower reading rates, longer fixations), and to then determine whether the adjustments would be sufficient to account for the remaining similarities (e.g., similar fixation landing-site distributions) and differences (e.g., slower detection of semantic plausibility violations) that have been observed between children and adults.

The results of these simulations were straightforward but surprising—only increasing the value of the parameter that controls the overall rate of lexical processing (i.e., from α_1 = 104 to α_1 = 208; see Equation 1) was sufficient to generate the all of the findings related to children's eye movements except one. The one exception was the finding that children are slower than adults at detecting semantic-plausibility violations, often detecting such violations very late, as evidenced by the fact that their first-pass measures (e.g., gaze durations) are unaffected by these violations but their second-pass measures (e.g., total viewing times) are (Joseph et al., 2008). To account for this final result, it was necessary to assume that the children also require more time to complete postlexical integration than adults. This assumption, in combination with the slower overall rate of lexical processing for children than adults, was sufficient for the model to simulate the finding that children's first-pass fixation-duration measures were unaffected by semantic plausibility violations, but that their second-pass measures were.

Thus on the basis of these simulations, one might conclude that the primary reason for the differences between the eye movements of beginning and skilled readers are differences in their proficiency in lexical—and to perhaps some lesser degree—postlexical processing. Although this conclusion is obviously tentative, it is interesting because it is consistent with prior claims that variation in the speed and accuracy of lexical processing is what mediates between-individual differences in reading skill (e.g., Ashby, Rayner, & Clifton, 2005; Perfetti, 2007; Shilling et al., 1998). That being said, we will now discuss how E-Z Reader has been used to examine the time course of lexical processing during reading.

The Time Course of Lexical Processing

Given the strong assumption that the rate of lexical processing accounts for both within- (i.e., developmental) and between-individual differences in reading skill, one might ask about the time course of lexical processing and how it might vary both within and between individuals. This question has been the subject of considerable empirical research during the past several decades (e.g., see Reichle, Tokowicz, Liu, & Perfetti, 2011; Reingold et al., 2012; Schilling et al., 1998), in no small part because the estimates of the time required have often varied quite considerably across tasks. For example, one method of estimating the speed of lexical processing is to record reaction times on behavioral tasks, such as naming and lexical decision tasks. These tasks produce reaction times of approximately 500 to 700 ms (e.g., Schilling et al., 1998), but it is important to note that this time also encompasses nonlexical processes, including the motor and decision processes that support lexical decisions, and articulatory processes required for naming. Likewise, fixation times during reading are not a pure measure of the speed of lexical processing, because the dependent measures like gaze duration can also reflect additional processes, such as postlexical integration.

In an effort to provide more precise estimates of the speed of lexical processing during reading, recent research has employed distributional analyses, such as ex-Gaussian fitting (Staub, White, Drieghe, Hollway, & Rayner, 2010) and survival analyses (Reingold et al., 2012) to demonstrate that lexical variables (e.g., word frequency) can produce rapid effects on fixation durations during reading (for a review of these findings, see Reingold, Sheridan, & Reichle, this volume). Such rapid lexical effects are consistent with other work that employed neuroimaging methodologies to demonstrate lexical effects within the range of 110 to 170 ms post stimulus onset (e.g., Assadollahi & Pulvermüller, 2001, 2003; Hauk, Davis, Ford, Pulvermüller, & Marslen-Wilson, 2006; Penolazzi, Hauk, & Pulvermüller, 2007; Reichle et al., 2011; Sereno, Brewer, & O'Donnell, 2003; Sereno, Rayner, & Posner, 1998; for a review, see Reichle & Reingold, 2013).

Although estimates of lexical processing time vary greatly, it is clear that a minimum interval of 100 to 150 ms is required before lexical processing is advanced enough to have a potential impact on fixation durations (Reichle & Reingold, 2013). However, because fixations are only 200 to 250 ms in duration, an important point of controversy has been whether lexical processing is fast enough to be the engine that drives eye movements (see Reingold et al., this volume). It is not immediately obvious how this could be true, given that lexical influences are subject to severe temporal constraints—they must occur after the 50 ms eye-mind lag that occurs at the start of the fixation and before the 100 to 150 ms required to program a saccade that occurs at the end of a fixation. Because of these temporal constraints, it has been historically argued that word identification is simply too slow to have an impact on eye movements (Bouma & de Voogd, 1974; Kolers, 1976). As a result, a few of the current models of eye movement control continue to assume that lexical processing plays only a minimal role in controlling eye movements (e.g., Feng, 2006; Yang, 2006).

Thus, given that the E-Z Reader model assumes that an early stage of lexical processing (i.e., L_1) is the engine that drives eye movements, it is important to demonstrate precisely how, according to the model, lexical processing can be rapid enough to have an impact on fixation durations. For example, the E-Z Reader model assumes that readers normally spend a substantial amount of time processing words to the right of fixation, thereby affording a significant amount of parafoveal processing (see Figure 19.1B). Specifically, the model predicts that, in many instances, the completion of lexical access (i.e., L_2) of word *n* allows attention to shift to word *n* + 1 before the eyes actually move to word *n* + 1, thereby making some amount of time available for the parafoveal processing of word *n* + 1 from word *n*. The duration of this preview time includes whatever time is available between when attention first shifts to word *n* + 1 and when new visual information from the fixation on word *n* + 1 actually reaches the brain. This preview-time interval (i.e., the interval between when attention first shifts to word *n* + 1 and the completion of the eye-mind lag from the new fixation on word *n* + 1) could potentially provide a substantial amount of time for readers to initiate lexical processing of word *n* + 1.

To examine the E-Z Reader model's predictions about parafoveal processing, Schotter, Reichle, and Rayner (2014) recently completed simulations to determine whether the duration of preview time predicted by the model is sufficiently long to explain two interesting but controversial phenomena: semantic-preview benefits and word *n* + 2 preview effects. The former controversy is about whether readers can obtain semantic information about word *n* + 1 while fixating word *n* (i.e., *semantic preview effects*) or whether it is instead only possible to obtain orthographic or phonological information. Although a number of studies have failed to show semantic-preview effects in English (Rayner, Balota, & Pollatsek, 1986; Rayner & Schotter, 2014; Rayner, Schotter, & Drieghe, 2014), a recent study by Schotter (2014) successfully demonstrated semantic-preview effects. This study used a gaze-contingent display-change method called the *boundary paradigm* (Rayner, 1975) to manipulate the letter information in the location of a target word prior to it being fixated. For example, prior to fixating the target word (e.g., *begin*), the reader might receive a preview that was identical to the target word (e.g., *begin*), a synonym of the target word (e.g., *start*), a semantically related word (e.g., *ready*), or an unrelated word (e.g., *check*). When the reader's eyes then crossed an invisible boundary to the left of a target word, the preview was immediately replaced by the target word. Using this paradigm, Schotter demonstrated that fixation durations on the target word were approximately the same for the identical- and synonym-preview conditions and that both of these conditions produced faster fixation times than the unrelated-preview condition. This pattern of results therefore suggests that semantic information can be extracted from the parafovea, thereby allowing the meaning of word synonyms to be somehow integrated.

The second controversy mentioned earlier refers to the debate about whether parafoveal processing from word *n* can extend as far as word *n* + 2 (i.e., *word n + 2 preview effects*), or whether it is instead only possible to obtain information about word *n* + 1. Although a number of studies have failed to show word *n* + 2 preview effects (Angele & Rayner, 2011; Angele, Slattery, Yang, Kliegl, & Rayner, 2008; Rayner, Juhasz, & Brown, 2007), these effects have been demonstrated under some circumstances, such as when word *n* + 1 is short and high in frequency (e.g., Kliegl, Risse, & Laubrock, 2007; McDonald, 2006; Radach, Inhoff, Glover, & Vorstius, 2013).

Although both semantic-preview effects and word *n* + 2 effects might intuitively seem to be at odds with a model such as the E-Z Reader model (because of its strong assumption that attention is only allocated to one word at a time), the simulations reported by Schotter et al. (2014) demonstrated that neither effect is necessarily inconsistent with the model. These simulations used the standard version of the model and its default parameter values (see Reichle et al., 2012) to examine the model's predictions about the time spent engaged in the lexical processing of parafoveal words.

To examine semantic-preview effects using the model, Schotter et al.'s (2014) first simulation used the mean lengths, frequencies, and predictabilities of both the pretarget words and the synonyms of the target words that were used in Schotter's (2014) experiment. The results of this simulation were informative: The mean probability of previewing word *n* + 1 (i.e., the word after the target word) was 0.94, the mean duration of that preview was 177 ms, and the mean probability of the word *n* + 1 preview advancing to the L_2 stage of lexical processing was 0.08. Because the L_2 stage is hypothesized to encompass semantic processing, this last simulation result suggests that the model predicts some amount of semantic preview on a modest but nontrivial proportion of trials—consistent with the results reported by Schotter (2014).

To examine word $n + 2$ preview effects, Schotter et al.'s (2014) second simulation examined the probability and time spent previewing word $n + 2$ from word n while varying the lengths, frequencies, and predictabilities of words n and $n + 1$ across a range of values. The key results of this second simulation were that the E-Z Reader model predicted some amount of parafoveal processing of word $n + 2$ on 20% of simulation trials, but this processing never advanced to the L_2 (i.e., semantic) stage. This pattern of results suggests that word $n + 2$ preview effects should be limited to orthographic (and perhaps some amount of phonological) processing, but not semantic processing. The results of this second simulation in combination with the first therefore suggest that the E-Z Reader model can accommodate modest-sized semantic-preview and word $n + 2$ preview effects, and the reason for this is that the model's assumptions afford a sufficient amount of time—but not too much time—for parafoveal processing of upcoming words.

Given the important role played by parafoveal processing in the E-Z Reader model, we were interested in knowing whether the model's predictions about preview time would be congruent with the empirical estimates reported by Reingold et al. (2012). In that study, the frequency (i.e., high vs. low) and preview availability (i.e., available vs. not available) of target words was manipulated using the boundary paradigm (Rayner, 1975), such that readers either saw a preview of the target word (i.e., the valid preview condition) or a pronounceable nonword (i.e., the invalid preview condition). A survival-analysis technique was then used to provide estimates of the earliest influence of the word-frequency manipulation on fixation times in the valid and invalid conditions. The key finding was that the earliest influence of word frequency occurred (on average) 145 ms after the start of fixation on the target words in the valid preview condition, but occurred 256 ms after the start of the fixation on the target words in the invalid preview condition. This suggests that preventing the parafoveal preview of the target words slowed their lexical processing by approximately 111 ms (i.e., 256 – 145 = 111 ms). Therefore, based on these empirical estimates, it was important to know whether a simulation using the stimuli from the Reingold et al. (2012) experiment might produce equally long preview times, which could explain why lexical processing was dramatically faster in the valid than invalid preview condition.

Our simulation was completed using the lengths, frequencies, and predictabilities of the pretarget and target words used by Reingold et al. (2012), using the 48 sentences of the Schilling et al. (1998) corpus as frames for these words. (The target words were always located at the sixth word position in the sentence frames). Because we were also interested in knowing how preview time might be modulated by hypothesized differences in reading skill, we completed the simulations using two rates of lexical processing. This was done using the same two values of the α_1 parameter (see Equation 1) that were used to simulate beginning and skilled readers (i.e., $\alpha_1 = 104$ ms for adults vs. $\alpha_1 = 208$ ms for children; see Reichle et al., 2013). This new simulation otherwise used all of the model's default parameter values (see Reichle et al., 2012) and 1,000 virtual participants per simulated condition.

Our simulation yielded mean preview times of 158 ms in the skilled reading condition (i.e., $\alpha_1 = 104$ ms) and 125 ms in the less-skilled reading condition ($\alpha_1 = 208$ ms). Both of these predicted values are similar to the mean preview times obtained in the simulations reported by Schotter et al. (2014). More importantly, our simulation results suggest an important link between the rate of lexical processing and the amount of time that is available for preview, such that a slower rate of lexical processing affords less time for parafoveal processing of upcoming words. In other words, if word n requires more time to process because of a slower rate of lexical processing (or alternatively, because the word is low frequency; Henderson & Ferreira, 1990), then there is necessarily a shorter interval of time available for the parafoveal processing of word $n + 1$ (see Figure 19.1B). Furthermore, our simulated manipulation of reading skill also markedly affected the time available for preview, reducing it by 33 ms (i.e., 158 – 125 ms) in the less-skilled condition. Finally, if one subtracts the duration of the eye-mind lag (i.e., 50 ms) from the 158 ms preview time in the skilled-reading condition, the resulting estimate of preview time (i.e., 108 ms) is very consistent with Reingold et al.'s (2012) estimate of 111 ms based on survival analyses of fixation times.

Because the E-Z Reader model was not explicitly designed to produce preview times of a particular duration, our simulation results are important because they show that simulated preview times in excess of 100 ms are a nonintuitive by product of the assumptions of the model. It is also impressive that the model's predictions about preview time can potentially accommodate a wide range of findings that the model was not originally designed to explain, such as semantic-preview effects, word $n + 2$

preview effects, and the results of survival analyses of fixation times. Moreover, our simulation underscores the importance of actually running simulations to test one's predictions, rather than simply assuming that a model can or cannot account for nonintuitive findings (Rayner, Pollatsek, & Reichle, 2003). In this vein, we believe that future efforts should examine the model's intriguing prediction that differences in reading skill (as indexed by differences in the rate of lexical processing) can influence the amount of time that is available for parafoveal processing during reading.

Conclusions

During the last decade, the E-Z Reader model has motivated a large amount of new empirical research (e.g., Inhoff, Eiter, & Radach, 2005; Kennedy, 2008; Mitchell, Shen, Green, & Hodgson, 2008; Reichle et al., 2011; Reingold & Rayner, 2006; Reingold et al., 2010; Staub, 2011; White, Warren, & Reichle, 2011). We believe that this is largely due to the fact that the model provides a simple theoretical framework for thinking about eye movement control in reading—a framework that is predicated on the basic assumption that words are (normally) identified one at a time and that the decisions about when to move the eyes are linked to an early stage of word identification. That being said, it is also important to acknowledge that the model fails to provide any deep account of the many component processes that are involved in guiding readers' eye movements (e.g., attention, lexical processing; for discussions of this, see Rayner et al., 2003; Reichle et al., 2009). We therefore also believe that future models of eye movement control will have to become more specific in their assumptions about how the various components involved in moving the eyes during reading (e.g., attention, lexical processing) are instantiated, perhaps by incorporating more detailed models of those processes within their frameworks.

Acknowledgments

The writing of this chapter was supported by a postdoctoral fellowship (PDF) awarded to H.S. from the Natural Sciences and Engineering Research Council of Canada (NSERC) and support from the Center for Vision and Cognition, University of Southampton.

References

Angele, B., & Rayner, K. (2011). Parafoveal processing of word $n + 2$ during reading: Do the preceding words matter? *Journal of Experimental Psychology: Human Perception and Performance*, *37,* 1210.

Angele, B., Slattery, T. J., Yang, J., Kliegl, R., & Rayner, K. (2008). Parafoveal processing in reading: Manipulating $n + 1$ and $n + 2$ previews simultaneously. *Visual Cognition*, *16,* 697–707.

Ashby, J., Rayner, K., & Clifton, C., Jr. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *Quarterly Journal of Experimental Psychology*, *58A,* 1065–1086.

Assadollahi, R., & Pulvermüller, F. (2001). Neuromagnetic evidence for early access to cognitive representations. *Neuroreport*, *12,* 207–213.

Assadollahi, R., & Pulvermüller, F. (2003). Early influences of word length and frequency: A group study using MEG. *Neuroreport*, *14,* 1183–1187.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research*, *19,* 967–983.

Blythe, H. I., & Joseph, H. S. S. L. (2011). Children's eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 643–662). Oxford, England: Oxford University Press.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research*, *49,* 1583–1591.

Bouma, H., & de Voogd, A. (1974). On the control of eye saccades in reading. *Vision Research*, *14,* 273–284.

Carpenter, R. H. S. (2000). The neural control of looking. *Current Biology*, *10,* R291–R293.

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112,* 777–813.

Feng, G. (2006). Eye movements as time-series random variables: A stochastic model of eye movement control in reading. *Cognitive Systems Research*, *7,* 70–95.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston, MA: Houghton Mifflin.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14,* 178–210.

Gough, P. B. (1972). One second of reading. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Reading by ear and eye* (pp. 331–358). Cambridge, MA: MIT Press.

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology*, *102,* 167–181.

Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage*, *30,* 1383–1400.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16,* 417–429.

Inhoff, A. W., Eiter, B. M., & Radach, R. (2005). Time course of linguistic information extraction from consecutive words during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *31,* 979–995.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, *40,* 431–439.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., Gathercole, S. E., & Rayner, K. (2008). Children's and adults' processing of anomaly and implausibility during reading: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, *61,* 708–723.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, *49,* 2078–2086.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87,* 329–354.

Kennedy, A. (2008). Parafoveal-on-foveal effects are not an artifact of mislocated saccades. *Journal of Eye Movement Research*, *2,* 1–10.

Klein, C., & Foerster, F. (2001). Development of prosaccade and antisaccade task performance in participants ages 6 to 26 years. *Psychophysiology*, *38,* 179–189.

Kliegl, R., & Engbert, R. (2005). Fixation duration before word skipping in reading. *Psychonomic Bulletin & Review*, *12,* 132–138.

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word *n* + 2. *Journal of Experimental Psychology: Human Perception and Performance*, *33,* 1250–1255.

Kolers, P. A. (1976). Buswell's discoveries. In R. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes* (pp. 371–395). Hillsdale, NJ: Erlbaum.

Matin, E. (1974). Saccadic suppression: A review. *Psychological Bulletin*, *81,* 899–917.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial fixations on words. *Vision Research*, *28,* 1107–1118.

McDonald, S. A. (2006). Parafoveal preview benefit in reading is only obtained from the saccade goal. *Vision Research*, *46,* 4416–4424.

Mitchell, D. C., Shen, X., Green, M. J., & Hodgson, T. L. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the Selective Reanalysis hypothesis. *Journal of Memory and Language*, *59,* 266–293.

O'Regan, J. K., & Lévy-Schoen, A. (1987). Eye-movement strategy and tactics in word-recognition and reading. In M. Coltheart (Ed.), *Attention and performance: Vol. 12. The psychology of reading* (pp. 363–384). Hillsdale, NJ: Erlbaum.

Penolazzi, B., Hauk, O., & Pulvermüller, F. (2007). Early semantic context integration and lexical access as revealed by event-related brain potentials. *Biological Psychology*, *74,* 374–388.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11,* 357–383.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52,* 1–56.

Radach, R., Inhoff, A. W., Glover, L., & Vorstius, C. (2013). Contextual constraint and *n* + 2 preview effects in reading. *Quarterly Journal of Experimental Psychology*, *66,* 619–633.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7,* 65–81.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124,* 372–422.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance*, *30,* 720–732.

Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic processing during eye fixations in reading. *Canadian Journal of Psychology*, *40,* 473–483.

Rayner, K., Juhasz, B. J., & Brown, S. J. (2007). Do readers obtain preview benefit from word *n* + 2? A test of serial attention shift versus distributed lexical processing models of eye movement control in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *33,* 230–345.

Rayner, K., Pollatsek, A., & Reichle, E. D. (2003). Eye movements in reading: Models and data. *Behavioral and Brain Sciences*, *26,* 507–518.

Rayner, K., & Reichle, E. D. (2010). Models of the reading process. *WIREs Cognitive Science*, *1,* 787–799.

Rayner, K., & Schotter, E. R. (2014). Semantic preview benefit in reading English: The effect of initial letter capitalization. *Journal of Experimental Psychology: Human Perception and Performance*, *40,* 1617–1628.

Rayner, K., Schotter, E. R., & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit in reading revisited. *Psychonomic Bulletin & Review*, *21,* 1067–1072.

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation during reading. *Perception & Psychophysics*, *31,* 537–550.

Reichle, E. D. (2011). Serial attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 767–786). Oxford, England: Oxford University Press.

Reichle, E. D., & Drieghe, D. (2013). Using E-Z Reader to examine word skipping during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39,* 1311–1320.

Reichle, E. D., & Laurent, P. A. (2006). Using reinforcement learning to understand the emergence of "intelligent" eye-movement behavior during reading. *Psychological Review*, *113,* 390–408.

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review*, *33,* 110–149.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105,* 125–157.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2012). Using E-Z Reader to simulate eye movements in non-reading tasks: A unified framework for understanding the eye-mind link. *Psychological Review*, *119,* 155–185.

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research*, *39,* 4403–4411.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, *26,* 445–476.

Reichle, E. D., & Reingold, E. M. (2013). Neurophysiological constraints on the eye-mind link. *Frontiers in Human Neuroscience*, *7,* 361.

Reichle, E. D., Tokowicz, N., Liu, Y., & Perfetti, C. A. (2011). Testing an assumption of the E-Z Reader model of eye-movement control during reading: Using event-related potentials to examine the familiarity check. *Psychophysiology*, *48,* 993–1003.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher-level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16,* 1–21.

Reilly, R., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research*, *7,* 34–55.

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z Reader model. *Psychological Science*, *17,* 742–746.

Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from survival analysis of fixation durations. *Cognitive Psychology*, *65,* 177–206.

Reingold, E. M., Yang, J., & Rayner, K. (2010). The time course of word frequency and case alternation effects on fixation times in reading: Evidence for lexical control of eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *36,* 1677–1683.

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, *26,* 1270–1281.

Schotter, E. R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language*, *69,* 619–633.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, *74,* 5–35.

Schotter, E. R., Reichle, E. D., & Rayner, K. (2014). Re-thinking parafoveal processing in reading: Serial-attention models can explain semantic preview benefit and N + 2 preview effects. *Visual Cognition*, *22,* 309–333.

Sereno, S. C., Brewer, C. C., & O'Donnell, P. J. (2003). Context effects in word recognition: evidence for early interactive processing. *Psychological Science*, *14,* 328–333.

Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a time-line of word recognition: Evidence from eye movements and event-related potentials. *Cognitive Neuroscience*, *9,* 2195–2200.

Staub, A. (2011). Word recognition and syntactic attachment in reading: Evidence for a staged architecture. *Journal of Experimental Psychology: General*, *140,* 407–433.

Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology: Human Perception and Performance*, *36,* 1280–1293.

Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, *30,* 415–433.

Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye movements in reading. *Psychonomic Bulletin & Review*, *14,* 770–775.

White, S. J., Warren, T., & Reichle, E. D. (2011). Parafoveal preview during reading: Effects of sentence position. *Journal of Experimental Psychology: Human Perception and Performance*, *37,* 1221–1238.

Yang, S.-N. (2006). An oculomotor-based model of eye movements in reading: The competition/interaction model. *Cognitive Systems Research*, *7,* 56–69.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46,* 441–517.

PART 4

Reading and Spelling Development

CHAPTER

20 How Children Learn to Read Words

Linnea C. Ehri

Abstract

A primary goal for beginners when they learn to read words is to secure spellings of the words to both their pronunciations and meanings in memory so that they can recognize the words immediately upon seeing them in or out of text. This requires orthographic mapping skill where knowledge of letter–sound units provides the connections for bonding spellings to pronunciations in memory. When readers apply decoding, analogizing, or prediction strategies to read unfamiliar spellings, this activates letter–sound connections and initiates bonding so that the words can be read from memory. When readers' mapping skill is activated as they read individual words in different, semantically rich sentences, spelling–meaning connections accumulate in memory. Development is portrayed by four phases, each characterized by the type of connections used to secure spellings to pronunciations in memory, ranging from visual/contextual, to partial alphabetic, to full graphophonemic, to consolidated syllabic and morphemic letter–sound spelling patterns.

Key Words: orthographic mapping, sight word learning, word spelling, developmental phases of word reading, vocabulary learning, decoding, analogizing, grapheme-phoneme relation, connectionist theory

When children learn to read, written language becomes connected to their growing competence with spoken language. In alphabetic writing systems the printed unit that maps directly and consistently onto spoken language is the word, so it is at a lexical level that print makes primary contact with children's linguistic knowledge as they read and comprehend text (Perfetti & Stafura, 2014). Much research has been conducted to clarify processes involved in children's acquisition of word reading ability. The focus of this chapter is on the English alphabetic writing system, although studies in other alphabetic systems are mentioned. Topics include the acquisition of written words, specifically spellings; various ways that children learn to read them; and how they are secured in memory by mapping phonemes, syllables, and morphemes in the pronunciations of words. In addition, the formation of connections between spellings and meanings is considered, particularly for words whose meanings require sentence contexts to be fully activated (e.g., prepositions, conjunctions, verbs).

Ways to Read Words

The English alphabetic system consists of graphemes symbolizing phonemes within words. *Graphemes* are single or double letters (e.g., ‹S›, ‹OU›, ‹TH›), and *phonemes* are the smallest units of sound within words (e.g., ‹SOUTH› has three phonemes). Connections between graphemes and phonemes provide foundational units that enable children to acquire word reading skill. The system also includes larger spelling patterns symbolizing spoken *rimes* (i.e., vowel and following consonants in a syllable such as /aʊθ/ in ‹SOUTH›), syllables, and *morphemes* (i.e., the smallest units of meaning, including root words and affixes). The English spelling system is variable and consists of alternative ways to represent phonemes and alternative sounds for graphemes. Spellings of some words are considered irregular when their pronunciations do not conform to the major grapheme-phoneme system. This variability and irregularity reduces the transparency of the English letter–sound system. As a result, reading acquisition is much delayed compared with more transparent writing systems such as Spanish (Seymour, Aro, & Erskine, 2003).

Children learn to read words in several ways (Ehri, 1998, 2014). If readers have never read the words before, they might apply one of three strategies. A decoding strategy involves applying their knowledge of the writing system to transform letters into sounds and to blend the sounds into a recognizable word. The letter units might consist of graphemes transformed into phonemes, or larger spelling units transformed into syllables or morphemes. An analogy strategy involves detecting within the spelling of an unfamiliar word (e.g., ‹PLUMP›) a letter pattern present in a known word (e.g., ‹UMP› in ‹JUMP›) and blending its sound with the other letters to pronounce the new word. A prediction strategy involves using partial letters and context cues such as pictures or adjacent words in a sentence to anticipate the identity of an unfamiliar word. These strategies might be applied to pronounce a recognizable word whose meaning children already know or a new word to be learned. Reading unfamiliar words is easier and more accurate when children already know the spoken forms and meanings of the words.

If children have practiced reading words and have retained their spellings in memory, the words are read from memory by sight. According to Ehri (1992, 1998, 2014), this involves storing connections between spellings, pronunciations, and meanings in memory. All words when practiced sufficiently become sight words that can be read immediately and automatically with little attention or effort (LaBerge & Samuels, 1974) and can be read as whole units rather than letter by letter (Ehri & Wilce, 1983). Being able to recognize words from memory facilitates text reading by making the word reading process unobtrusive and allowing attention to focus on the meaning of the text rather than on figuring out the words. Although word reading strategies are no longer needed to identify words known by sight, they may still function as backup sources used by readers to verify that the words pronounced match the spellings in print and fit the context.

Decoding

Decoding words requires knowledge and skills. Readers of English must process letters from left to right. They must possess knowledge of grapheme-phoneme relations, blending skill to merge the separate phonemes into a coarticulated whole, and vocabulary knowledge to recognize the meanings of the words they pronounce. Learning to convert graphemes into phonemes and to hold the sounds in memory long enough to blend them in the correct order requires instruction and practice. These are the skills taught to beginners in instructional programs to teach phonics systematically (Adams, 1990; Ehri, Nunes, Stahl, & Willows, 2001). The ability to decode is commonly assessed by having students read pseudowords, but this only taps the steps of sounding out and blending and omits meaning recognition.

At the outset of learning to read, beginners face the task of reading many unfamiliar printed words whose spoken forms are known. Applying a decoding strategy yields a blend of phonemes that may or may not match a known word. This is particularly true for irregularly spelled words such as ‹SAID› and ‹COME›, which might be decoded to rime with ‹PAID› or ‹HOME›. Tunmer and Chapman (2012) have shown that students who decode irregular words by testing approximations until they find a real word that fits the context are more accurate decoders. An example is decoding ‹STOMACH› as "sto-match," then "sto-mack," then "stu-muck." This preparedness for determining correct pronunciations is referred to as having a *set for variability*. Possession of a larger vocabulary enables greater success. Elbro, de Jong, Houter, and Nielsen (2012) showed that having a set for variability is important for recognizing not only irregular words but also more regularly spelled words, particularly when graphemes can symbolize more than one phoneme (e.g., ‹G› as /d͡ʒ/ in ‹GEM› vs. /g/ in ‹GET›, ‹A› as /ɔ/ in ‹TALL› vs. /æ/ in ‹PAL›).

The blending process is complicated also by the fact that when readers pronounce *stop consonants* in isolation (e.g., /b/, /d/, /t/, /p/ articulated by halting the air stream briefly), they must attach a vowel, usually schwa (e.g., /bə/), which then has to be deleted to produce the blend (e.g., /bə/ /æ/ /tə/ blended to form "bat") (Liberman, Shankweiler, Fisher, & Carter, 1974). This problem is avoided with graphemes that stand for *continuant consonants* (e.g., /s/, /m/, /n/, /f/), whose sound is not stopped and can remain connected during blending without adding a schwa (e.g., compare /bə-æ-tə/ for ‹BAT› to /mmænn/ for ‹MAN›). Sounding out letters is less disruptive to blending and word recognition when the stop consonant plus schwa occurs only at the end of the word (e.g., /ssopə/ for ‹SOAP›). This suggests that learning to decode is easier when beginners are taught with continuant consonants so that the phonemes can be pronounced and blended without any schwas causing breaks between phonemes.

As beginners' knowledge of orthographic regularities grows, larger spelling patterns are learned as units of print. This includes consonant clusters; rime units such as ‹ACK› and ‹UMP› that recur across words; high-frequency words such as ‹ON›, ‹IT›, and ‹UP› that are subunits of longer words; spellings of common syllables; and bound morphemes such as ‹ING›, ‹ED›, and ‹EST›. Only 37 rime spellings such as ‹ICK›, ‹-AIN›, ‹ELL›, ‹OKE›, and ‹UG› are found in nearly 500 English primary-grade words, indicating their value in decoding words (Mayzner & Tresselt, 1965). Treiman, Goswami, and Bruck (1990) provided evidence for the value of rime spelling units for decoding. They examined readers' ability to pronounce two types of consonant-vowel-consonant (CVC) nonwords: those containing rime spellings present in several real words (i.e., ‹TAIN›, ‹GOACH›), and those containing uncommon rime spellings (i.e., ‹GOAN›, ‹TAICH›). Both sets contained the same grapheme-phoneme units. Children and adults pronounced the words with common rime spellings more accurately than those with less common rime spellings.

Knowing multiletter spelling units enables readers to decode multisyllabic words. When words include more than a few letters, sounding out and blending separate grapheme-phoneme units becomes more difficult because it is harder to remember all the sounds in order to blend them. Being able to decode multiple letters as single units eases the decoding task by reducing the number of units to blend and thus speeds word reading (Juel, 1983). This is especially beneficial in English. Its writing system consists of more than 100 letters or letter combinations that symbolize forty-plus phonemes including fifteen vowels spelled in multiple ways (Moats, 2000). Decoding that is conducted with larger spelling units is more accurate, because patterns that are deemed irregular when viewed graphophonically become regular when the letters are learned as part of larger spelling units recurring across words—for example, ‹IGHT› in ‹SIGHT›, ‹LIGHT›, ‹MIGHT›, or vowel + ‹STLE› in ‹WHISTLE›, ‹CASTLE›, ‹WRESTLE›.

Once a sequence of letters is learned as a spelling unit, it may function as a symbol for more than one pronunciation; for example, the rime spelling ‹EAD› in ‹BEAD›, ‹READ›, ‹BREAD›, ‹HEAD›, pronounced either /id/ or /ɛd/. Although ‹EAD› pronounced /id/ is considered regular because the vowel digraph ‹EA› pronounced /i/ conforms to the major grapheme-phoneme system, ‹EAD› pronounced /ɛd/ occurs sufficiently often in different words to be considered systematic at the level of rime units. An extreme example of rimes with consistent spellings but highly variable sounds across words is ‹OUGH›, which occurs in ‹COUGH›, ‹TOUGH›, ‹THROUGH›, ‹THOUGH›, and ‹BOUGH›, with five different pronunciations. These examples reveal a source of consistency limited to spelling patterns without predictable pronunciations. Even though they are irregular graphophonemically, their orthographic consistency may help readers spell the words once they know the patterns.

Although readers' decoding skill is likely to benefit from knowing many grapheme-phoneme relations and spelling patterns, according to Tunmer and Nicholson (2011) there are too many to teach explicitly—probably several hundred, as estimated by Gough and Hillinger (1980). They suggest that children learn many of these implicitly once they receive explicit instruction in basic grapheme-phoneme relations and use them to build their sight word vocabularies. As sight words are retained in memory, their spelling-sound mappings provide the basis for inducing regularities (Thompson, Fletcher-Flinn, & Cottrell, 1999). This is thought to explain how readers acquire more extensive knowledge of the English writing system than they are taught explicitly.

Decoding performs several functions for readers. It enables them to pronounce words never read before. It also provides backup verification that the pronunciations of words read in other ways contain phonemes that match the graphemes seen in print. Importantly, it provides learning trials to establish written words in memory for sight word reading. Share (2004b, 2008) refers to decoding as a *self-teaching mechanism*. By decoding words, readers can teach themselves to read words from memory.

Analogizing

In order to read unfamiliar words by analogy to familiar words, students need to possess a bank of known spellings in memory. The greater the size of students' sight vocabularies, the more likely they are to possess a relevant analog that can be used (Leslie & Calhoon, 1995). In addition, students need to be aware of subunits within written and spoken words, such as rimes, in order to apply units in known words to read unknown words.

Use of this analogical strategy has been demonstrated in various ways. One way is to have students read nonwords that are analogous to irregularly spelled real words, such as ‹DUSY›-‹BUSY›, ‹BUITAR›-‹GUITAR›, ‹MONGUE›-‹TONGUE›. In these cases, analogizing can be differentiated from decoding because the two strategies yield different pronunciations (i.e., /dusi/ vs. /dɪzi/, /bʌtar/ vs. /bɪtar/, /mɑngu/ vs. /tʌŋ/, respectively). Another way is to show evidence of analogizing in a spelling task. Campbell (1983) found that priming students by having them read familiar real words (e.g., ‹BRAIN› or ‹CRANE›) influenced their spelling of analogous nonwords (e.g., /pren/ spelled ‹PRAIN› vs. ‹PRANE›).

Goswami (1986) studied analogizing in beginning readers. She displayed and pronounced a clue word (e.g., ‹BEAK›) and then asked children to read analogous and nonanalogous test words and nonwords (e.g., ‹BEAN›, ‹BEAL›, ‹PEAK›, ‹NEAK›, ‹LAKE›, ‹PAKE›). They read more analogous than nonanalogous words correctly, especially if the clue and test words shared rimes. Ehri and Robbins (1992) examined analogizing based on clue words that were stored in memory rather than exposed during the word reading task. Beginners were taught to read several clue words and then were shown either of two sets of test words containing the same letter–sound relations as the clue words. One analogous set also shared rime spellings with the training words. Children read the analogous words more accurately than the nonanalogous words. However, the benefit was limited to students who could decode nonwords, suggesting that some decoding skill was needed to read test words by analogy to words stored in memory. Nondecoders tended to misread test words as the originally taught clue words, indicating that confusion resulted from partial memory for letters, making clue and test words look like the same words.

Analogizing may be based not on a single word but on a neighborhood of known words having the same spelling patterns and subunit pronunciations. This has been shown to influence word reading. Khanna, Cortese, and Birchwood (2010) examined whether children's analogizing could be strengthened by teaching them to read sets of rime-based words whose spellings deviated from graphophonemic rules that did not take context into account (e.g., ‹DEAF› and ‹STROLL› deviating from spellings of vowels in ‹LEAF› and ‹DOLL›). Groups of words with the same rime-based spellings and pronunciations were taught and practiced together (e.g., ‹FIND›, ‹BLIND›, ‹KIND›, ‹MIND›), and the teacher explained how they deviated from grapheme-phoneme rules. Also, groups of regularly spelled words were explained and taught together (e.g., ‹GLOBE›, ‹LOBE›, ‹PROBE›, ‹ROBE›). On pre- and posttests containing nonwords spelled analogously to the real words taught, students showed significant increases in analogizing on both types of rime spellings.

Programs have been developed to teach beginners to read words by analogy. One example is the key word method developed at Benchmark School to help struggling readers (Gaskins et al., 1988). Over the course of a year, students were taught to read 120 high-frequency words with common spelling patterns and to use them to read unfamiliar words by analogy. Several years later the program was modified because some students had trouble remembering spellings of the key words (Gaskins et al., 1996–97). The solution was to teach students how graphemes symbolized phonemes in the key words. Students practiced counting the phonemes in spoken words, then they looked at spellings, matched graphemes to phonemes, and reconciled mappings when there were more letters than phonemes (e.g., final ‹E› marking a long vowel or digraphs such as ‹SH›). Then they practiced spelling the words from memory. This procedure was expected to enhance students' memory for the key words (Ehri, 1992). A longitudinal comparison of the two approaches (Ehri, Satlow, & Gaskins, 2009) revealed that students taught the modified key word program outperformed students in the original program on measures of reading and spelling words during the first two years. However, the difference was diminished during Years 3 and 4, when students in the original program caught up to the modified group in reading words.

Prediction

To apply a prediction strategy to read unfamiliar words in text, readers may access background knowledge, draw from information they have read up to that point or look ahead in the text, sound out initial letters or match letters to familiar words, or look at pictures on the page. Prediction is the easiest among the word reading strategies to apply and appears early in development, although it may not be accurate. Reading errors have been interpreted to indicate prediction. If children's misreadings fit the sentence structure and meaning of the text, this suggests that they are using context to predict. If sounds in a substituted word correspond to some of the letters in print, this indicates that the reader's prediction is influenced by the word's spelling (Biemiller, 1970). A *cloze task* requiring readers to fill in missing words in a text assesses use of a prediction strategy.

Prediction serves several purposes during text reading. It helps readers read unfamiliar words that are difficult to decode. It speeds the reading of familiar words. It provides backup confirmation that words read in other ways fit the context and initiates self-checking if they do not fit (Clay, 1985). For readers with weak decoding skill, it improves their ability to read words in text. Goodman (1967) labeled reading a psycholinguistic guessing game. He claimed that good readers use context to predict the words in text and attend to letters only as necessary. However, Stanovich (2000) conducted many studies showing that it is not good readers but instead poorer readers who rely on context and prediction to compensate for deficient decoding skill when they read words in text. Good readers are skilled decoders and can recognize words accurately and quickly in or out of text, so reliance on context is not necessary. In support of this, Nicholson (1991) found that beginning and poorer readers read words better in context than in isolation, but 8-year-old good readers showed no significant benefit of context.

There have been heated disagreements among educators about how beginning reading should be taught. One battle has involved the relative importance of prediction versus decoding as the most effective approach for teaching beginners to read (Stanovich, 2000). Proponents of the *whole-language* approach have regarded predictable books as most appropriate for beginning readers. Words in predictable texts are read primarily using context, pictures cues, and repetitive sentence stems. Prediction is a major reading strategy and receives priority over a decoding strategy in instructional programs such as Reading Recovery (Clay, 1985; Tunmer & Hoover, 1993). In contrast, systematic *phonics* programs instruct beginners to pay primary attention to letter–sound relations to decode words in books with vocabulary tailored to their letter–sound knowledge as it accumulates.

Tunmer and Hoover (1993) conducted an experiment to compare the standard Reading Recovery (RR) program with a modified version of RR that included decoding instruction. The participants were first graders who were at risk in learning to read. Although the RR program did not teach children to sound out and blend words, it did teach children to spell the sounds in words. By the end of instruction when students had reached mastery levels qualifying them to discontinue the programs, the two groups did not differ on word reading and writing posttests. However, students who had been taught decoding took significantly less time to complete the program than students in the standard RR program, with means of 42 vs. 57 lessons. In other words, the instruction provided in the standard RR program was 37% less efficient. These findings suggest the importance of combining instruction in the use of prediction with decoding instruction.

Reading Words From Memory by Sight

Once the written forms of words become familiar and are stored in memory, they are read by sight. A common belief is that only high-frequency words and irregular words are read by sight. However, this appears not to be true. Rather evidence suggests that all words are read from memory by sight once readers have practiced reading them.

One compelling finding comes from the Stroop task (Stroop, 1935). Readers are shown written words naming colors, but the font is printed in a color that is discrepant with the color named (e.g., the word ‹RED› is printed in green letters). People are instructed to name the color of the letters and ignore the color named by the word. In another form of the task, readers are shown drawings of objects with a printed word naming a different object planted on the drawing, such as a horse with the letters ‹COW› printed on the picture, or an apple with ‹ORANGE› printed on it. Readers are instructed to name the pictures as quickly as possible and ignore the words.

Studies show that readers find it impossible to ignore the words. Regardless of their intent, the presence of discrepant words slows them down in naming the colors or pictures, much more than the presence of nonwords printed on the pictures (Rosinski, Gollinkoff, & Kukish, 1975). Responses are especially slow when the picture and word are in the same semantic category, as in the horse-cow and apple-orange examples. The explanation for the slowdown is that sight of familiar words activates their pronunciations and meanings quickly and automatically in memory, faster than the names of the colors or pictures can be retrieved, and this competition impedes retrieval of the color and picture names. Interference from sight words emerges as young as first grade (Guttentag & Haith, 1978).

Another important property of sight word reading is that the words are pronounced as single units without pauses between letters or word parts, referred to as *unitization*. Evidence for this comes from studies showing that readers can read words containing multiple letters as quickly as they can read single-syllable digits. In a study measuring children's reaction times to initiate pronunciations of single items presented on slides, Ehri and Wilce (1983) showed that second- and fourth-grade good readers read familiar object words such as ‹MAN›, ‹CAR›, ‹DOG›, ‹BALL›, ‹HAT› as quickly as they named digits such as 2, 6, 3, 5, 4, 9. This indicates that the words were read as single units rather than by sounding out the separate letters. Children read unfamiliar nonwords such as ‹JAD›, ‹TUK›, ‹NEL›, ‹FUP›, ‹MIG› much more slowly than they read the familiar words, showing the advantage of reading words as single units from memory over using a decoding strategy. The same pattern of findings was evident among children who had learned to read in Portuguese, a more transparent writing system. Even though familiar words could have been read accurately by decoding because grapheme-phoneme correspondences in the writing system are highly predictable, children read them as single units from memory (Defior, Cary, & Martos, 2002).

Reading words from memory is more efficient than application of strategies. When readers read text, their attention is focused on comprehending the meaning. If they have to stop and decode, or analogize, or predict, their attention is shifted away from meaning while the word is identified. However, if the words can be read from memory, there is little disruption. Enabling students to read words automatically from memory is essential for improving text reading skill.

ORTHOGRAPHIC MAPPING

Research has changed our explanation of how sight words are learned. An earlier view was that readers use visual cues and memorize the shapes of words to remember how to read them (Barron, 1981; Henderson, 1980). However, visual cues cannot be the explanation for several reasons. The shape envelopes and visual features of words are not sufficiently distinctive to discriminate among all the thousands of words stored in readers' memory. Readers should mistake similarly shaped words, yet evidence shows that word reading is highly accurate. Similarly spelled words are not often confused.

If visual cues were the basis for remembering words, lots of practice would be required because the connections between spellings and pronunciations would be arbitrary, not systematic. However, evidence shows that readers store written words in memory very quickly, without much practice (Ehri, 1980; Ehri & Saltmarsh, 1995). In one study, first graders required four exposures to words to be able to read them faster than alternative homophonic spellings of the same words they had not seen (Reitsma, 1983). In another study (Share, 2004b), third graders required only a single exposure to retain information about the words' letters in memory. They read novel words aloud in stories. Each word contained a target sound that could be spelled in either of two ways. Then, three, seven, or thirty days later, children's memory for the target letters they had seen was tested by having them write the words. Memory for these letters was significantly greater than chance, even as long as thirty days later. To explain word learning that occurs this easily, readers need to possess a powerful mnemonic system, one that works like very strong glue to stick the spellings of words in memory.

Ehri (1992) hypothesized that readers remember sight words by forming connections between the spellings of individual words and their pronunciations (phonological representations). The glue that bonds them is provided by readers' knowledge of the mapping system consisting of grapheme-phoneme relations that secure graphemes in spellings to phonemes detected in pronunciations of words. For example, four connections secure the graphemes in ‹STOP› to phonemes in the pronunciation, /s/-/t/-/ɑ/-/p/. Three connections secure the graphemes in ‹CHECK› to its phonemes, /t͡ʃ/-/ɛ/-/k/. Note that connections are systematic, not arbitrary. Connections would not be formed if the spelling ‹BOT› was pronounced /t͡ʃ/-/ɛ/-/k/.

As grapheme-phoneme relations are applied to retain the spellings of words in memory, those letter combinations that are read frequently become consolidated into larger multiletter sequences. These sequences are built out of grapheme-phoneme units that become unitized to represent rimes, syllables, or morphemes (e.g., ‹ED›, ‹ION›, ‹ING›, ‹CON›, ‹ENT›, ‹MENT›). These serve as units that bond spellings to pronunciations in memory. Larger units reduce the number of connections needed to remember multisyllabic words; for example, four graphosyllabic units rather than nine to ten graphophonemic units in ‹AD MIN IS TER›. Meanings are bonded to spellings as well. Together, these identities form an amalgam representing individual words in memory. This view portrayed by Ehri (1992, 2014) resembles views of other theorists (Perfetti & Hart, 2002; Rack, Hulme, Snowling, & Wightman, 1994; Share, 2008). This explanation may be difficult to understand because it is counterintuitive. It involves the retention of visual forms in memory by phonological means and thus conflicts with the traditional view that visual memory is separate from phonological memory. However, results of functional magnetic resonance imaging studies are consistent with this view. They show that a visual word form area becomes linked to a phonological area in the brain when children learn to read (Frost et al., 2008).

One argument against this view rests on the claim that letters in irregularly spelled words lack the grapheme-phoneme consistency needed to store the words in memory, so they must be remembered in some other way. However, this is not true. Most letters in irregularly spelled words are regular and conform to the mapping system, either at the grapheme-phoneme level (e.g., all letters but ‹I› in ‹FRIEND› and ‹W› in ‹ANSWER› or at the level of multiletter units (e.g., ‹EAD› in ‹DEAD›). According to connectionist theory, irregularly spelled words are stored in memory in the same way as regularly spelled words.

In order to form connections and secure spellings of words in memory, requisite knowledge and skills are needed. Beginning readers need segmentation skill to analyze pronunciations into phonemes. They need knowledge of the writing system, principally grapheme-phoneme correspondences. They need to apply their graphophonemic knowledge to connect graphemes in spellings to phonemes in pronunciations to form the bond that is retained in memory, referred to as *graphophonemic* or *orthographic mapping*. In addition, the phonological representations of the words have to be sufficiently precise in order for graphemes to link up to the phonemes that they symbolize (Elbro, 1997). For example, the pronunciation "going to" is more precise and maps onto ‹GOING TO› more accurately than "gonna." Also, readers need to know the meanings of the words so they become bonded to their spellings and pronunciations in memory. Once readers know the graphophonemic system, they do not even need to see a word to have definite ideas about its letters. Just hearing the word activates expectations about its spelling (Stuart & Coltheart, 1988).

Readers with the requisite knowledge can retain sight words in memory as a result of several word reading events. If students decode the word by sounding out and blending letters, or if they analogize, this will activate connections and secure the spelling in memory. If students are told how to pronounce an unfamiliar spelling they are viewing, the connections can become activated in memory (Rosenthal & Ehri, 2008). If students use context plus partial letters to predict an unfamiliar word, connections between the spelling, pronunciation, and meaning will be activated provided they know the spelling system.

Not only reading but also spelling words creates orthographic representations in memory. Reading and spelling are very closely related (Ehri, 1997). In studies examining whether training in the formation of connections to read words also improves spelling, findings are typically positive. In correlational studies, coefficients between the two skills are very high, typically above .70. Reading words and spelling words are both governed by processes that include knowledge of the orthographic system and use of this knowledge to form connections and retain individual words in memory. This suggests the importance of teaching reading and spelling to strengthen both skills.

Learning to read unfamiliar words from memory presents problems for struggling readers. One problem involves phonological difficulties of various kinds (Shankweiler et al., 1995). Studies have shown that students with a reading disability may have limited phonemic awareness and weak phonological working memory, and their phonological representations of words may be imprecise (Elbro, Borstrøm, & Petersen, 1998). Another problem is that they have not mastered the major grapheme-phoneme relations so this limits their ability to phonologically decode unfamiliar words (Rack, Snowling, & Olson, 1992). As a result, they lack the requisite skills for forming complete connections between spellings and pronunciations of words to store them in memory. The connections are partial and incomplete. When they encounter unfamiliar words in text, they compensate for poor decoding skill by predicting words using partial letters and context cues (Stanovich, 2000; Rosenthal & Ehri, 2011). As a result, they do not retain fully connected words in memory to support accurate sight word reading.

Phase Theory of Development

Ehri (2005a, 2005b) has distinguished four phases to identify significant advances that occur as children learn to read words from memory. They are the *prealphabetic, partial alphabetic, full alphabetic,* and *consolidated alphabetic* phases. The phases are labeled to reflect the type of alphabetic knowledge that predominates in the spelling–sound connections that are formed to bond spellings to their pronunciations in memory. During the prealphabetic phase, the connections are visual or contextual and do not involve letter sounds. During the partial alphabetic phase, some of the letters in words are connected to their sounds. During the full alphabetic phase, more complete grapheme-phoneme connections are formed. During the consolidated phase, larger spelling patterns involving syllables and morphemes are used to form connections. The transition from one phase to the next is gradual rather than discrete and stage-like. At any point the child may be using more than one type of connection, although one type predominates. Development is governed by the child's knowledge of the alphabetic writing system and its use to read words, not by age or grade level.

The course of development of phases from prereading to fluent reading is summarized in Table 20.1. Specific acquisitions are listed including requisite knowledge and subskills, characteristics of sight word memory, strategies for reading unfamiliar words, and spelling.

Prealphabetic Phase

During the earliest period, children cannot sound out and blend letters in words. They cannot read text independently. They do not use letter–sound connections to invent spellings of words or to remember how to read or spell words, even though they might know some letters. They write words with random letters, pseudoletters, or memorized letters. Often they know the letters in their own names, but these are memorized, not connected to sounds in the name. Children can pretend to read books they have heard over and over, but they are only reciting what they have memorized. This is evidenced by their inability to point to individual words as they recite the text (Ehri & Sweet, 1991; Morris, 1983). If they can read any words, they do so by relying on salient visual-semantic cues such as the two eyeballs in ‹LOOK› (Gough, Juel, & Griffith, 1992). Environmental print is read from contextual cues in or around the written words such as the golden arches to read ‹McDONALDS›, not from letters.

In one study, we selected familiar signs and labels that young children could read in their environment (Masonheimer, Drum, & Ehri, 1984). We altered letters in the labels—for example, ‹X› replaced the initial ‹P› in ‹PEPSI› to form ‹XEPSI›. Prealphabetic readers did not notice the change, even when we asked them whether there was a mistake. Most still read the label as "Pepsi." Although they knew about 60% of letter names, they did not use them. This shows that they were reading the environment rather than the print.

Because the surrounding cues are especially salient in environmental print, this may distract prealphabetic readers from noticing letters. Studies of personal name recognition have shown that prereaders do use letters to read their own names and some classmates' names, even when the names are removed from classroom lockers and presented in isolation (Levin & Ehri, 2009; Share & Gur, 1999; Treiman & Broderick, 1998). However, the letters are memorized visual forms rather than connected to pronunciations of the names.

Children in the prealphabetic phase are essentially nonalphabetic readers. Their feats of reading are performed by using cues that do not involve the letter–sound system. To transition to the next phase, children need to acquire alphabetic skills. One of the best predictors of children's success in learning to read during kindergarten and first grades is their knowledge of letter names or sounds when they enter kindergarten (Share, Jorm, Maclean, & Matthews, 1984). Roberts (2003) found that teaching letter names to preschoolers facilitated their use of letter sounds to read words from memory. Piasta, Petscher, and Justice (2012) studied preschoolers' letter knowledge and their success in learning to read in first grade. They identified an optimal benchmark for predicting success: knowing eighteen uppercase and fifteen lower case letters at the end of preschool.

Many letter names contain relevant sounds in their names; for example, ‹B› contains /b/. If children know the names, it is easier to teach them sounds found in the names (Share, 2004a). An effective way to teach letter–sound relations is with embedded picture mnemonics (Ehri, Deffner, & Wilce, 1984; Shmidman & Ehri, 2010). This involves presenting pictures of objects or characters whose inherent shapes resemble the shapes of letters and who have names that begin with the sound of the letters, such as a drawing of a snake resembling ‹S›, a drawing of a table resembling ‹T›, or a drawing of mountains with peaks resembling ‹M›. Teaching

Table 20.1 Summary of the Emergence of Knowledge, Skills, and Strategies characterizing Ehri's Prealphabetic, Partial Alphabetic, Full Alphabetic, and Consolidated Alphabetic Phases of Development in Learning to Read and Spell Words.

Prealphabetic	Partial Alphabetic	Full Alphabetic	Consolidated Alphabetic
REQUISITE KNOWLEDGE AND SUBSKILLS			
Limited or no letter knowledge	Most letter names and some grapheme-phoneme correspondences (GPs)	Major GPs and some larger spelling units	Many graphosyllabic and morphemic spelling units
Lack of phoneme segmentation	Partial phoneme segmentation	Full phoneme segmentation	
No GP mapping Growing knowledge of spoken language: pronunciations, syntax, meanings of words	Partial GP mapping; correct directional orientation to print Growing knowledge of spoken language continues	Complete GP mapping Growing knowledge of spoken language continues	Graphosyllabic and morphemic as well as GP mapping Growing knowledge of spoken language continues
SIGHT WORD MEMORY			
Reading words by remembering salient visual or context cues; semantic substitution errors; no letter–sound connections; memory unreliable except for personal name	Reading words by remembering partial GP connections; confusing similarly spelled words	Reading words by remembering full GP connections; accuracy, automaticity, and unitization emerging	Reading words by remembering larger spelling units as well as GP connections; accuracy, automaticity, unitization established for known words
STRATEGIES TO READ UNFAMILIAR WORDS			
No word decoding ability Cannot analogize Words predicted from visual cues, context, pictures	No word decoding ability Cannot analogize Words predicted using initial letters and context	Growing ability to decode unfamiliar words using GPs Limited use of analogizing due to small sight vocabulary Prediction to support and confirm words decoded or read by analogy	Proficient decoding of unfamiliar words using GPs or larger units Greater use of analogizing as sight word vocabulary grows Prediction to support and confirm words decoded or read by analogy
SPELLING			
Nonphonetic spellings of unfamiliar words using scribbling, pseudoletters, or letters No memory for correct spellings except for personal name	Partial phonetic spellings of unfamiliar words using letter names or GPs Limited memory for correct spellings	Complete phonetic spellings of unfamiliar words using GPs Good memory for correct spellings of many known words	Graphosyllabic and morphemic units as well as GPs to spell unfamiliar words; Proficient memory for correct spellings of known words

Note. GP refers to grapheme-phoneme relations.

these associations makes learning letter sounds easier for children than teaching arbitrary associations.

Partial Alphabetic Phase

To move into the next phase, not only do children need to acquire letter name and sound knowledge but also they need to become aware of sounds present in spoken words, beginning with initial sounds. They need to recognize the connection between some of the letters in the spellings of words and corresponding sounds in pronunciations. Because their knowledge is limited, their ability to read and spell words is partial and incomplete. When they invent spellings they write only some of the sounds they hear, typically the first and final sounds. Memory for correct spellings of words eludes them because they lack full knowledge of the spelling system. They have not learned to sound out and blend words. They can guess words from partial letters and context clues. They can remember how to read words by forming connections between some of the letters and sounds in words, much like their invented spellings. For example, to remember how to read ‹SOUP›, they might connect ‹S› and ‹P› to initial and final sounds detected in the pronunciation. These letters may be selected because the relevant sounds are detected in their letter names. Other letters are not remembered, at least not by connecting graphemes to phonemes. One reason is that segmenting the phonemes in words is incomplete. For example, the two phonemes in consonant clusters (e.g., /st/ in ‹STOP›) are hard to detect, so both letters may not be remembered. Another reason is that vowel spellings are more complex and may not be known. Because the connections are partial, word reading and spelling are not very accurate and similarly spelled words may be mixed up.

Ehri and Wilce (1985) selected children in the prealphabetic and partial alphabetic phases and compared their word reading to see whether the two groups would use different types of cues to read the words. Children in the prealphabetic phase knew only a few letter sounds and could not read any words. Children in the partial phase knew most of the letter sounds and could read a few easy words, but they could not decode words. They were given several practice trials to learn to read two types of simplified spellings, one type spelled with letters that were distinctive visually but bore no relationship to sounds in the words (e.g., ‹wBc› for *elephant*, ‹FO› for *arm*). The other type contained letter-name and letter-sound cues, much like partial invented spellings (e.g., ‹LFT› for *elephant*, ‹RM› for *arm*). Children in the prealphabetic phase learned to read the visual spellings more easily than the letter-sound spellings, whereas children in the partial alphabetic phase learned letter-sound spellings more easily than the visual spellings. Other researchers have replicated these findings (de Abreu & Cardoso-Martins, 1998; Roberts, 2003; Treiman & Rodriguez, 1999). These studies show that when children learn the names or sounds of alphabet letters, they become able to read words in a different, more effective way, a way that uses the graphophonemic mnemonic system for remembering how to read words.

The partial phase depicts what beginning readers are able to do when their reading abilities are limited. They have some knowledge of the alphabetic writing system and its application to read and spell, but they lack the competence needed to decode words and to store the complete forms of written words in memory. Students with reading disabilities exhibit characteristics of the partial alphabetic phase in their reading and spelling.

To transition from the partial to the full phase, three requisites are needed for more effective sight word learning: letter-sound knowledge, segmentation of words into phonemes, and mapping letter–sound connections within words. Ehri and Wilce (1987b) showed that partial phase readers who were taught to spell words by segmenting them into phonemes and selecting corresponding letters learned to read words from memory better than students who just practiced the individual letter–sound relations. Boyer and Ehri (2011) showed that teaching partial phase readers to spell words by distinguishing the articulatory features of phonemes as well as letter sounds enhanced beginners' success in learning to read words from memory compared with beginners taught to spell with letter sounds but not articulatory features. Children who received articulatory training used mirrors to observe how their mouths moved when pronouncing phonemes in words, and they learned to use tiles displaying mouth positions to segment words into phonemes—for example, selecting a picture of the lips closed and then a picture of the lips smiling to depict the two phonemes in /mi/. Boyer and Ehri suggest that motoric gestures are more central defining properties of phonemes than acoustic features, based on the motor theory of speech perception (Liberman, 1992). These findings suggest the importance of teaching students to monitor their mouth movements as well as sounds in words during phonemic awareness instruction.

The aim of systematic phonics instruction is to teach the knowledge and skills necessary to move beginners to the full alphabetic phase. In a meta-analysis, Ehri et al. (2001) found that systematic phonics instruction boosted sight word reading, decoding, and reading comprehension more than other kinds of instruction including whole-word and whole-language instruction. Effects were especially pronounced in kindergarten and first grade, when students were just beginning to read. Studies have shown that phonics-trained beginning readers are less apt to guess unfamiliar words based on partial letters and context cues and are more apt to apply a decoding strategy in reading words than children receiving whole-word beginning reading instruction (Barr, 1974–75; Carnine, Carnine, & Gersten, 1984).

Full Alphabetic Phase

Children become full alphabetic phase readers when they can learn sight words by forming complete connections between graphemes in spellings and phonemes in pronunciations. This is possible because they have learned the major grapheme-phoneme correspondences and they can segment pronunciations into phonemes. When readers apply this knowledge to individual words to form connections, graphemes in spellings become fully bonded to phonemes in pronunciations along with meanings and are stored together as an amalgam in memory. In the full phase, with phonics instruction, children learn to sound out and blend letters to read unfamiliar words. This decoding strategy serves as an important means of getting sight words into memory by activating the relevant connections.

The difference between partial and full phase readers in sight word learning was shown by Ehri and Wilce (1987a). Kindergarteners in the partial alphabetic phase received either of two types of training. One group was taught to process all the grapheme-phoneme connections in words so that they would function like full phase readers. They practiced reading and rereading twelve sets of similarly spelled words and nonwords, such as ‹BAP›, ‹DAT›, ‹LAB›, ‹PAM›, ‹RAS›, ‹SAN›, ‹TAD›. The other group received training in the same grapheme-phoneme relations, but these were practiced individually rather than blended in words. As a result of this training, children in this group remained partial phase readers. At the end of training both groups practiced learning to read fifteen words over several trials. They read each word and were corrected if wrong. The words were similarly spelled, making it hard to remember them using partial cues; for example, ‹BEND›, ‹BIB›, ‹BLAST›, ‹BLOND›. Those who had received full-phase instruction learned to read on average 90% of the words within three trials. In contrast, the partial phase readers read on average only 30% of the words by Trial 3. Their main problem was mixing up similarly spelled words, showing the limitation of partial cue reading. These findings underscore the great advantage to readers when they can form full connections in learning to read words by sight.

Several other capabilities distinguish readers in the full phase (see Table 20.1). Their word reading is more accurate than that of partial phase readers. Their sight word vocabularies are expanding rapidly. They can recognize words automatically and faster. As the store of sight words increases, they can learn the strategy of reading words by analogy to words they already know. They are able to invent more complete grapheme-phoneme spellings of words. It is easier to remember the correct spellings of words that are consistent with their knowledge of the spelling system. They can read text independently when the text contains mostly words they know by sight or can figure out by decoding, analogizing, or prediction.

A few years ago, a controversy arose about whether phase theory applies to children learning to read words in transparent orthographies, which can be decoded easily. Wimmer and Hummer (1990) tested the decoding ability of German-speaking beginning readers and found no one in the partial phase. They suggested that this phase was absent in transparent writing systems. However, their beginners had already received six months of systematic phonics instruction. These children may have been observed too late to detect evidence of an earlier and possibly brief partial phase. Cardoso-Martins (2001) also examined this question with two groups of Portuguese beginning readers in Brazil. One group was taught with a whole-word approach and the other with a phonics approach. Three months into the school year, students in the whole-word group read at the partial phase. They could read familiar words but could not decode nonwords. In contrast, the phonics group had received decoding instruction and could read words and nonwords at the full phase. These findings indicate that partial phase readers can be found in transparent writing systems if they are tested early and if they have not been taught to decode words.

Share (2008) has studied decoding as a self-teaching mechanism to store sight words in memory. He and others have shown that students remember the specific spellings of unfamiliar words they read in text and do not confuse them with plausible alternative spellings not seen. Share (2004b) found that one exposure to words was sufficient to improve third graders' memory for letters in the words as long as a month later. One possible reason that beginners who possess full phonics skills can retain sight words in memory with minimal experience reading the words is offered by Stuart and Coltheart (1988). They suggest that students' grapheme-phoneme knowledge leads them to expect specific connections between spoken and written words and that this facilitates memory for new spellings. That readers can use spelling–sound expectations to learn words was shown by Ehri and Wilce (1979). They gave second graders several trials to learn spoken CVC nonwords paired with numbers. Those who were instructed to imagine how the nonwords were spelled recalled the spoken nonwords better over the learning trials than those who simply repeated the spoken forms without any mention of spellings.

Consolidated Alphabetic Phase

As readers retain more and more sight words in memory they move toward the consolidated phase of development. Letter patterns that recur in different words become familiar and unitized, such as the ‹AMP› in ‹CAMP›, ‹DAMP›, ‹LAMP›, ‹CHAMP›. Other examples of larger units are spellings of vowel-consonant endings, syllables, root words, prefixes, and suffixes. These letter chunks are used to form connections in learning multisyllabic words. For example, ‹IN TER ES TING› might be segmented into these letter units to represent syllables in the word's pronunciation and these connections retained in memory along with meaning.

To examine sight word learning during the consolidated phase, Bhattacharya and Ehri (2004) studied older struggling readers (mean age 13.8 years) who were reading at a third-grade level as measured by a standardized word reading test. They examined whether teaching them to read multisyllabic words such as ‹SUBSTITUTION›, ‹CONFERENCE›, ‹DEMOCRATIC› by forming syllabic connections would improve their sight word reading. Students practiced reading 100 multisyllabic words by dividing the words into spoken syllables and matching the spoken to the written syllables. Different places for breaking words into syllables were accepted as long as the vowels were separated and spoken forms matched the spelling. A control group practiced reading the same words, but they read them as whole words and read them extra times. On posttests the syllable group outperformed the whole-word group in reading words, in decoding nonwords, and in spelling words. This provides evidence for the consolidated phase of development. It shows that practice forming connections between spoken and written syllabic units in words improves students' word reading skills and also their spelling memory.

To compare the contribution of smaller graphophonemic and larger morphemic units to word reading fluency and comprehension, Nunes, Bryant, and Barros (2012) assessed English-speaking 8- and 9-year-olds who presumably had moved into the consolidated phase of development. Use of graphophonic units and morphemic units were assessed with word, pseudoword reading, and spelling tasks. Results of hierarchical regression analyses showed that both the graphophonic and morphemic factors explained significant unique variance on measures of sight word reading, text reading fluency, and comprehension after controlling for age and verbal IQ. However, greater variance was explained by morphemic knowledge. These findings suggest that both small and large units contribute to orthographic connections in lexical memory, with larger units playing a greater role than smaller units among children in the consolidated phase.

Role of Context in Learning to Read Words

Phase theory addresses the course of development in learning to read words as children progress from nonreaders to skilled readers who are able to read text independently and fluently. Fluency is evidenced when students can read a text aloud with sufficient speed, appropriate phrasing, and expression (Kuhn & Stahl, 2003). An important achievement contributing to fluency is being able to recognize single written words as units immediately upon seeing them, with their pronunciations and meanings recognized automatically. Automatic recognition is explained in part by orthographic mapping, which secures the spellings of individual words to their pronunciations in memory as units ready to be activated when spellings are seen in text. However, another part of the explanation involves the activation of connections between words in memory.

According to Ehri's (1978, 1980) amalgamation theory as well as other connectionist theories, syntactic and semantic identities become connected to individual spellings when words are combined with other words in sentences and texts, and meanings are comprehended. These encounters build up potential syntactic functions and meanings. Which among them is activated depends on the particular context in which the spelling appears and is read. Researchers have paid less attention to children's acquisition of connections between spellings, syntactic functions, and meanings of words, yet this is an important part of development.

All words are not the same. A common way to teach beginners to read words has been to have them practice reading words in isolation on cards. Although this approach may activate connections between graphemes and phonemes, it neglects the formation of connections between spellings and their syntactic and semantic identities. This may not be a problem for beginners reading semantically rich content words (e.g., ‹TABLE›, ‹GHOST›, ‹HORSE›) whose pronunciations activate relatively complete meanings. However, many words depend on the presence of other words for their syntactic functions and meanings to be activated. This includes words such as past-tense irregular verbs (e.g., ‹SAID›, ‹WENT›, ‹DID›, ‹CAME›), prepositions, conjunctions, and other function words (e.g., ‹FROM›, ‹WITH›, ‹IS›, ‹ARE›, ‹AND›, ‹FOR›, ‹THE›). To connect syntactic and semantic identities of context-dependent words such as these to their spellings, children must practice reading them in text. This is especially important for developing fluency because context-dependent words provide the glue that connects words within sentences. Moreover, context-dependent words are among the first that children learn to read. They are the high-frequency words that appear on preprimer and primer lists.

Results of a study by Duff and Hulme (2012) suggest that context-dependent words may be harder for beginners to learn than semantically rich words when practiced in isolation. They taught 5- and 6-year-olds to read high-imagery and low-imagery words over six learning trials. Children pronounced the words but were not told their meanings. Children learned the high-imagery words more easily. However, whether imagery was the cause remains unclear. All the high-imagery words were concrete nouns (i.e., ‹LADDER›, ‹RIVER›, ‹JACKET›, ‹KNIFE›), whereas many of the low-imagery words were context-dependent words that were not nouns (i.e., ‹HIDDEN›, ‹LOOKED›, ‹NEVER›, ‹BETTER›, ‹BECAUSE›).

Studies show that if function words are heard in isolation, prereaders may not recognize them as real words. Although young children can combine content and function words to produce meaningful sentences, Huttenlocher (1964) and Holden and MacGinitie (1972) found that they had difficulty segmenting sentences into separate words. Children could pick out content words, but not function words. Often they would group articles or prepositions with adjacent content words.

Ehri (1975) gave young children several tasks to assess their facility with context-dependent words. In one task, children were directed to embed individual spoken words in sentences. Prereaders had difficulty with words such as "ran" and "and," but not with "wagon." Some changed the word to "run" or embedded "ran" in a name such as "Randy." After the tester modeled a sentence containing "and," "The boy and his dog walked home," one child generated a parallel meaning but failed to include the target word, saying "The girl is walking home with a cat." When asked to segment sentences into words, children overlooked function words such as "the," "to," "is," "my," "has," and "of." In contrast to prereaders who showed a lack of awareness of the identities of function words, beginning readers performed these tasks easily. They recognized the words spoken in isolation, they could embed them in meaningful sentences, and they could segment spoken sentences into their words.

In another study, Ehri (1976) gave kindergarten and first-grade readers and prereaders a paired-associate word learning task in which familiar spoken content and function words were paired with meaningless squiggles. No spellings were shown. For example, one set included the words "milk," "small," "came," "of," and "and." The task was to recall each word when its squiggle was shown over several learning trials with corrective feedback. Both readers and prereaders learned the nouns and adjective associations more quickly than the verbs, prepositions, and conjunctions. However, it took the prereaders an especially long time to learn the past-tense verb and function words compared with the nouns and adjectives, and much longer to learn them than the readers, even though the two groups were matched in age.

In explaining these findings, Ehri (1976) suggests that learning to read words in text and retaining spellings of the words in memory serve to enhance beginning readers' awareness of the lexical structure of spoken language and the separate identities of words, particularly context-dependent words. Contrast the spoken form, "Gimme apiecea cake" to its written form "Give me a piece of cake." Achieving word consciousness by seeing the spellings of words in print serves to render children's implicit knowledge of individual words explicit and available for use in reading and writing. This is particularly true for context-dependent words. Learning the spellings of these words converts them from transitory, opaque, conflated spoken utterances to separate, concrete visual lexical forms having distinctive orthographic, phonological, syntactic, and semantic identities. Their presence in the structure of a sentence activates their syntactic function. Other words in the sentence and text activate their relevant meanings.

Studies show that giving children practice reading context-dependent words in sentences facilitates the formation of connections between spellings, syntactic, and semantic identities compared with reading the same words in isolation. For example, Ehri and Wilce (1980) had first graders practice reading ten unfamiliar words such as ‹MIGHT›, ‹WHICH›, ‹ENOUGH›, ‹GAVE›. Either they read the words in sentences, or they read the words in isolation and then heard the sentence contexts. On posttests, those who had read the words in sentences were better able to embed the context-dependent words in complete, meaningful sentences than those who read the words in isolation. Also, the sentence readers were better at detecting which of four context-dependent words was present in a spoken sentence (e.g., "gave" in "The green frog gave vegetables to the hungry rabbit").

According to amalgamation theory (Ehri, 1978, 1980), when children learn to read context-dependent words in text, their syntactic functions and meanings are activated and become bonded to their spellings. To illustrate, when the words ‹FOR›, ‹BY›, ‹TO›, ‹THROUGH›, ‹WOULD› are read in meaningful sentences, syntactic and semantic connections to spellings are formed, with the result that these words are not mistaken for their homophonic counterparts ‹FOUR›, ‹BUY›, ‹TWO›, ‹THREW›, ‹WOOD›. Ehri and Roberts (1979) studied this learning process with beginning readers. In one condition, first graders were taught to read eight homophones in sentence contexts. This was expected to activate the words' syntactic and semantic identities when the words were read. In the other condition, children read the same words in isolation on cards and then heard the sentence contexts containing the words spoken by the experimenter. Because syntactic and semantic identities were not active when the words were read in isolation, it was expected that connections to spellings were less likely to be formed. Some of the word pairs taught were ‹WHICH›-‹WITCH›, ‹CHOOSE›-‹CHEWS›, ‹BALD›-‹BAWLED›. Results on a sentence production posttest showed that the context group was better able to embed the words in appropriate sentences than the isolation group, who mixed up the homophones by placing them in sentences that were appropriate for the other homophone. These findings indicate that having children read homophones in isolation on flash cards is flawed. This practice undermines the formation of connections between spellings and appropriate meanings of the words. It may be especially deleterious for context-dependent words such as ‹WHICH›, because its pronunciation is likely to activate the incorrect meaning, 'witch,' as a result of its semantic salience out of context.

In the two preceding studies comparing the effects of learning to read words in sentence contexts versus in isolation, beginning readers who read words out of context learned more about the spellings of the words than children in the context conditions as shown on posttests following word learning. This suggests that orthographic mapping was more thorough when children read words in isolation. Lacking any context cues to predict words, children had to employ a decoding strategy and attend to all the letters to read the words. Also, they might have looked at isolated words longer than when the words were buried in sentences. Findings of the two studies (Ehri & Roberts, 1979; Ehri & Wilce, 1980) suggest that these two word reading experiences contribute differently to the building of word identities in memory among beginning readers. Reading words in isolation may be better for learning spellings, whereas reading words in context may be better for learning syntactic and semantic identities.

In her developmental model, Chall (1983) considered text reading practice as the most important way to build fluency. According to her, children need to read text that contains familiar concepts and ideas so that they "can concentrate attention on the printed words, usually the most common, high-frequency words" (p. 18) and on matching them to their knowledge of language and the world. In this way, beginners are enabled to connect high-frequency, context-dependent words to their knowledge of language in order to recognize and read them as separate meaningful lexical units.

Building Vocabulary From Written Words

Children learn new vocabulary words not only by hearing them, but also by reading them in text. As they advance through the grades, a greater proportion of their word knowledge comes from reading experiences rather than from exposure to spoken language. Children's books contain more rare words than spoken conversations among adults (i.e., 31 rare words per 1,000 versus 17 rare words per 1,000, respectively) (Cunningham, 2005). One advantage of learning vocabulary words from print is that when spellings of the words are decoded they become bonded to pronunciations and meanings in memory.

Rosenthal and Ehri (2008) conducted a study to determine whether showing students the spellings of new vocabulary words would improve their memory for the words and their meanings. Second and fifth graders were taught the spoken forms and meanings of several unfamiliar, low frequency nouns. Fifth graders were taught ten multisyllabic words, such as ‹VIBRISSA› (the whiskers on a cat) and ‹TAMARACK› (a big tree), whereas second graders were taught six shorter words, such as ‹SOD› (wet, grassy ground) and ‹PAP› (soft mushy food for babies). In each experiment, the nouns were pronounced, defined, embedded in spoken sentences, and depicted in drawings on cards. Children were given several trials with feedback to learn the pronunciations and meanings of the words. In one condition, spellings appeared on the cards during study and feedback periods but not when children recalled the words. In the control condition, the same procedures were followed but students were not shown spellings. To compensate, they pronounced the words extra times. Results were clear in showing the benefit of spellings at both grade levels. Students learned the pronunciations and meanings of the words more quickly over the learning trials when they had seen their spellings. In the fifth-grade experiment, both stronger and weaker readers remembered pronunciations much better when they had been exposed to their spellings, but the advantage was greater for the stronger readers. Others have shown the benefit of exposure to spellings on vocabulary learning as well (Ricketts, Bishop, & Nation, 2009).

Gathercole (2006) has suggested that superior *phonological working memory* for novel words explains why good readers are better at building their vocabularies than poor readers. Phonological working memory refers to the temporary storage of phonological representation of words in memory. However, findings in the preceding study suggest that orthographic knowledge may be more important than phonological memory. The better readers outperformed the weaker readers by very little in remembering pronunciations over the learning trials when they only heard and spoke the novel words but did not see spellings. This indicates only a small difference in phonological memory. However, better readers were far superior to weaker readers in remembering pronunciations when spellings were seen during learning. This suggests that superior ability to connect spellings to pronunciations in memory via orthographic mapping explains why good readers build larger vocabularies than poor readers, not phonological working memory.

Conclusion

Research on learning to read words is extensive. This chapter has touched on some of this research. Discussion has focused on underlying theory and distinctions that are central to understanding development. Children learn to read words in several ways. When spellings are unfamiliar the words might be read by applying strategies, including decoding, analogizing, and prediction. When spellings become familiar, they are read from memory by sight. Orthographic mapping is a key concept that explains how the spellings of words are retained in memory and enable sight word reading. This happens when written units in individual spellings, either graphemes or larger graphosyllabic or graphomorphemic units, are recognized as symbols for units in pronunciations of phonemes, syllables, subsyllables such as rimes, or morphemes. Recognition of these symbolic relations comes from readers' knowledge of the general writing system. When unfamiliar words are seen and pronounced, graphophonological connections are activated. This secures spellings of words to their pronunciations in memory.

Four phases of development portray the emergence of sight word reading, from prealphabetic to partial, full, and consolidated alphabetic phases. Each phase is defined by the predominant type of unit that connects the spellings of words to their pronunciations in memory. Connections evolve from visual and contextual nonalphabetic connections to partial alphabetic letter–sound connections, to more complete graphophonemic connections to connections involving consolidated multiletter-sound spelling patterns. Once readers can perform orthographic mapping, their vocabulary learning can benefit. Seeing the spellings of new words can strengthen their memory for pronunciations and meanings of the words.

Spellings of words also become imbued with syntactic and semantic information. Syntactic identities are connected to spellings when the words are read in sentences and their grammatical roles are processed. This is especially important for words such as prepositions, auxiliary verbs, and conjunctions, whose function is to identify syntactic relations among words. Also semantic identities become amalgamated to spellings during text reading as connections are formed linking the spellings to other words and to text meanings in memory.

These connections reside in memory as potential meanings of spellings, any of which is ready for activation when the spelling appears in a specific context. Readers' spelling vocabularies support oral text reading fluency to the extent that the spellings are recognized automatically and are well connected to other words and their meanings.

Future Directions: Questions to Address

According to phase theory, beginners must learn to connect graphemes in the spellings of words to phonemes in their pronunciations to retain sight words in memory during the full phase. More research is needed to clarify the transition from children knowing individual letters and their names or sounds to using this knowledge to retain sight words in memory effectively. At the outset, much practice is needed to begin forming connections spontaneously when words are read. What forms of training and practice promote this growth? How important is phonemic segmentation with attention to articulation?

According to the view of word learning proposed here, when children learn to read words, the spellings of words become imbued with syntactic functions and multiple meanings. Acquisition is governed by the sentences and passages where the words are read. More and more syntactic and semantic connections accumulate in memory creating potential meanings that become activated when the words appear in particular contexts. More research is needed to explore the course of acquisition in connecting spellings to syntactic and semantic identities of words. What sorts of contexts facilitate this learning and for which types of words? Second-language learners of English have special difficulty learning the proper use of function words. To what extent does reading these words in text facilitate their acquisition?

Knowledge of the letter–sound mapping system has been shown to facilitate memory for the spellings of individual words. However, spelling English words correctly from memory is difficult because the system is variable, with many sources of regularity as well as irregularity. What underlies memory for the correct spellings of English words? How many different knowledge sources are influential? How can spelling instruction be improved? These are questions that deserve further exploration.

References

Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Barr, R. (1974–75). The effect of instruction on pupil reading strategies. *Reading Research Quarterly, 10,* 555–582.

Barron, R. (1981). Development of visual word recognition: A review. In G. Mackinnon & T. Waller (Eds.), *Reading research: Advances in theory and practice* (pp. 119–158). New York, NY: Academic Press.

Bhattacharya, A., & Ehri, L. (2004). Graphosyllabic analysis helps adolescent struggling readers read and spell words. *Journal of Learning Disabilities, 37,* 331–348.

Biemiller, A. (1970). The development of the use of graphic and contextual information as children learn to read. *Reading Research Quarterly, 6,* 75–96.

Boyer, N., & Ehri, L. (2011). Contribution of phonemic segmentation instruction with letters and articulation pictures to word reading and spelling in beginners. *Scientific Studies of Reading, 15*, 440–470.

Campbell, R. (1983). Writing nonwords to dictation. *Brain and language, 19,* 153–178.

Cardoso-Martins, C. (2001). The reading abilities of beginning readers of Brazilian Portuguese: Implications for a theory of reading acquisition. *Scientific Studies of Reading, 5,* 289–317.

Carnine, L., Carnine, D., & Gersten, R. (1984). Analysis of oral reading errors made by economically disadvantaged students taught with a synthetic phonics approach. *Reading Research Quarterly, 19,* 343–356.

Chall, J. (1983). *Stages of reading development*. New York, NY: McGraw Hill.

Clay, M. (1985). *The early detection of reading difficulties*. Auckland, New Zealand: Heinemann.

Cunningham, A. (2005). Vocabulary growth through independent reading and reading aloud to children. In E. Hiebert & M. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 45–68). Mahwah, NJ: Erlbaum.

de Abreu, M., & Cardoso-Martins, C. (1998). Alphabetic access route in beginning reading acquisition in Portuguese: The role of letter-name knowledge. *Reading and Writing: An Interdisciplinary Journal, 10,* 85–104.

Defior, S., Cary, L., & Martos, F. (2002). Differences in reading acquisition development in two shallow orthographies: Portuguese and Spanish. *Applied Psycholinguistics, 23,* 135–148.

Duff, F., & Hulme, C. (2012). The role of children's phonological and semantic knowledge in learning to read words. *Scientific Studies of Reading, 16,* 504–525.

Ehri, L. (1975). Word consciousness in readers and prereaders. *Journal of Educational Psychology, 67,* 204–212.

Ehri, L. (1976). Word learning in beginning readers and prereaders: Effects of form class and defining contexts. *Journal of Educational Psychology, 68,* 832–842.

Ehri, L. (1978). Beginning reading from a psycholinguistic perspective: Amalgamation of word identities. In F. Murray (Ed.), *The recognition of words* (pp. 1–33). Newark, DE: International Reading Association.

Ehri, L. (1980). The development of orthographic images. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 311–338). London, England: Academic Press.

Ehri, L. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading Acquisition* (pp. 107–143). Hillsdale, NJ: Erlbaum.

Ehri, L. (1997). Learning to read and learning to spell are one and the same, almost. In C. Perfetti, L. Rieben, & Fayol, M. (Eds.), *Learning to spell: Research, theory and practice across languages* (pp. 237–269). Mahwah, NJ: Erlbaum.

Ehri, L. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. Metsala & L. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 3–40). Mahwah, NJ: Erlbaum.

Ehri, L. (2005a). Development of sight word reading: Phases and findings. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 135–154). Oxford, England: Blackwell.

Ehri, L. (2005b). Learning to read words: Theory, findings and issues. *Scientific Studies of Reading, 9,* 167–188.

Ehri, L. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading, 18,* 5–21.

Ehri, L. C., Deffner, N. D., & Wilce, L. S. (1984). Pictorial mnemonics for phonics. *Journal of Educational Psychology, 76,* 880–893.

Ehri, L., Nunes, S., Stahl, S., & Willows D. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71,* 393–447.

Ehri, L., & Robbins, C. (1992). Beginners need some decoding skill to read words by analogy. *Reading Research Quarterly, 27,* 12–26.

Ehri, L., & Roberts, K. (1979). Do beginners learn printed words better in contexts or in isolation? *Child Development, 50,* 675–685.

Ehri, L., & Saltmarsh, J. (1995). Beginning readers outperform older disabled readers in learning to read words by sight. *Reading and Writing: An Interdisciplinary Journal, 7,* 295–326.

Ehri, L., Satlow, E., & Gaskins, I. (2009). Grapho-phonemic enrichment strengthens keyword analogy instruction for struggling readers. *Reading and Writing Quarterly, 25,* 162–191.

Ehri, L., & Sweet, J. (1991). Fingerpoint-reading of memorized text: What enables beginners to process the print? *Reading Research Quarterly, 26,* 442–462.

Ehri, L., & Wilce, L. (1979). The mnemonic value of orthography among beginning readers. *Journal of Educational Psychology, 71,* 26–40.

Ehri, L., & Wilce, L. (1980). Do beginners learn to read function words better in sentences or in lists? *Reading Research Quarterly, 15,* 451–476.

Ehri, L., & Wilce, L. (1983). Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology, 75,* 3–18.

Ehri, L., & Wilce, L. (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly, 20,* 163–179.

Ehri, L., & Wilce, L. (1987a). Cipher versus cue reading: An experiment in decoding acquisition. *Journal of Educational Psychology, 79,* 3–13.

Ehri, L., & Wilce, L. (1987b). Does learning to spell help beginners learn to read words? *Reading Research Quarterly, 22,* 47–65.

Elbro, C. (1997). Early linguistic abilities and reading development: A review and a hypothesis about underlying differences in distinctiveness of phonological representations of lexical items. *Reading and Writing: An Interdisciplinary Journal, 8,* 453–485.

Elbro, C., Borstrøm, I., & Petersen, D. (1998). Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly, 33,* 36–60.

Elbro, C., de Jong, P., Houter, D., & Neilsen, A. (2012). From spelling pronunciation to lexical access: A second step in word decoding? *Scientific Studies of Reading, 16,* 341–359.

Frost, S., Sandak, R., Mencl, W., Landi, N., Moore, D., Della Porta, G.,... Pugh, K. (2008). Neurobiological and behavioral studies of skilled and impaired word reading. In E. Grigorenko & A. Naples (Eds.), *Single-word reading: Biological and behavioral perspectives* (pp. 355–376). Mahwah, NJ: Erlbaum.

Gaskins, I., Downer, M., Anderson, R., Cunningham, P., Gaskins, R., & Schommer, M. (1988). A metacognition approach to phonics: Using what you know to decode what you don't know. *Remedial and Special Education, 9,* 36–41.

Gaskins, I. W., Ehri, L. C., Cress, C., O'Hara, C., & Donnelly, K. (1996–97). Procedures for word learning: Making discoveries about words. *The Reading Teacher, 50,* 312–328.

Gathercole, S. (2006). Non-word repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27,* 513–543.

Goodman, K. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist, 6,* 126–135.

Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology, 42,* 73–83.

Gough, P., & Hillinger, M. (1980). Learning to read: An unnatural act. *Bulletin of the Orton Society, 30,* 179–196.

Gough, P., Juel, D., & Griffith, P. (1992). Reading, spelling and the orthographic cipher. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 35–48). Hillsdale, NJ: Erlbaum.

Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology, 42,* 73–83.

Guttentag, R., & Haith, M. (1978). Automatic processing as a function of age and reading ability. *Child Development, 49,* 707–716.

Henderson, L. (1980). Wholistic models of feature analysis in word perception: A critical examination. In P. Kolers (Ed.), *Processing of visible language* (Vol. 2, pp. 207–218). New York, NY: Plenum.

Holden, M., & MacGinitie, W. (1972). Children's conceptions of word boundaries in speech and print. *Journal of Educational Psychology, 63,* 551–557.

Huttenlocher, J. (1964). Children's language: Word-phrase relationship. *Science, 143,* 264–265.

Juel, C. (1983). The development and use of mediated word identification. *Reading Research Quarterly, 18,* 306–327.

Khanna, M., Cortese, M., & Birchwood, K. (2010). Learning new words affects nonword pronunciation in children. *Scientific Studies of Reading, 14,* 407–439.

Kuhn, M., & Stahl, S. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95,* 3–21.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293–323.

Leslie, L., & Calhoon, J. (1995). Factors affecting children's reading of rimes: Reading ability, word frequency, and rime neighborhood size. *Journal of Educational Psychology, 87,* 576–586.

Levin, I., & Ehri, L. (2009). Young children's ability to read and spell their own and classmates' names: The role of letter knowledge. *Scientific Studies of Reading, 13,* 249–273.

Liberman, A. (1992). The relation of speech to reading and writing. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 167–177). Amsterdam, the Netherlands, North-Holland.

Liberman, I., Shankweiler, D., Fisher, F., & Carter, B. (1974). Reading and the awareness of linguistic segments. *Journal of Experimental Child Psychology, 18,* 201–212.

Masonheimer, P., Drum, P., & Ehri, L. (1984). Does environmental print identification lead children into word reading? *Journal of Reading Behavior, 16,* 257–272.

Mayzner, M., & Tresselt, M. (1965). Tables of single-letter and digram frequency counts for various word-length and letter position combinations. *Psychonomic Monograph Supplements, 1,* 13–32.

Moats, L. (2000). *Speech to print: Language essentials for teachers*. Baltimore, MD: Brookes.

Morris, D. (1983). Concept of word and phoneme awareness in the beginning reader. *Research in the Teaching of English, 17,* 359–373.

Nicholson, T. (1991). Do children read words better in context or in lists? A classic study revisited. *Journal of Educational Psychology, 83,* 444–450.

Nunes, T., Bryant, P., & Barros, R. (2012). The development of word recognition and its significance for comprehension and fluency. *Journal of Educational Psychology, 104,* 959–973.

Perfetti, C., & Hart, L. (2002). The lexical quality hypothesis: In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 67–86). Amsterdam, the Netherlands: Benjamins.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18,* 22–37.

Piasta, S., Petscher, Y., & Justice, L. (2012). How many letters should preschoolers in public programs know? The diagnostic efficiency of various preschool letter-naming benchmarks for predicting first-grade literacy achievement. *Journal of Educational Psychology, 104,* 945–958.

Rack, J., Hulme, C., Snowling, M., & Wightman, J. (1994). The role of phonology in young children's learning to read words: The direct mapping hypothesis. *Journal of Experimental Child Psychology, 57,* 42–71.

Rack, J., Snowling, M., & Olson, R. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly, 27,* 28–53.

Reitsma, P. (1983). Printed word learning in beginning readers. *Journal of Experimental Child Psychology, 75,* 321–339.

Ricketts, J., Bishop, D., & Nation, K. (2009). Orthographic facilitation in oral vocabulary acquisition. *Quarterly Journal of Experimental Psychology, 62,* 1948–1966.

Roberts, T. (2003). Effects of alphabet letter instruction on young children's word recognition. *Journal of Educational Psychology, 95,* 41–51.

Rosenthal, J., & Ehri, L. (2008). The mnemonic value of orthography for vocabulary learning. *Journal of Educational Psychology, 100,* 175–191.

Rosenthal, J., & Ehri, L. (2011). Pronouncing new words aloud during the silent reading of text enhances fifth graders' memory for vocabulary words and their spellings. *Reading and Writing: An Interdisciplinary Journal, 24,* 921–950.

Rosinski, R., Golinkoff, R., & Kukish, K. (1975). Automatic semantic processing in a picture-word interference task. *Child Development, 46,* 243–253.

Seymour, P., Aro, M., & Erskine, J. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94,* 143–174.

Shankweiler, D., Crain, S., Katz, L., Fowler, A., Liberman, A., Brady, S., … Shaywitz, B. (1995). Cognitive profiles of reading-disabled children: Comparison of language skills in phonology, morphology and syntax. *Psychological Science, 6,* 149–156.

Share, D. L. (2004a). Knowing letter names and learning letter sounds: A causal connection. *Journal of Experimental Child Psychology, 88,* 213–233.

Share, D. L. (2004b). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology, 87,* 267–298.

Share, D. L. (2008). Orthographic learning, phonological recoding, and self-teaching. In R. Kail (Ed.), *Advances in child development and behavior* (pp. 31–81). New York, NY: Elsevier.

Share, D. L., & Gur, T. (1999). How reading begins: A study of preschoolers' print identification strategies. *Cognition and Instruction, 17,* 177–213.

Share, D. L., Jorm, A. F., Maclean, R., & Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Educational Psychology, 76,* 1309–1324.

Shmidman, A., & Ehri, L. (2010). Embedded picture mnemonics to learn letters. *Scientific Studies of Reading, 14,* 159–182.

Stanovich, K. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York, NY: Guilford.

Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

Stuart, M., & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition, 30,* 139–181.

Thompson, B., Fletcher-Flinn, C., & Cottrell, D. (1999). Learning correspondences between letters and phonemes without explicit instruction. *Applied Psycholinguistics, 20,* 21–50.

Treiman, R., & Broderick, V. (1998). What's in a name? Children's knowledge about the letters in their own names. *Journal of Experimental Child Psychology, 70,* 97–116.

Treiman, R., Goswami, U., & Bruck, M. (1990). Not all nonwords are alike: Implications for reading development and theory. *Memory & Cognition, 18,* 559–567.

Treiman, R., & Rodriguez, K. (1999). Young children use letter names in learning to read words. *Psychological Science, 10,* 334–338.

Tunmer, W., & Chapman, J. (2012). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studies of Reading, 16,* 122–140.

Tunmer, W., & Hoover, W. (1993). Phonological recoding skill and beginning reading. *Reading and Writing: An Interdisciplinary Journal, 5,* 161–179.

Tunmer, W., & Nicholson, T. (2011). The development and teaching of word recognition skill. In M. Kamil, P. Pearson, E. Moje, & P. Afflerbach (Eds.), *Handbook of reading research* (Vol. 4, pp. 405–431). New York, NY: Routledge.

Wimmer, H., & Hummer, P. (1990). How German-speaking first graders read and spell: Doubts on the importance of the logographic stage. *Applied Psycholinguistics, 11,* 349–368.

CHAPTER

21 Children's Spelling Development: Theories and Evidence

S. Hélène Deacon *and* Erin Sparks

Abstract

This chapter reviews empirical findings about children's spelling development, with a focus on alphabetic writing systems. The chapter describes the extent to which research evidence accords with the predictions made by three prominent models of spelling development: phonological, constructivist, and statistical learning. Within this framework, models are evaluated for their ability both to describe children's spelling across development and to explain developmental change by specifying underlying mechanisms. The review offers insight into the current state of our knowledge of children's spelling development, gained through years of empirical research. This work has furthered our understanding of children's developing sensitivity to spelling regularities based on the phonology, morphology, and orthography of words. Yet the review also highlights a clear need for further research in order to clarify points of disagreement between existing models; this pursuit will benefit from spelling research that covers a greater diversity of writing systems.

Key Words: spelling, phonology, morphology, orthography, statistical learning, constructivism

In the midst of a book on reading, we are delighted to write a chapter on spelling, specifically on its development. Spelling is often considered to be more difficult than reading (e.g., Mommers, 1987). One reason, among many, is that spelling requires recall of all of the individual letters in a word in the correct order, rather than recognition based on whole words or parts of words. The study of children's spelling is of interest for this reason but also because it offers us insight into how children think about the way in which language is represented on the page. Contrary to early ideas about learning to spell (e.g., Hillerich, 1977; Horn, 1960), there is much more to spelling than rote memorization.

We organize our review around core models of spelling development. These models are important descriptive and explanatory tools that provide broad brushstrokes of what to expect of typical spelling development and also provide rubrics with which to organize the growing body of experimental research. We evaluate the ability of models of spelling development to achieve two important goals, both of which are demanded of any model of development. First, a model must accurately describe children's behavior at different points in development. In terms of spelling development, this means that it needs to accurately describe children's spelling from their earliest attempts to the near-perfect transcription of complex words. Second, an effective model must go beyond description to provide explanations for what causes change; effectively, it needs to specify developmental mechanisms. We juxtapose the findings of empirical research against the descriptions and predictions of the most prominent models of spelling development in order to gain insight into the current status of our knowledge of children's spelling development.

This review, like most models of spelling development, focuses on what we know about children learning to spell in alphabetic writing systems. Certainly, regardless of the written language being learned, developing readers and spellers are faced with the task of working out how the oral language is represented in written form. In the written form of many languages, phonology is an important aspect of this representation, with considerable variation in how information about sound is encoded. Alphabetic writing systems like English use letters to represent phonemes (the smallest units of sound in a language), whereas syllabic writing systems like Bengali use symbols to represent entire syllables. In morphosyllabic writing systems like Chinese, characters represent units of meaning rather than sound, but most characters contain phonetic radicals that often reflect some aspect of the pronunciation of the word. In written languages that reliably map symbols to sounds (e.g., Finnish, which has one-to-one sound-to-letter correspondences), spelling words according to phonological regularities alone may suffice. However, in many languages (e.g., English) the mapping between written symbols and sounds is imperfect, in part because the writing system reflects nonphonological regularities. For example, English spelling includes morphological regularities, reflecting the smallest units of meaning in language; consider the past tense suffix *-ed*, which is spelled the same way in *jumped* and *played* despite different pronunciations (/t/ and /d/, respectively). Spelling also captures orthographic regularities, which in alphabetic writing systems reflect patterns of allowable letter combinations and the contexts in which they occur; for example, English consonants are more likely to appear as doublets at the end of a word (e.g., *full*) than at the beginning of a word (e.g., **fful*). Determining how these various regularities are represented in one's writing system is a major hurdle in learning to read and spell. The study of children's spellings, especially their errors, provides us with a rare window into their developing thoughts about this connection.

Phonological Perspective

The phonological perspective describes spelling development in terms of children's increasingly sophisticated ability to map sounds to letters, demonstrating that children do more than just memorize spellings. Under this view, the alphabetic principle is central; children's biggest challenge when learning to spell in an alphabetic writing system lies in grasping that letters represent the phonemes in speech. Having gained this insight, children apply it to the task of spelling. Although models under this perspective predict that children eventually use other spelling regularities, appreciating letter–sound correspondences is key, an emphasis that is in line with the phonological perspective's focus on learning to spell in alphabetic writing systems.

The phonological perspective has been advanced in varying forms by several theorists (e.g., Ehri, 1997, this volume; Frith, 1985; Gentry, 1982; Henderson, 1985). Although their accounts of spelling development differ slightly, the substantive common ground lies in the remarkably similar series of phases or stages through which children are predicted to develop (for consistency, we use the term stages throughout). Initially, children's spellings are nonphonological and may be characterized as either logographic, in which familiar words are spelled as unanalyzed wholes (Frith, 1985), or as random strings of letters (Gentry, 1982). Fundamentally, these early spellings are argued to lack a systematic connection to the words' sounds. Children go on to produce phonologically based spellings, first with partial and then with complete phonological representations. In the models' final stages, children are predicted to master spelling regularities that go beyond phonology to include morphological and orthographic patterns, allowing children to successfully spell complex words (Ehri, 1997, this volume; Gentry, 1982). As such, models under the phonological perspective attempt to document the progression of spelling abilities from children's earliest attempts to fully competent spelling. In our review and evaluation, we focus on two models that are representative of the phonological perspective (Ehri, 1997; Gentry, 1982).

According to these models, children initially produce spelling attempts that pay no heed to the alphabetic principle. Phonological theories of spelling development recognize these prephonological spellings; Ehri (1997) does so in her prealphabetic stage and Gentry (1982) in his precommunicative stage. As examples, Gentry (1982) describes a 4-year-old child who spelled the English phrase *welcome home* as ‹SSHIDCA›, and Kessler and colleagues (2013) describe a young child who spelled the Portuguese word *bicicleta* 'bicycle' as ‹ORP›. These examples show an understanding that words should be represented with letters, but critically, the letters chosen are not connected to the words' sounds.

More rigorous support for the existence of prephonological spellings comes from analyses of spelling produced to dictation by Brazilian children in their first year of preschool (average age of 4 years, 3 months; Kessler, Treiman, & Cardoso-Martins, 2013). Children were asked to spell a short list of words. For each of these words, the researchers generated a list of all phonologically plausible spellings—specifically, spellings that represented each phoneme in the word, in the correct order, with a letter or digraph commonly used to spell that sound in Portuguese. Children were given credit for the highest degree of proximity between their own spelling and any one of these phonologically plausible spellings; summing these proximity scores across all items yielded each child's total phonological spelling score. Next, the researchers compared each child's phonological spelling score with chance by running simulations that randomly rearranged the child's own spellings multiple times; this established the likelihood that a child's performance could have been achieved simply as a result of random spellings. Almost half of the children in the sample (~45%) did not produce phonologically based spellings at a rate greater than would be expected by chance (Kessler et al., 2013; see also Pollo, Kessler, & Treiman, 2009). Advocates of the phonological perspective characterize these early, prephonological spellings as "a random stringing together of letters of the alphabet" (Gentry, 1982, p. 193). As we will discuss later, it is not always accurate to say that children's prephonological spellings are entirely random (see Kessler et al., 2013). However, it does seem that many young children produce spellings that lack a systematic relationship to the sounds of the words being spelled.

An important shift occurs when children begin to understand the alphabetic principle. At this point, children realize that written spellings represent spoken words (as was the case in their prephonological spellings) and, critically, that particular letters represent sounds within those words. This marks their transition to the partial alphabetic stage (Ehri, 1997) or the semiphonetic stage (Gentry, 1982). As these names suggest, children's knowledge and application of letter–sound correspondences is burgeoning but not yet complete. Again, Gentry (1982) provides an illustrative example, in which a young child spelled the word *telephone* as ‹TLEFN›. Here the letters plausibly represent many, but not all, of the word's sounds.

Some of children's early phonological spellings draw on their knowledge of letter names. We see this demonstrated in the earlierexample: the letter ‹L›'s name is pronounced /ɛl/, a phoneme sequence that appears in the word *telephone*. The child has spelled that phoneme sequence using only the letter ‹L›, neglecting to include a letter to represent the vowel. Indeed, both naturalistic and experimental spelling research shows that children in many societies do learn letter names at an early age, and are able to use this knowledge in their early attempts at phonologically-based spelling (see Treiman & Kessler, 2003 for review). For example, beginning spellers of English are more likely to include a vowel in their naturalistic spelling attempts when the vowel sounds the same as a letter name (Treiman, 1993). Similarly, when asked to spell a nonword like /vɑr/ (pronounced to rhyme with *jar*), young kindergarteners and first graders frequently produced the spelling ‹VR›, omitting the vowel and using the letter ‹R› to represent the sound sequence /ɑr/. They were less likely to omit the vowel from a word that did not contain a letter name (Treiman, 1994). Similar evidence for the use of letter names in spelling has been found in learners of Hebrew (Levin, Patel, Margalit, & Barad, 2002) and Portuguese (Pollo, Kessler, & Treiman, 2005). These findings corroborate observations from invented spellings that led theorists to suggest that children use letter names in their early, semiphonetic spellings (Gentry, 1982). However, in considering children's use of letter names in spelling, we need to bear in mind growing evidence that they may not do so as widely as predicted by theory (e.g., Pollo et al., 2005; Treiman, 1994). Children's letter-name spellings reflect an emerging, but not yet complete, phonological strategy.

Building on this initial phonological strategy, the phonological perspective argues that children go on to produce spellings that more completely reflect the sounds of a word. Ehri (1997) refers to this as the full alphabetic stage, whereas Gentry (1982) refers to it as the phonetic stage. As an example, at this point in spelling development, the child who previously spelled *telephone* as ‹TLEFN› now produced ‹TALAFON› (Gentry, 1982), representing each phoneme in the word in a phonologically plausible way. This example is illustrative of a large body of research (e.g., Read, 1975) showing that children's invented spellings are largely efforts to represent the phonological features of words. Phonologically based invented spellings have been documented for children learning to spell in many alphabetic writing systems, including English (e.g., Lombardino, Bedford, Fortier, Carter, & Brandi, 1997; Read, 1975), Greek (Porpodas, 2001), Swahili (Alcock & Ngorosho, 2003), and Chinese pinyin (Shen & Bear, 2000). Even spellings that at first glance seem quite unconventional often have a clear phonological basis. For example, children's invented spellings often reflect subtle phonetic distinctions, as when they spell *dragon* as ‹JRAGON›, classifying the first sound as /d͡ʒ/ and spelling it accordingly (e.g., Treiman, 1985). Such spellings reflect children's efforts to map what they know about letters onto what they hear in spoken words. At this stage of development, proponents of the phonological perspective argue that children spell strictly on the basis of sound, overlooking other conventions (e.g., spelling the suffix *–ed* in *jumped* phonetically, as in ‹JUPT›; Gentry, 1982).

In the final stages of models within the phonological perspective, children move from producing spellings that only reflect phonological principles to producing spellings that acknowledge principles beyond phonology. Gentry (1982) notes this progression in his transitional and correct stages, and Ehri (1997) describes it in her consolidated alphabetic stage. It is at this point that the phonological perspective would first expect morphological and orthographic regularities to be reflected in children's spelling. As examples, Gentry (1982) notes that correct stage spellers are able to accurately use word structure, including prefixes and suffixes (reflecting morphological regularities), and silent and doubled consonants (reflecting orthographic regularities). Having reached this final stage, children are seen as skilled spellers.

The phonological perspective has had a profound impact on many aspects of our understanding of spelling development for children learning to spell in alphabetic writing systems. Perhaps its most important contribution is the insight that children learning to spell in these writing systems use their understanding of phonology in their spelling attempts, an idea that has influenced both theory and educational practice. As descriptive accounts of spelling development, these phonologically based models have strengths: as noted earlier, once children grasp that letters map onto sounds in words, their invented spellings are quite often phonologically driven. This gives the approach clear face validity. Turning to the models' explanatory power, learning letter–sound correspondences is a plausible mechanism for children's spelling development, at least for the transition from prealphabetic spelling to partial and full alphabetic spelling. Indeed, longitudinal evidence points to the important role of letter-sound knowledge in young children's spelling development during the early school years: Children's letter-sound knowledge uniquely predicts their subsequent ability to produce phonologically reasonable spellings, and this, in turn, predicts their later conventional spelling accuracy (Caravolas, Hulme, & Snowling, 2001). Furthermore, instruction designed to augment knowledge of letter–sound correspondences seems to promote spelling development: Over the course of their first-grade year, spelling improved at a faster rate among children who received more letter-sound instruction than among those who received less letter-sound instruction (Foorman, Francis, Novy, & Liberman, 1991).

Despite these strengths, models within the phonological perspective focus on somewhat narrow aspects of spelling, limiting their ability to account for the full scope of spelling development. Consider children's earliest, prephonological spelling attempts. From a strictly phonological standpoint, spellings like ‹SSHIDCA› indeed seem to be random (as suggested by phonological models; Gentry, 1982), in that they do not suggest any knowledge of the relationship between letters and sounds. Importantly, though, being nonphonological is not the same thing as being entirely random. As we will review in some detail when discussing the statistical learning perspective, children's spellings reflect many regularities beyond phonology that are present in the written language to which they have been exposed (e.g., Treiman, Kessler, & Bourassa, 2001). Indeed, a key conclusion of the Kessler et al. (2013) study described earlier was that many children's prephonological spellings reflected statistical patterns in the written language; those children were not merely selecting random letters. This point emphasizes a recurrent issue with the phonological perspective: While it rightly emphasizes phonological factors as important to children's spelling development, it tends to overlook other factors.

This somewhat restricted focus is also evident when we consider the phonological perspective's account of later spelling development. For children learning to spell, in many alphabetic writing systems, even mastery of phonological conventions leaves a great deal to be learned; many languages, such as English, have a great number of words for which spellings depend on nonphonological regularities. Ehri and Gentry both acknowledge this in the final stages of their respective models by noting that children begin to apply morphological and orthographic principles. However, in contrast to the careful detail in which these models describe and explain the changing representation of phonology in children's spelling attempts, they are quite vague about children's acquisition of other principles. For example, in later spelling development Ehri (1997) describes a general process of consolidation by which, through exposure to conventional spellings, recurring letter patterns come to be recognized and reproduced as unitized chunks. This consolidation is described as applying to both orthographic patterns (e.g., written rime units) and morphological patterns (e.g., suffixes), without differentiating between the two. Although learning letter–sound correspondences may be the mechanism that drives learning alphabetic spelling in early stages of development, it is unlikely to be entirely responsible for children's learning of morphological or orthographic regularities.

Models within the phonological perspective have made an undeniably important contribution to the study of children's spelling development. However, it seems clear that there is more to children's spelling than just phonology. Other models, which we review in the following section, are better able to characterize the nonrandom nature of children's early spellings and to make clear predictions about children's learning of nonphonological spelling regularities.

Constructivist Approach to Spelling Development

The constructivist approach has a key strength in positing a mechanism that applies to all domains of children's learning and development: that of the construction of ideas. In his classic work, Jean Piaget (e.g., 1950, 1954) argued that children actively construct their knowledge by generating and testing hypotheses. Children construct new, more sophisticated rules as they attempt to reconcile their initially relatively simple rules with their observations of the world. Such learning is characterized by stages, in that children's thinking is dominated by their constructed rule at each stage. Here we review two dominant models that have applied this approach to explaining children's development in learning to spell in alphabetic writing systems.

Ferreiro's Universal Hypotheses

Following the Piagetian tradition, Ferreiro was interested in children's hypotheses about how writing works (Ferreiro, 1978; Ferreiro & Teberosky, 1982). Ferreiro chose the term "writing" to highlight the importance of understanding children's ideas about how writing works in general, rather than their ideas about individual orthographies. Accordingly, Ferreiro predicted that children would generate hypotheses about how writing works that might not necessarily apply to their specific orthography. For example, Ferreiro and Teberosky (1982) suggested that prereaders believe that a word must involve a minimum number of letters and several different letters in order to be readable; these are the minimum quantity and within-word variation hypotheses, respectively. According to Ferreiro and Teberosky, as children confront more words they gain the insight that spelling variation, in either quantity or variety of letters, does not seem to be connected to the meaning of individual words; this leads children to a phonological approach to spelling. They argue that at this point children move onto the syllabic stage, during which they hypothesize that individual letters represent syllables. Evidence of syllabic spellings has been reported for children learning to spell in alphabetic scripts including Spanish, Portuguese, and Italian (Ferreiro & Teberosky, 1982; Nunes Carraher & Rego, 1984; Rego, 1999). This syllabic approach to spelling is phonological in that children's hypothesis about writing is that the letters represent sounds; however, children in this stage do not yet appreciate that letters represent a unit of sound smaller than a syllable in the specific writing system that they are learning. Children finally move to an alphabetic approach to spelling in which they map individual letters onto phonemes as a result of their experience that the number of letters in words does not accord to the number of syllables. According to this approach, children's spelling progresses through a process of testing and rejecting hypotheses.

There is clear value in Ferreiro's constructivist approach. Perhaps most importantly, it encouraged researchers and educators alike to consider the possibility that children's early writings are more than random strings of letters unconnected with later spelling development (Gentry, 1982). This theory has also been considered to be universally applicable to children learning to spell any alphabetic writing system (Ferreiro, Pontecorvo, & Zucchermaglio, 1996).

Despite these clear theoretical strengths, the empirical data to support this theory is somewhat lacking, even when tested with children learning Romance languages, with whom it was initially developed. As examples, recent studies have shown that even children who are demonstrably prephonological spellers do not avoid one- and two-letter spellings as often as predicted by the minimum quantity hypothesis. In fact, the number of one- and two-letter spellings created by Portuguese and English children was remarkably similar to the number of times such spellings occur in texts (~20% of the time; Pollo et al., 2009). There is more evidence for the within-word variation hypothesis; a recent study showed that Portuguese and English children are less likely to use the same two letters in a row to write a word than would be expected by chance, showing that they do tend to avoid repeating the same letters in their spellings (Pollo et al., 2009). However, Portuguese writers showed a greater such tendency than English writers, reflecting the relative frequency of repeated letters in the two languages (Pollo et al., 2009). It seems that the hypotheses proposed by this model are not universal, and they might be influenced by the input to which children are exposed.

In terms of stages applicable to older children, several studies have failed to find evidence for a syllabic stage in English (Kamii, Long, Manning, & Manning, 1990; Pollo et al., 2009) and in Portuguese (Cardoso-Martins, Corrêa, Lemos, & Napoleão, 2006; Pollo et al., 2005, 2009; Treiman, Pollo, Cardoso-Martins, & Kessler, 2013). The lack of such evidence in Portuguese is most surprising, given that a syllabic spelling stage had been previously reported for children learning to spell in this language (Nunes Carraher & Rego, 1984; Rego, 1999). Indeed, children's use of syllables in spelling Portuguese was argued to reflect, at least in part, the syllable-based rhythm in Romance languages (Kamii et al., 1990). For example, the presence of a single consonant following the vowel in Romance languages might highlight the prominence of syllables more than in English, where vowels are often followed by multiple consonants (Blevins, 1995). However, in Pollo et al.'s (2009) study, neither Portuguese nor English children categorized as prephonological spellers used more letters to write two-syllable words than they did to write one-syllable words. An alternative possibility is that young children learning to write Romance languages may use two vowels to represent two-syllable words (e.g., Rego, 1999) because letter names frequently occur in their language (Pollo et al., 2005; see also Cardoso-Martins & Batista, 2003; Treiman & Kessler, 2003). However, a recent empirical test of this hypothesis suggests that this is not the case (Treiman et al., 2013). We address this issue in more detail in our discussion of statistical learning models, and we remember that there appears to be little empirical evidence for an early syllabic stage.

To review, it is not clear that the description of children's writing provided by Ferreiro's model of spelling development holds up to empirical scrutiny. The experimental evidence for the universal hypotheses is not overwhelming, leading one to consider alternative interpretations of the patterns observed in children's early spellings. Further, when we think about the many different writing systems in which children could learn to spell, these hypotheses appear far less universal; it is not clear how they could be applied to children learning to spell in nonalphabetic writing systems, such as Japanese and Chinese. In the face of these descriptive challenges, one is also led to be skeptical of the mechanism put forward. Nevertheless, as we will see, the notion of children testing hypotheses can be applied in other ways to describe and explain children's spelling development.

Nunes and Bryant's Stage Model

In contrast to Ferreiro's focus on children's early understanding of writing, Nunes and Bryant's model (e.g., Bryant & Nunes, 1998; Bryant, Nunes, & Aidinis, 1999; Nunes, Bryant, & Bindman, 1997a, 1997b) focuses on children's learning about spelling "beyond the first steps" (e.g., Nunes & Bryant, 2009, p. 1). This model makes predictions about how children learn about regularities beyond phonology, and it has been tested most explicitly in terms of children's spelling of morphemes.

Applying Piaget's model of knowledge construction to children's spelling development in alphabetic writing systems, Nunes and Bryant (2009; Nunes et al., 1997a, 1997b) argued that children's spelling follows stages. Following an initial period of nonsystematic spellings, children appear to have the hypothesis that spelling represents the sounds in words. After observing the spellings of words through their own reading and spelling attempts, children learning to spell languages such as English and French realize that the spellings of a large number of words deviate from this phonological rule. Consequently, children experiment with exceptions to their rule, specifically by using alternative spellings for sounds. This active experimentation with new spellings leads children to discover that there is a basis for many exception spellings: that spellings are based on morphemes as well as phonemes. According to Nunes and Bryant, children construct this new, and quite advanced, rule through their own reading and writing, rather than as a result of explicit teaching.

Support for this model comes from a longitudinal study of children's spelling of the past-tense inflection in English (Nunes et al., 1997b). The rule distinguishing regular and irregular past-tense verbs in English is relatively simple. If the sound is the same in the root and past-tense form of the verb (as it is in *walk* and *walked*), then the verb is regular and the ending must be spelled with *-ed*. If the sound is different (as it is in *keep* and *kept*), then the verb is classified as irregular and the ending is spelled phonetically (as *-t* or *-d*). Given that this rule is not taught in schools, if children learn about it they are likely to have constructed this knowledge themselves.

Nunes and Bryant (Nunes et al., 1997b) tracked the spelling of past-tense verbs (and control words) by a group of 6-, 7-, and 8-year-old children over the course of almost two years. The authors were able to categorize the majority of children's spellings into five stages. The first stage is nonphonetic, in which children did not appear to spell the endings of words systematically. This stage is characterized by omissions and nonphonetic spellings. The second stage is phonetic, with sound-based spellings of the endings (e.g., *-t* for /t/ and *-d* for /d/). In the third stage, children again spelled word endings inconsistently: this time sometimes with *-ed* and sometimes phonetically. Critically, children used *-ed* both correctly (e.g., ‹kissed›) and incorrectly (e.g., ‹feled› and ‹sofed›). Nunes and Bryant argued that children adjusted their rule to accommodate the exception spelling *-ed* but they did not yet understand the basis of this alternative spelling. In the fourth stage, children restricted *-ed* to the ends of verbs but used it for both regular and irregular ones. At this point, Nunes and Bryant argued that children understood that *-ed* is used with past-tense verbs, but did not distinguish between regular and irregular verbs. In the fifth stage, children spelled the endings of both regular and irregular past tense verbs and nonverbs accurately (see also Nunes et al., 1997a). Nunes and colleagues argued that children at this final stage have a morphological rule to determine the basis for the spelling of *-ed*.

Support for Nunes and Bryant's model comes from several aspects of their data. First, 90% of the children could be categorized into stages in the model. Second, the children in the more advanced stages were older and more advanced readers than those in the lower stages. Third, most children moved forward, and not backward, in the stages across the longitudinal study. Accordingly, this model of development captures the progress of the majority of children in mastering the past tense suffix.

Perhaps most compellingly, additional support for Nunes and Bryant's model comes from studies of children learning to spell other languages. All of these studies investigated the effects of morphology, suffixes in particular, on children's spellings of word endings. A similar pattern of development has emerged in studies of children's spelling in Greek (Bryant, Nunes, & Aidinis, 1999; Chliounaki & Bryant, 2007), French (Fayol, Thenevin, Jarousse, & Totereau, 1999; Totereau, Thenevin, & Fayol, 1997), and Dutch (Notenboom & Reitsma, 2007). In each case, albeit with different designs, these studies find that younger children appear to spell predominantly phonologically, older children use alternative spellings, and children who are older still represent suffixes with reasonable accuracy. Nunes and Bryant interpret this pattern of results as reflecting children's acquisition of a morphological rule, at least in these studies of suffixes.

Chliounaki and Bryant's (2007) recent empirical study provides the most direct test of Nunes and Bryant's hypothesis that experience in reading and spelling drives their construction of a rule. Chliounaki and Bryant evaluated whether children's success at spelling inflections in real words drives their ability to spell inflections in pseudowords. They examined children's spelling of word-final vowels in Greek for which there is a morphological basis. For example, word-final /o/ is spelled as ‹ο› if the word is a masculine or neuter gender noun or adjective (as in ‹νερο› water) and as ‹ω› if it is a first-person singular present verb (as in ‹γραω› 'I write'; Harris & Giannouli, 1999). Thus, identifying morphemes in words can allow one to choose between phonologically plausible alternatives. Across two time points in their longitudinal study, Chliounaki and Bryant (2007) found that children's ability to spell inflections correctly in real words predicted their later ability to spell inflections in pseudowords beyond their earlier ability to spell inflections. This relationship did not emerge in the other direction, nor did it emerge with children's spelling of other word stems. The authors suggest that "specific learning, thus, may give children the knowledge that they need to construct morphological spelling rules" (p. 1370).

The constructivist approach, as envisioned by Nunes and Bryant, provides a comprehensive description of children's spellings as they move from predominantly phonological spelling to the accurate representation of complex spelling features, such as suffixes (e.g., Nunes et al., 1997b). This is one clear advantage of the model: its ability to explain children's spelling development "beyond the first steps" (Nunes & Bryant, 2009, p. 1).

There is also beginning evidence for the mechanism by which children master these morphological and orthographic spelling features (Chliounaki & Bryant, 2007), demonstrating that the strengths of the model extend beyond description to prediction. Further, it is compelling that this mechanism is domain-universal, one that children use to construct their knowledge about the world in general (e.g., Piaget, 1950, 1954).

However, some aspects of children's spellings do not concur with the predictions of this model. Consider, for example, Nunes, Bryant, and Bindman's (1997b) comprehensive study of children's spelling of past-tense verbs. The authors argue that there are several key tests of a stage theory. One of these is that it should be possible to classify the vast majority of children into one stage at each testing point. This appears to be the case; the spellings of 90% of the children accorded with predictions of the model, with the spellings of only 10% not doing so. However, Nunes, Bryant, and Bindman also note that a key test of a stage model is that children should move forward and not backward in the stages over time. A significant number of children moved backward in the stages between the testing sessions; as an example, over a quarter of children in the upper two stages moved backward between testing sessions. Certainly, this pattern could be attributable to random error or inconsistency in children's spellings. However, based on Nunes, Bryant and Bindman's own criteria, the existence of a substantial number of backsliders casts some doubt on stage-based learning.

As we will see in detail in the next section, data from other researchers suggest that young children's spellings are influenced by more than what is strictly predicted by the model of Nunes and Bryant. Children's very early spellings cannot be considered to be random (e.g., Treiman et al. 2001). Furthermore, once children's spellings do become phonological in nature, they are not strictly so; there is widespread evidence that young children's spellings reflect both morphological and orthographic regularities (e.g., Treiman & Kessler, 2006). Recently, Nunes and Bryant (2009) have argued that spelling in each stage of their model might be best described as a bias; for example, young children have many strategies available to them but might initially be biased toward the use of a phonological strategy. Such an argument is consistent with both the evidence that young children's spelling of suffixes is predominantly phonological (e.g., Gentry, 1982) and the evidence that young children's spellings are also influenced by orthographic and morphological features (Treiman & Cassar, 1996). However, such an interpretation of stages is somewhat at odds with classic interpretations, which focus on qualitative differences, rather than biases, as distinguishing children's thinking in different stages.

Finally, Nunes and Bryant's constructivist model predicts that children construct a rule, and yet there is conflicting evidence on this front. These and other researchers have offered good evidence that children achieve accurate spelling of morphemes that can be described with a rule (such as regular past-tense verbs are spelled with *-ed*; e.g., Bryant, Nunes, & Snaith, 2000). However, as we will see, even in these cases there is clear evidence that children's and adults' spellings are influenced by features of the surrounding context (e.g., Kemp & Bryant, 2003). Of course, children could rely on rules specific to each phonological or orthographic context that they encounter, but this would not reflect the acquisition of abstract rules considered to be the cornerstone of traditional conceptualizations of rules. These approaches suggest that rules should be abstract such that they should operate independently of surface features (e.g., Anderson, 1993). Together, these studies lead us to question the developmental mechanism put forward in this model.

Statistical Learning

Broadly, statistical learning refers to learning about the frequency with which features occur and co-occur. Saffran and her collaborators coined this term in describing their seminal studies of lexicon formation (1996), in which 8-month-old infants used the likelihood of syllable co-occurrence to extract "words" from an artificial language. In less than two decades, the statistical learning approach has transformed the study of child development. Its appeal lies in part in its breadth; it offers a single mechanism that can be applied across a diverse range of domains, such as learning about auditory (Saffran et al., 1999; Saffran et al., 2005), visual (Fiser & Aslin, 2001, 2002), and tactile (Conway & Christiansen, 2005) stimuli (see Perruchet & Pacton, 2006, for a review). Regardless of the writing system being learned, the statistical learning approach would predict that children's spellings should reflect the regularities present in the data available to them, possibly leading to sensitivities to a variety of information at the same time. We review here data on the application of this approach to explaining development in children learning to spell in alphabetic writing systems (Pollo, Treiman, & Kessler, 2007).

An early pattern to which young children are exposed repeatedly is their own name. The impact of this increased frequency of exposure is reflected in the success with which children identify letters that occur in their own first name; this pattern has emerged in several different languages, including English (Treiman & Broderick, 1998), Hebrew (Levin & Aram, 2004), and Portuguese (Pollo et al., 2009). And children's early spellings, which might otherwise appear to be random (Gentry, 1982), appear to include disproportionately more of the letters from their own names (Bloodgood, 1999; Treiman et al., 2001).

As children are exposed to an increasing number of letters and words (e.g., Robins, Treiman, & Rosales, 2014; Roy-Charland, Saint-Aubin, & Evans, 2007), other patterns emerge in their spellings. Recent evidence demonstrates that the influence of letter names on children's spelling (e.g., Levin et al., 2002; Treiman, 1994), long cited in support of the phonological perspective, is influenced by the frequency with which letter names appear in words and the relative frequency of individual letters. As an example, words in certain languages such as Portuguese are more likely to contain letter names than words in other languages such as English. Words in Portuguese also contain more vowels than those in English. In Pollo et al.'s study (2005), 4- to six-year-old English- and Portuguese-speaking children were asked to spell words, all of which had two vowels. Some of the words were pronounced with one vowel letter name and others with two (e.g., /i/ in *bunny* versus /o/ and /i/ in *pony*). Children in both languages used more vowels and were more phonologically accurate in spelling words containing two letter names than those containing one letter name. This demonstrates that children learning to spell both languages use a letter-name strategy, and yet there is also an influence of the statistics present in the learning situation. Portuguese-speaking children were more likely than English-speaking children to use vowels in their spellings, even though the words that the children spelled had the same structure. This difference may reflect the higher rate of occurrence of vowel letter names in words in Portuguese in comparison with English (Pollo et al., 2005). Patterns present in the input appear to influence the spellings of young children.

Children also seem to be sensitive to the frequency with which letters co-occur. As an example, even very young writers of both English and French are sensitive to which letters can occur as doublets and which cannot. Cassar and Treiman (1997) showed that 6-year-old English-speaking children were more likely to choose nonwords as word-like when these nonwords contained allowable doublets over nonallowable doublets, both for vowels (e.g., *heek* vs. *haak*) and consonants (e.g., *yill* vs. *yihh*). Pacton and colleagues reported a similar pattern of findings with French-speaking children (Pacton, Perruchet, Fayol, & Cleeremans, 2001). Similarly, young children's spellings of vowels are influenced by the surrounding consonants, just as their spellings of consonants are influenced by the adjacent vowels (Hayes, Treiman, & Kessler, 2006; Treiman & Kessler, 2006). These patterns could suggest that children are using context-sensitive rules. Regardless of the interpretation, such early sensitivity to orthographic patterns contrasts directly with the predictions of both phonological and constructivist models that predict that children's early spellings should be dominated by representation of phonology (e.g., Ehri, 1997; Frith, 1985).

Pacton and colleagues (2001) provided perhaps the most direct test of whether sensitivity to orthographic features can be explained by children's extraction of abstract, general rules, as suggested by the constructivist approach (Bryant, Nunes, & Snaith, 2000). In French, one could describe consonant doublet regularities with a rule: Doubled consonants only occur in word-medial positions. And indeed, children were more likely to suggest that pseudowords were word-like when they included doublets in the medial position than in the beginning or end positions (e.g., *nullor* versus *nnulor*). Critically, though, children's performance was affected by the frequency with which specific letters are doubled, a factor that should not impact performance if children rely exclusively on a rule (e.g., Smith, Langston, & Nisbett, 1992). Children were more likely to choose pseudowords as word-like if they contained frequently doubled consonants than never doubled consonants (e.g., *tummet* versus *tukket*). Although it is possible that the children might have extracted a rule specifying which consonants double, such a rule would be relatively limited in its scope. Further, it would not adhere to traditional criteria of rule-based learning, one of which is that performance should not be affected by frequency (e.g., Smith et al., 1992). It seems then that children's spellings of orthographic features do not reflect the extraction of a rule, at least not one that applies across a wide range of contexts.

The ability of the statistical learning approach to explain how children learn morphological features of spelling is perhaps its most contentious; it is in this domain that theorists argue that children should rely on rules (e.g., Bryant et al., 2000). This rule-based learning is based in part on classic models of morpheme processing treating them as discrete units (Taft & Forster, 1975). In contrast, newer conceptualizations of morphology suggest that morphemes can be described in graded terms; they reflect a statistical co-occurrence of sound, letters, and meaning, or between phonology, orthography, and semantics, respectively (Seidenberg & Gonnerman, 2000). For example, the words *teach* and *teacher* are likely to appear close together in text and in texts with similar meanings (according to approaches such as latent semantic analysis; Landauer, Laham, & Foltz, 1998), and the overlap of the initial letters ‹teach› and the sound /tiːt͡ʃ/ along with the meaning 'instruct' could lead to the creation of a strong association between the forms *teach* and *teacher*. Similarly, children might see *-er* and frequently hear its pronunciation in combination with the meaning 'someone who does something.' These co-occurrences could lead children to use consistent spellings for these units (*teach* and *-er*). Accordingly, statistical learning might account for learning of these morphological regularities.

One line of evidence supporting the possibility that statistical learning accounts for learning of morphological regularities comes from findings of morphological effects in the spellings of young children. Across several studies, the spellings of children as young as 5 and 6 years of age appear to be influenced by words' morphological structures (e.g., Turnbull, Deacon, & Kay-Raining Bird, 2011); children are more accurate in their spelling of parts of morphemes (Kemp, 2006; Treiman, 1993; Treiman & Cassar, 1996; Treiman, Cassar, & Zukowski, 1994; Treiman & Cassar, 1996; Walker & Hauerwas, 2006; see also Byrne, 1996; Levin & Korat, 1993) and whole morphemes (Deacon & Bryant, 2006) than comparable letter patterns that are not morphemes. And a recent study points to impacts of both orthographic and semantic frequency on the emergence of these morphological effects in young children's spellings. Deacon and Leung (2013) found that young children were more accurate in choosing correct spellings for word-endings that were morphemes than for those that were not specifically when these endings were the frequent *-er* spelling (e.g., *painter* vs. *corner*), but not the less frequent *-or* spelling (e.g., *actor* vs. *alligator*). Further, these morphemic effects emerged for *–er* in derived forms for the 6- and 7-year-old children (e.g., *painter* versus *corner*), but in inflected forms only with the 8- to 9-year-old children (e.g., *shorter* versus *corner*). We attributed this difference in the timing of these morphemic effects' emergence to semantic frequency; there are far more derived forms than inflected forms in English, based on grade-level counts in children's books (Zeno, 1995). It seems that even when children's spellings are argued to be primarily phonological in nature (Gentry, 1982), the spellings reflect morphological regularities in a manner that reflects frequency of exposure to these forms.

A second line of evidence that statistical learning can account for learning of morphological regularities comes from findings that even adult spellers do not rely on simple abstract rules. As we noted earlier, traditional conceptualizations of rules entail that they operate independently of surface features, such as the surrounding letter and sound contexts (e.g., Anderson, 1993). If children did rely on such rules, then they would simply spell pseudowords presented as plural with an *-s*. They do not appear to do so. Kemp and Bryant (2003) showed that child and adult spellers were twice as likely to use *-s* to spell pseudowords presented as plurals (e.g., "those smees") when the penultimate sound was a consonant than when it was a long vowel (e.g., *preens* vs. *smees*; see Pacton, Fayol, & Perruchet, 2005, for similar French data). This large variation in performance based on the adjacent sound or letter shows that spellers do not rely on an abstract rule, such as plurals are spelled with *-s*. Their spellings are, instead, consistent with the input to which they are exposed; *-s* follows consonants far more frequently than long vowels in English. Certainly there was some sensitivity to the fact that plurals are spelled with *-s*; university-educated adults and better-spelling children were more likely to use *-s* when the long-vowel pseudowords were presented as plurals than as nonplurals (e.g., "those smees" vs. "that smeese"; see also Mitchell, Kemp, & Bryant, 2011). However, the effects of the adjacent sound or letter on spelling choices show that this knowledge does not always operate in a manner that adheres to traditional conceptualizations of rules, at least in the paradigm employed in these studies.

Further evidence that surface features affect spelling accuracy comes from studies of both French and Dutch spellers. Contrary to the predictions of traditional conceptualizations of rule-based learning, these studies demonstrate effects of the presence of a homophonous form and of the surrounding semantic and syntactic context (e.g., Hupet, Fayol, & Schelstraete, 1998; Largy, Fayol, & Lemaire, 1996; Notenboom & Reitsma, 2007; Pacton, 2004; Sandra, 2010; Sandra, Frisson, & Daems, 1999). As an example, adults and children are more likely to misspell words that have a homophonous form (e.g., *timbre* 'stamp' can be both a noun and verb, and so has two plural forms: *timbres* and *timbrent*) than those that do not (e.g., *nuage* 'cloud' is only a noun, and so has only one plural form, *nuages*; Largy et al., 1996). Further, the rate with which these errors occur is influenced by the words that precede the target (Pacton, 2004). As with the effects of the adjacent sound, the influence of features of the surrounding context demonstrates that even adults' spellings do not appear to reflect the application of an abstract rule that is independent of surface features.

We have reviewed empirical evidence demonstrating that the frequency of exposure to features in the written language impacts children's spelling. Careful experimental design has brought to light findings of nonrandomness in very young writers' spellings (e.g., Pollo et al., 2009) as well as orthographic and morphological effects in the spellings of children (e.g., Deacon & Leung, 2013) whose spelling would be considered by some to be phonologically based (e.g., Gentry, 1982). Further, there is evidence of sensitivity to orthographic features, even when a rule can be applied and even by mature writers (Kemp & Bryant, 2003). Across all of these findings, the statistical learning approach claims that children's and adults' spellings reflect the patterns in the print to which they are exposed. There is appeal in the single mechanism put forward by statistical learning, applicable across both domains and periods of development. Accordingly, statistical learning offers a parsimonious explanation; in the domain of spelling, statistical learning approaches put forward that children's spellings reflect regularities present in the input (e.g., Perruchet & Pacton, 2006).

As with any approach, there are challenges to explanations offered by statistical learning approaches. First, one cannot dismiss the evidence that children's spellings do appear, at face value, to be initially largely random and later phonologically driven (e.g., Gentry, 1982; Read, 1986), as predicted by both the phonological models and Nunes and Bryant's model. This criticism does not discount the subtle effects of morphological and orthographic features that have emerged (e.g., Treiman & Cassar, 1996), and yet there is a real-world validity to these descriptions of children's early spellings. Second, perhaps the most daunting challenge lies in outlining clear implications for educators. New research begins to outline educational applications; a recent study showed that children are more likely to later develop spelling difficulties if their early spellings accord less rather than more closely with the regularities in the print to which they are exposed (Kessler et al., 2013). Attention to the randomness (or conversely the regularities) in young children's spellings might then be an indicator of very early progress in learning to spell. And yet beyond this there are few explicit applications of this approach to the educational world, despite clear evidence that children have a great deal to learn (e.g., Nunes et al., 1997b; see also Bryant, Devine, Ledward, & Nunes, 1997). Statistical learning relies on implicit learning, and it is not yet clear how this might be augmented with explicit teaching or even whether it is appropriate to do so. A final challenge for the statistical learning model comes from findings that children's metalinguistic skills, such as morphological awareness, predict their acquisition of the spellings of specific morphological features (e.g., Nunes et al., 1997b; but see da Mota, 1996). These findings are hard to reconcile with predictions that the primary determinant of performance should be frequency of exposure. Educational implications and the impacts of metalinguistic skills are two areas in which greater clarity is required for the statistical learning approach.

Conclusion

Taken together, the models put forward to date each have strengths and weaknesses in their ability to describe and explain the spelling development of children learning to spell in alphabetic writing systems. The phonological perspective (Ehri, 1997; Gentry, 1982) highlights the importance of learning letter-sound regularities at early stages of spelling development. Ferreiro's (e.g., Ferreiro & Teberosky, 1982) and Nunes and Bryant's (Nunes et al., 1997a) constructivist models each provide more detailed descriptions of children's spelling prior to and following this phonological stage, respectively. The statistical learning approach brings a mechanistic explanation to the recently emerging evidence of children's sensitivity to regularities in their writing system at ages at which their spellings are thought to be either random or phonological. The latter two sets of models suggest domain-general mechanisms that could underlie many aspects of development, including spelling.

And yet, exploring the challenges of each model points to gaps that remain in our current understanding of spelling development. Perhaps the most glaring issue lies in the relatively limited scope of the dominant models of spelling development; all focus on children learning to spell in alphabetic writing systems in which letters are mapped onto phonemes, albeit with variability in the reliability of this mapping. Accordingly, there is a lack of specificity (and, in many cases, clear omission) in the models' description and prediction of spelling development for children learning to spell in languages that use other types of writing systems. Some of these models can be more easily applied across writing systems than others. For example, the statistical learning approach's predictions regarding the impact of frequently occurring features on children's spelling development should apply regardless of the nature of these features. Supporting evidence comes from findings on the frequency of exposure to the spelling of one's own name; like children learning to write in alphabetic writing systems, 4-year-old Mandarin speaking children are more accurate in spelling the characters in their own name compared with other characters (Yin & Treiman, 2013; but see also Treiman & Yin, 2011).

The extent to which other models can be applied to children learning to spell in nonalphabetic writing systems is less clear. For example, the primary mapping in Chinese writing is at the level of the morpheme, rather than the phoneme (Shu, Chen, Anderson, Wu, & Xuan, 2003); as such, predictions of a phonological stage or phase would probably miss the mark for children learning to spell in Chinese. Indeed, that is the case in some of the first evidence emerging on this front. A recent study found that the dominant error type for young children learning to spell in Chinese was the substitution of a synonymous character or the omission of a character altogether (70% to 90% of all errors; Tong, McBride-Chang, Shu, & Wong, 2009); such errors are thought to occur at the morphemic level. Phonological errors, such as homophonous spellings of characters, were relatively rare (3% to 4% of all errors; see also Shen & Bear, 2000). These few examples highlight the importance of taking cross-linguistic diversity seriously by continuing to test models of spelling development across languages (see Caravolas & Samara, this volume; Ho, Yau, & Au, 2003).

We began this review by noting the many insights to be gained by studying children's spelling development, and we end by encouraging further empirical study. Not only will such inquiry advance our understanding of spelling development but also it will be necessary to clarify points of disagreement between models. This pursuit will particularly benefit from research across languages, including those that do not use an alphabetic writing system (e.g., Chinese), in order to build on the empirical foundations discussed here and provide key tests of the generalizability of models of spelling development.

References

Alcock, K. J., & Ngorosho, D. (2003). Learning to spell a regularly spelled language is not a trivial task—patterns of errors in Kiswahili. *Reading and Writing: An Interdisciplinary Journal, 16,* 635–666.

Anderson, J. R. (1993). *Rules of mind.* Hillsdale, NJ: Erlbaum.

Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 206–244). Cambridge, MA: Blackwell.

Bloodgood, J. W. (1999). What's in a name? Children's name writing and literacy acquisition. *Reading Research Quarterly, 34,* 342–367.

Bryant, P., Devine, M., Ledward, A., & Nunes, T. (1997). Spelling with apostrophes and understanding possession. *British Journal of Educational Psychology, 67,* 91–110.

Bryant, P., & Nunes, T. (1998). Learning about the orthography: A cross-linguistic approach. In S. G. Paris & H. M. Wellman (Eds.), *Global perspectives for education: Development, culture and schooling* (pp. 171–192). Washington, DC: American Psychological Association.

Bryant, P., Nunes, T., & Aidinis, A. (1999). Different morphemes, same spelling problems: Cross-linguistic developmental studies. In M. Harris, & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 112–133). Cambridge, England: Cambridge University Press.

Bryant, P., Nunes, T., & Snaith, R. (2000). Children learn an untaught rule of spelling. *Nature, 403,* 157–158.

Byrne, B. (1996). The learnability of the alphabetic principle: Children's initial hypotheses about how print represents spoken language. *Applied Psycholinguistics, 17,* 401–426.

Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language, 45,* 751–774.

Cardoso-Martins, C., & Batista, A. C. E. (2003, April). *The role of letter name knowledge in learning to connect print to speech: Evidence from Brazilian Portuguese-speaking children*. Paper presented at the meeting of the Society for Research in Child Development, Tampa, FL.

Cardoso-Martins, C., Corrêa, M. F., Lemos, L. S., & Napoleão, R. F. (2006). Is there a syllabic stage in spelling development? Evidence from Portuguese-speaking children. *Journal of Educational Psychology, 98,* 628–641.

Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology, 89,* 631–644.

Chliounaki, K., & Bryant, P. (2007). How children learn about morphological spelling rules. *Child Development, 78,* 1360–1373.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 24–39.

da Mota, L. (1996). *The role of grammatical knowledge in spelling.* Unpublished doctoral dissertation, University of Oxford, Oxford, UK.

Deacon, S. H., & Bryant, P. (2006). Getting to the root: Young writers' sensitivity to the role of root morphemes in the spelling of inflected and derived words. *Journal of Child Language, 33,* 401–417.

Deacon, S. H., & Leung, D. (2013). Testing the statistical learning of spelling patterns by manipulating semantic and orthographic frequency. *Applied Psycholinguistics, 34,* 1093–1108.

Ehri, L. (1992). Review and commentary: Stages of spelling development. In S. Templeton & D. R. Bear (Eds.), *Development of orthographic knowledge and the foundations of literacy: A memorial festschrift for Edmund H. Henderson* (pp. 307–332). Hillsdale, NJ: Erlbaum.

Ehri, L. (1997). Learning to read and learning to spell are one and the same, almost. In C. A. Perfetti, L. Reiben, & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 237–269). Mahwah, NJ: Erlbaum.

Fayol, M., Thenevin, M. G., Jarousse, J. P., & Totereau, C. (1999). From teaching to learning French written morphology. In T. Nunes (Ed.), *Learning to read: An integrated view from research and practice* (pp. 43–64). Dordrecht, the Netherlands: Kluwer.

Ferreiro, E. (1978). What is written in a written sentence? A developmental answer. *Journal of Education*, *160,* 25–39.

Ferreiro, E., Pontecorvo, C., & Zucchermaglio, C. (1996). *Pizza* or *piza*? How children interpret the doubling of letters in writing. In C. Pontecorvo, M. Orsolini, B. Burge, & L. Resnick (Eds.), *Children's early text construction* (pp. 145–163). Mahwah, NJ: Erlbaum.

Ferreiro, E., & Teberosky, A. (1982). *Literacy before schooling*. Portsmouth, NH: Heinemann.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12,* 499–504.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28,* 438–467.

Foorman, B. R., Francis, D. J., Novy, D. M., & Liberman, D. (1991). How letter-sound instruction mediates progress in first-grade reading and spelling. *Journal of Educational Psychology, 83,* 456–469.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, M. Coltheart, & J. Marshall (Eds.), *Surface dyslexia* (pp. 301–330). London, England: Erlbaum.

Gentry, J. R. (1982). An analysis of developmental spelling in GNYS AT WRK. *The Reading Teacher, 36,* 339–363.

Harris, M., & Giannouli, V. (1999). Learning to read and spell in Greek. In M. Harris, & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 51–70). Cambridge, England: Cambridge University Press.

Hayes, H., Treiman, R., & Kessler, B. (2006). Children use vowels to help them spell consonants. *Journal of Experimental Child Psychology*, *94,* 27–42.

Henderson, E. (1985). *Teaching spelling*. Boston, MA: Houghton Mifflin.

Hillerich, R. L. (1977). Let's teach spelling—not phonetic misspelling. *Language Arts, 54,* 301–307.

Ho, C. S.-H., Yau, P. W.-Y., & Au, A. (2003). Development of orthographic knowledge and its relationship with reading and spelling among Chinese kindergarten and primary school children. In C. McBride-Chang & H.-C. Chen (Eds.). *Reading development in Chinese children* (pp. 51–71). London, England: Praeger.

Horn, E. (1960). Spelling. In C. W. Harris (Ed.), *Encyclopedia of educational research* (3rd ed., pp. 1337–1354). New York, NY: Macmillan.

Hupet, M., Fayol, M., & Schelstraete, M. (1998). Effects of semantic variables on the subject-verb agreement processes in writing. *British Journal of Psychology, 89,* 59–75.

Kamii, C., Long, R., Manning, M., & Manning, G. (1990). Spelling in kindergarten: A constructivist analysis comparing Spanish-speaking and English-speaking children. *Journal of Research in Childhood Education, 4,* 91–97.

Kemp, N. (2006). Children's spelling of base, inflected, and derived words: Links with morphological awareness. *Reading and Writing: An Interdisciplinary Journal, 19,* 737–765.

Kemp, N., & Bryant, P. (2003). Do beez buzz? Rule-based and frequency-based knowledge in learning to spell plural –s. *Child Development, 74,* 63–74.

Kessler, B., Pollo, T. C., Treiman, R., & Cardoso-Martins, C. (2013). Frequency analyses of prephonological spellings as predictors of later success in conventional spelling. *Journal of Learning Disabilities, 46,* 252–259.

Landauer, T., Laham, D., & Foltz, P. (1998). *Latent semantic analysis*. http://lsa.colorado.edu.

Largy, P., Fayol, M., & Lemaire, P. (1996). The homophone effect in written French: The case of verb-noun inflection errors. *Language and Cognitive Processes, 11,* 217–255.

Levin, I., & Aram, D. (2004). Children's names contribute to early literacy: A linguistic and a social perspective. In D. Ravid & H. Bat-Zeev Shyldkrot (Eds.), *Perspectives on language and language development* (pp. 219–239). Dordrecht, the Netherlands: Kluwer.

Levin, I., & Korat, O. (1993). Sensitivity to phonological, morphological, and semantic cues in early reading and writing in Hebrew. *Merrill-Palmer Quarterly, 39,* 213–232.

Levin, I., Patel, S., Margalit, T., & Barad, N. (2002). Letter names: Effect on letter saying, spelling, and word recognition in Hebrew. *Applied Psycholinguistics, 23,* 269–300.

Lombardino, L. J., Bedford, T., Fortier, C., Carter, J., & Brandi, J. (1997). Invented spelling: Developmental patterns in kindergarten children and guidelines for early literacy intervention. *Language, Speech, and Hearing Services in Schools, 28,* 333–343.

Mitchell, P., Kemp, N., & Bryant, P. (2011). Variations among adults in their use of morphemic spelling rules and word-specific knowledge when spelling. *Reading Research Quarterly, 46,* 119–133.

Mommers, M. J. C. (1987). An investigation into the relation between word recognition skills, reading comprehension and spelling skills in the first two years of primary school. *Journal of Research in Reading, 10,* 122–143.

Notenboom, A., & Reitsma, P. (2007). Spelling Dutch doublets: Children's learning of a phonological and morphological spelling rule. *Scientific Studies of Reading, 11,* 133–150.

Nunes, T., & Bryant, P. (2009). *Children's reading and spelling: Beyond the first steps*. Oxford, England: Wiley-Blackwell.

Nunes, T., Bryant, P., & Bindman, M. (1997a). Learning to spell regular and irregular verbs. *Reading and Writing: An Interdisciplinary Journal, 9,* 427–449.

Nunes, T., Bryant, P., & Bindman, M. (1997b). Morphological spelling strategies: Developmental stages and processes. *Developmental Psychology, 33,* 637–649.

Nunes Carraher, T., & Rego, L. R. B. (1984). Desenvolvimento cognitivo e alfabetização [Cognitive development and alphabetization]. *Revista Brasileira de Estudos Pedagógicos, 63,* 38–55.

Pacton, S. (2004, July). *Children's use of syntactic information in spelling.* Paper presented at the meeting of the Society for the Scientific Study of Reading, Amsterdam, the Netherlands.

Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities. *Child Development, 76,* 324–339.

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General, 130,* 401–426.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10,* 233–238.

Piaget, J. (1950). *The psychology of intelligence*. New York, NY: Harcourt, Brace.

Piaget, J. (1954). *The construction of reality in the child*. New York, NY: Basic Books.

Pollo, T. C., Kessler, B., & Treiman, R. (2005). Vowels, syllables, and letter names: Differences between young children's spelling in English and Portuguese. *Journal of Experimental Child Psychology, 92,* 161–181.

Pollo, T. C., Kessler, B., & Treiman, R. (2009). Statistical patterns in children's early writing. *Journal of Experimental Child Psychology, 104,* 410–426.

Pollo, T. C., Treiman, R., & Kessler, B. (2008). Three perspectives on spelling development. In E. J. Grigorenko & A. Naples (Eds.), *Single-word reading: Cognitive, behavioral, and biological perspectives* (pp. 175–189). Mahwah, NJ: Erlbaum.

Porpodas, C. D. (2001). Cognitive processes in first grade reading and spelling of Greek. *Psychology: The Journal of the Hellenic Psychological Society, 8,* 384–400.

Read, C. (1975). *Children's categorization of speech sounds in English*. (NCTE Research Report No. 17). Urbana, IL: National Council of Teachers of English.

Read, C. (1986). *Children's creative spelling.* London, England: Routledge and Kegan Paul.

Rego, L. R. B. (1999). Phonological awareness, syntactic awareness and learning to read and spell in Brazilian Portuguese. In M. Harris & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 71–88). Cambridge, England: Cambridge University Press.

Robins, S., Treiman, R., & Rosales, N. (2014). Letter knowledge in parent-child conversations. *Reading and Writing: An Interdisciplinary Journal, 27,* 407–429.

Roy-Charland, A., Saint-Aubin, J., & Evans, M. A. (2007). Eye movements in shared book reading with children from kindergarten to grade 4. *Reading and Writing: An Interdisciplinary Journal, 20,* 909–931.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926–1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70,* 27–52.

Saffran, J. R., Reeck, K., Niebuhr, A., & Wilson, D. (2005). Changing the tune: The structure of the input affects infants' use of absolute and relative pitch. *Developmental Science, 8,* 1–7.

Sandra, D. (2010). Homophone dominance at the whole-word and sub-word levels: Spelling errors suggest full-form storage of regularly inflected verb forms. *Language and Speech, 53,* 405–444.

Sandra, D., Frisson, S., & Daems, F. (1999). Why simple verb forms can be so difficult to spell: The influence of homophone frequency and distance in Dutch. *Brain and Language, 68,* 277–283.

Seidenberg, M. S., & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences, 4,* 353–361.

Shen, H. H., & Bear, D. R. (2000). Development of orthographic skills in Chinese children. *Reading and Writing: An Interdisciplinary Journal, 13,* 197–236.

Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development, 74,* 27–47.

Smith, E. E., Langston, C., & Nisbett, R. E. (1992). The case for rules in reasoning. *Cognitive Science, 16,* 1–40.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14,* 638–647.

Tong, X., McBride-Chang, C., Shu, H., & Wong, M.-Y. (2009). Morphological awareness, orthographic knowledge, and spelling errors: Keys to understanding early Chinese literacy acquisition. *Scientific Studies of Reading, 13,* 426–452.

Totereau, C., Thenevin, M. G., & Fayol, M. (1997). The development of the understanding of number morphology in written French. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory and practice across languages* (pp. 97–114). Mahwah, NJ: Erlbaum.

Turnbull, K., Deacon, S. H., & Kay-Raining Bird, E. (2011). Mastering inflectional suffixes: A longitudinal study of beginning writers' spellings. *Journal of Child Language, 38,* 533–553.

Treiman, R. (1985). Phonemic awareness and spelling: Children's judgments do not always agree with adults'. *Journal of Experimental Child Psychology, 39,* 182–201.

Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. New York, NY: Oxford University Press.

Treiman, R. (1994). Use of consonant letter names in beginning spelling. *Developmental Psychology, 30,* 567–580.

Treiman, R., & Broderick, V. (1998). What's in a name: Children's knowledge about the letters in their own name. *Journal of Experimental Child Psychology, 70,* 97–116.

Treiman, R., & Cassar, M. (1996). Effects of morphology on children's spelling of final consonant clusters. *Journal of Experimental Child Psychology, 63,* 141–170.

Treiman, R., Cassar, M., & Zukowsi, A. (1994). What types of linguistic information do children use in spelling? The case of flaps. *Child Development, 65,* 1318–1337.

Treiman, R., & Kessler, B. (2003). The role of letter names in the acquisition of literacy. In R. Kail (Ed.), *Advances in child development and behavior* (Vol. 31, pp. 105–135). San Diego, CA: Academic Press.

Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology, 98,* 642–652.

Treiman, R., Kessler, B., & Bourassa, D. (2001). Children's own names influence their spelling. *Applied Psycholinguistics, 22,* 555–570.

Treiman, R., Pollo, T. C., Cardoso-Martins, C., & Kessler, B. (2013). Do young children spell words syllabically? Evidence

from learners of Brazilian Portuguese. *Journal of Experimental Child Psychology, 116,* 873–890.
Treiman, R., & Yin, L. (2011). Early differentiation between drawing and writing in Chinese children. *Journal of Experimental Child Psychology, 108,* 786–801.
Walker, J., & Hauerwas, L. B. (2006). Development of phonological, morphological, and orthographic knowledge in young spellers: The case of inflected verbs. *Reading and Writing, 19,* 819–843.
Yin, L., & Treiman, R. (2013). Name writing in Mandarin-speaking children. *Journal of Experimental Child Psychology, 116,* 199–215.
Zeno, S. (1995). *The educator's word frequency guide.* Brewster, NJ: Touchstone Applied Science Associates.

CHAPTER

22 Learning to Read and Spell Words in Different Writing Systems

Markéta Caravolas *and* Anna Samara

Abstract

There is strong evidence that word reading and spelling ability in English is founded on three core skills, namely knowledge of letters of the alphabet, awareness of phonemes in spoken words, and rapid automatized naming (RAN) of visual stimuli. It is suggested here that these abilities represent cognitive constructs that constitute the triple foundation of literacy in all languages. The chapter reviews the research carried out in different writing systems to assess the extent to which this triple foundation provides a good language-general model of early literacy development. The evidence is considered in the context of potentially important moderating, language-specific influences of orthographic variables, especially symbol–sound mapping consistency. It is proposed that the triple foundation model, conceptualized as (1) knowledge of the functional symbol set of the orthography, (2) awareness of the speech units to which orthographic symbols map, and (3) efficient mappings between the representational systems of orthographic symbols and their related speech units, provides a universally valid description of the cognitive architecture underlying early literacy development.

Key Words: reading, spelling, cross-linguistic, foundation model, predictor, alphabetic, nonalphabetic, orthography, phoneme awareness

Understanding how literacy skills are acquired is important to all literate societies. Cross-linguistic studies of reading and spelling development contribute to this goal by uncovering the universal and the language-specific aspects of this learning process. At the broadest level, literacy development depends on the interplay between the cognitive abilities of the learner and the nature of the writing system being learned. While the development of the cognitive competencies needed for learning to read and spell is assumed to be similar across languages and cultures (e.g., Samuelsson et al., 2005), orthographies vary along a number of important dimensions, including the type, number, and complexity of their writing symbols, and these variations may specifically influence the learning process (e.g., Kessler & Treiman, this volume; Perfetti & Dunlap, 2008; Seidenberg, 2011; Ziegler & Goswami, 2005). To take an obvious example, learning the thousands of characters of Chinese surely takes longer than learning the twenty-six letters of the English alphabet. However, reading and writing are complex activities, and it is important to understand the more subtle ways in which learning might (or might not) be affected by linguistic and orthographic variation. In the present chapter, we review the literature in search of the essential cognitive abilities that constitute the foundations of literacy in any orthography. We also seek to identify those features of orthographies that influence the patterns and rates of reading and spelling development in language-specific ways. Studies that have taken a direct cross-linguistic approach are prioritized.

The study of the foundations of literacy concerns that phase during which children progress from having no reading or spelling/writing abilities to the point at which they become independent decoders and spellers, able to tackle the reading and writing of single words and simple texts (even if imperfectly) on their own. For researchers and educators, it is important to know what cognitive skills enable the learner to make this transition into literacy. We consider the research evidence within the theoretical framework of the *triple foundation model* (an elaboration on Byrne's (1998) *dual foundation model,* to which we return in a later section). According to this view, the ability to acquire word reading and writing skills, regardless of the language and writing system, depends crucially on three core precursor abilities: the ability to have conscious awareness of and the ability to manipulate oral sublexical units (e.g., phonemes, syllables, morphemes), eventually those which correspond to the basic symbols of one's orthography; the ability to learn the functional set of writing symbols of one's orthography (e.g., letters, syllabographs, logographs); and the ability to establish and use quick and efficient connections between the linguistic units and their corresponding orthographic units (see Figure 22.1). We report research in different languages and orthographies that is relevant to the proposed triple foundation model, examine how well the findings fit the model, and consider what questions they raise for further research.

The triple foundation model constitutes a causal, predictive hypothesis about the architecture that drives early literacy development; such theoretical models are typically tested in correlational, often longitudinal studies that use statistical modeling techniques such as path analysis, structural equation modeling, and growth curve analysis. Accordingly, the present review focuses heavily on correlational and longitudinal research of the predictors of reading and writing across orthographies and to a lesser extent on experimental and intervention research. The latter approaches are often used to uncover how written (and oral) language input affects the patterns of performance (e.g., types of reading and spelling errors, error rates) and the use of strategies in reading and writing across languages (e.g., phonologically versus semantically or lexically weighted processes; small-grain versus larger-grain reading units). Such studies are beyond the scope of the present chapter and have been reviewed elsewhere (e.g., Caravolas, 2005; Perfetti, Liu, & Tan, 2005; Share, 2008; Ziegler & Goswami, 2005).

Much of the research evidence in the area of early literacy comes from studies of languages that represent spoken words with alphabetic letters (such as those of

Language-General Model

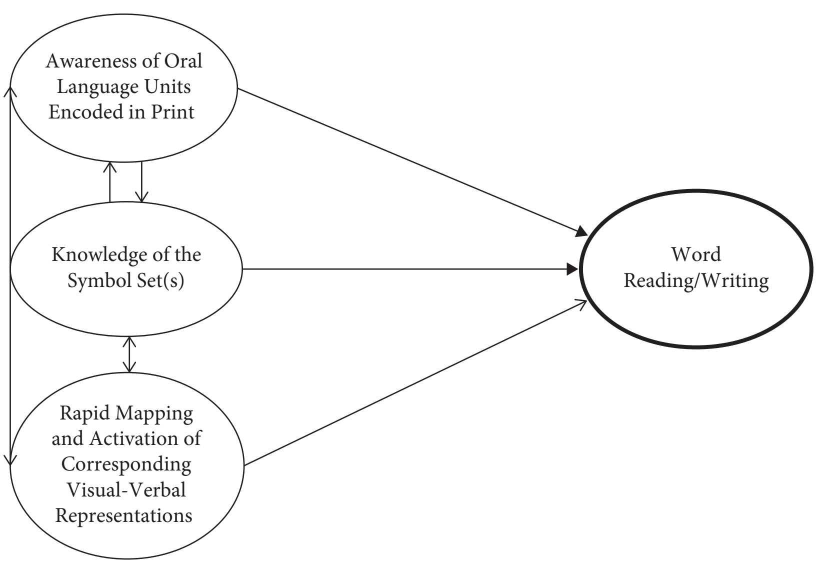


Fig. 22.1 The proposed language-general, *Triple Foundation Model* of early literacy, contains two representational systems: sublexical representations (estimated by phonological awareness and/or morphological awareness) and orthographic symbol representations (estimated by knowledge of functional symbols sets, such as letters, syllabographs, logographs) which reciprocally promote growth in each other, and, a mechanism for the efficient mappings between them (estimated by RAN). Each component skill uniquely promotes growth in early reading and spelling skills.

the Latin alphabet of many European orthographies), and especially from the study of English-speaking populations (e.g., Share, 2008). Accordingly, in the first part of this chapter we focus on evidence from English and other alphabetic orthographies. In the second part, we broaden the review to the growing body of research on early literacy in the nonalphabetic systems of logographies (e.g., Chinese) and alphasyllabaries (e.g., Korean, Kannada). Then, summarizing across the findings, we consider the implications for a language-general understanding of the foundations of literacy development.

Learning to Read and Spell in Alphabetic Orthographies

Features of Alphabetic Orthographies

Alphabets generally are relatively small sets of visually simple symbols, called letters, that correspond to phonemes. Languages that frequently use more than one letter to represent a given phoneme, like English, tend to have fewer alphabet letters than those that have created dedicated letter symbols to represent all or most of the phonemes of the spoken language, like Czech. For example, to represent the phonemes /t͡ʃ/, /ʃ/, /iː/, usually English uses the digraphs ‹ch›, ‹sh›, ‹ea›, while Czech has the dedicated letters with diacritics ‹č›, ‹š›, ‹í› in its alphabet (as distinct from ‹c›, ‹s›, ‹i›, which represent /t͡s/, /s/, /i/, respectively). Despite such differences, the variation in the number of letters across orthographies is quite restricted: in the range of 20 to 40 (see Kessler & Treiman, this volume). To our knowledge, the effect of alphabet size on learning to read and write has not been considered in cross-linguistic research. Nevertheless, learning to recognize and to write letters of the alphabet in all their aspects (upper- and lowercase, by their names and their corresponding speech sounds, and in different scripts—print or cursive) is the focus of much direct instruction during the first year or two of schooling in many countries (e.g., Caravolas et al., 2012, supplement), and children can vary widely in their ability to learn the alphabet.

Despite similarities in sizes of alphabets, alphabetic orthographies vary in the complexity of their *phonographemic* (henceforth sound–letter) and *graphophonemic* (henceforth letter–sound) consistency. By consistency, we refer to the extent to which multiple pronunciations exist for a particular letter or grapheme when reading, and conversely, the extent to which multiple spellings exist for a particular speech unit (e.g., phoneme or syllable) when spelling. Inconsistency in letter–sound mappings can arise for a variety of reasons. For instance, many alphabetic orthographies preserve some *morphological* information about words at the expense of phoneme–grapheme consistency. In English, for example, the vowel letter digraph ‹ea› corresponds to two different phonemes, /i/ and /ɛ/, in the morphologically related words ‹st*ea*l› and ‹st*ea*lth›. Another important source of inconsistency in some orthographies is the effect of *lexical stress*. For example, unstressed vowels in English are frequently pronounced as the schwa vowel /ə/, which can map to any short vowel spelling in the language (e.g., ‹scall*o*p› /skæləp/, ‹cact*u*s› /kæktəs/, ‹rock*e*t› /rɒkət/), and often the written vowel has no audible phonemic counterpart at all (e.g., ‹mas*o*n› /meˈsn̩/, ‹hamm*e*r› /hæmr̩/ (e.g., Treiman, Berch, & Weatherston, 1993). Also, the adoption of *foreign words* with their original spellings (e.g., /ʃæmpeˈn/ spelled as ‹champagne› in English), and the retention of *historical spellings* that reflect archaic word pronunciations (e.g., the ‹k› in ‹*knee*›, once pronounced, is now phonologically redundant) can entail inconsistency. When letters have no correspondence in speech, they produce *opacity* in the spelling system; this is more frequent in systems that have not been reformed recently, such as English and French.

To the extent that they are present within orthographies, the preceding factors all reduce spelling–sound and sound–spelling consistency. Learners may be explicitly taught to resolve such ambiguities by codified orthographic rules, and they may learn many types of orthographic patterns and constraints implicitly (e.g., Cassar & Treiman, 1997; Pacton, Fayol, & Perruchet, 2002; Samara & Caravolas, 2014). However, from the beginner's point of view, all else being equal, letters and letter sequences (graphemes) with only one possible pronunciation are more consistent and easier to learn to read than those with several possible pronunciations. Conversely, speech sounds with only one possible spelling are more consistent and easier to spell than those with several possible spellings.

Alphabetic orthographies can be described in terms of a continuum of consistency with languages such as Finnish and Turkish at the high consistency end, having mainly one-letter-to-one-sound mappings, and others, such as French and English, at the low consistency end, having many-to-one letter-sound and sound-letter mappings. There are numerous ways to estimate consistency. It is possible to simply count the number of possible letters and letter strings (graphemes) that may represent the phonemes in a language (e.g., in English, the vowel /ɒ/ can be represented by the graphemes ‹*o*›, ‹*a*›, ‹*ough*›, ‹*aw*› as well as a few others, while in the relatively consistent Czech orthography this same vowel is only represented by ‹o›. More informative and precise measures have been developed, including consistency estimates of larger units such as *rimes* (vowel and ensuing consonant(s) in a syllable) (Peereman & Content, 1999; Ziegler, Jacobs, & Stone, 1996), and of graphemes and phonemes (e.g., Kessler & Treiman, 2001; Peereman, Lété, & Sprenger-Charolles, 2007). At the letter–sound level, the estimates reflect the proportion of words (weighted by their frequency) in which a given letter–sound or sound–letter correspondence occurs (e.g., the phoneme /i/ spelled as ‹ee› in words like ‹feed›, ‹meet›) relative to all of the words containing the phoneme /i/ regardless of its spelling (e.g., ‹fee*t*›, ‹m*ea*l›, ‹*sc*en*e*›). The corpora of Peereman et al. (2007) for French and of Kessler and Treiman (2001) for American English include data from child-directed printed materials, and they report consistency information across as well as within specific word positions. Kessler and Treiman's (2001) work demonstrated that in English, individual letter–sound and sound–letter correspondences become considerably more predictable (consistent) if the adjacent grapheme or phoneme in the syllable and its position are taken into account. For example, knowing that /iː/ is followed by /t/ improves the predictability of its spelling. However, in the earliest stages of learning to read and spell, children focus more on the attributes of small-grain (letter–sound) units in words, and only later in development make use of the larger units and letter environments (e.g., Caravolas, Kessler, Hulme, & Snowling, 2005; Hayes, Treiman, & Kessler, 2006; Pacton et al., 2002; Rittle-Johnson & Siegler, 1999).

Databases with sophisticated consistency estimates based on children's materials are available for only a few languages to date. Some are in development (e.g., West Slavic lexical database; Kessler & Caravolas, 2011), and many languages still completely lack corpus-based calculations of consistency. A simpler approach that has been used in several cross-linguistic studies is the measure of onset entropy (e.g., Borgwaldt, Hellwig, & deGroot, 2005; Ziegler et al., 2010), where only the initial letter of a word is taken into consideration, and the probability of each of its possible pronunciations is calculated and summed. Initial letters with only one possible pronunciation have an entropy value of 0; this value increases as the number of possible pronunciations increases. Entropy estimates are useful for estimating the consistency ranking of orthographies (e.g., Ziegler et al., 2010).

To this point, we have discussed orthographic consistency as a factor differentiating the complexity of alphabetic orthographies. However, linguistic variables such as syllable structure can also vary considerably among languages, and this is of course reflected in word spellings in alphabetic orthographies. In one of the first multilanguage studies comparing early literacy attainment among learners of fourteen different languages, Seymour, Aro, and Erskine (2003) posited that two main factors determine the rates and patterns of growth of reading and spelling in alphabetic orthographies: syllabic complexity and orthographic depth (including inconsistent letter–sound correspondences and morphological influences). Languages with predominantly open syllables (consonant-vowel, or CV) and relatively few clusters in the onset position, as tends to be the case in Romance languages, were classified as having simple syllable structure; those containing many closed syllables (consonant-vowel-consonant, or CVC) and consonant clusters in both the onset and coda position, as is common in Germanic languages, were considered to have complex syllable structure. However, while syllable structure has been found to impact children's performance in tasks such as phoneme awareness and spelling, with better performance on structures that occur more frequently in their language (e.g., Caravolas & Bruck, 1993; Durgunoğlu & Öney, 1999), consistency seems to be the stronger driver of differences in the rates of early reading and spelling development. For example, Seymour et al.'s (2003) data showed that syllable structure was not a strongly discriminating factor of overall literacy attainment, as learners of some Romance languages (e.g., Portuguese, French) were among the lowest achieving groups on comparable measures of reading, while learners of some Germanic languages (German, Swedish, Dutch) performed on a par with the highest scoring simple syllable structure groups (e.g., Finnish, Greek, Italian). These trends have been replicated in more focused comparisons of pairs of languages with similar syllable structures but differences in orthographic consistency. For example, given reading lists of comparable words and nonwords (e.g., Wimmer & Goswami, 1994) or identical (cognate) words (Landerl, Wimmer, & Frith, 1997), comparisons of German and English primary school children have repeatedly shown much lower attainments for the English-speaking groups. Defior, Martos, and Cary

(2002) demonstrated that the impact of inconsistency is measurable even when the two languages and cultures are much more closely matched, as in the case of Spanish and Portuguese. These languages have similar syllable structures; however, the Portuguese orthography is subtly less consistent than Spanish. Defior et al. (2002) found that given parallel items (i.e., nonwords derived from number words: e.g., ‹ninco› derived from ‹cinco›, 'five'), the Portuguese children read more slowly than their Spanish peers at every grade level from first to fourth. They read less accurately in first and second grades catching up in third grade. Thus when syllable structure is very similar across stimuli, as well as across phonological systems, but orthographic consistency differs between languages, the less consistent orthographies are associated with slower and more error-prone literacy development. Taking the opposite tack, Caravolas and Landerl (2010) examined first-grade learners of Czech—a language with complex syllable onset structures but simple (predominantly open) coda structures—and Austrian German—a language with relatively complex syllables, in particular in codas—on parallel tests of phoneme awareness, reading, and spelling at the start and end of the school year. Both orthographies have relatively high letter–sound consistency. While the groups (well matched on a host of demographic and educational variables) demonstrated clear language-specific patterns of performance on oral phoneme awareness tasks, each performing better on the complex units with which they had greater experience, the groups were indistinguishable in their overall attainments in reading and spelling at the start (both near floor) as well as at the end (both highly proficient) of the school year. Thus, differences in syllable complexity between languages with relatively consistent orthographies do not seem to be associated with notable differences in early literacy attainments in reading speed or accuracy (see also Share, 2008, p. 585).

In sum, factors such as morpheme–grapheme constancy, lexical stress, and historical and foreign spellings all affect letter–sound and sound–letter consistency in language-specific ways. While children may learn to resolve some of these inconsistencies more easily than others, it is likely that in the initial stages of learning their progress will be hampered by all sources of inconsistency. Indeed, studies show that inconsistency is a major determinant of language-specific differences in certain aspects of learning to read and spell alphabetically, and most certainly of differences in the rate of acquisition (e.g., Seymour et al., 2003). However, much less agreement exists about the impact of letter–sound consistency on the cognitive architecture that underpins and drives early reading and writing development in different languages. We begin to examine this issue with a review of what has been learned from English.

Foundations of Literacy in English

A wealth of evidence from studies of English shows that during the early phases of reading and spelling development, children learn to apply the *alphabetic principle.* That is, they rely extensively on their ability to identify the letters in words, to activate their corresponding pronunciations, and to assemble them into best approximations of the words they are attempting to read. Conversely, when spelling, beginners identify the sounds in words, connect these to their corresponding letter symbols, and assemble them into best approximations of the intended words (Ehri, 1997, 2014; Treiman, 1993). These findings were expressed within the dual foundation model of early literacy proposed by Brian Byrne (1998), in which knowledge of the letters of the alphabet and conscious awareness of the phonemes (speech sounds) constituting spoken words are separately emerging skills, mutually promoting growth in each other, and independently contributing to individual variations in growth in reading and spelling abilities (e.g., Hulme, Caravolas, Málková, & Brigstocke, 2005). Longitudinal studies of the predictors of word reading (e.g., Muter, Hulme, Snowling, & Stevenson, 2004) and spelling development (e.g., Caravolas, Hulme, & Snowling, 2001) over the first three years of schooling have confirmed the validity and the robustness of this model with English-speaking populations. In these studies, letter knowledge and phoneme awareness, along with the *autoregressors* (earlier measures of the skills being predicted) measured at the start of schooling, accounted for approximately two-thirds of the variation in reading and spelling over the first two years of formal instruction. Moreover, several studies have demonstrated that training children in phoneme awareness and letter knowledge skills is associated with improvements in reading and spelling skills (e.g., Ball & Blachman, 1991; Bradley & Bryant, 1983; Hatcher, Hulme, & Ellis, 1994). In a recent intervention study of British English school beginners whose weak oral language skills placed them at risk of poor literacy attainment, Bowyer-Crane et al. (2008) showed that twenty weeks of daily teaching of letters knowledge, phoneme awareness, and basic reading skills led to improvements in their literacy skills; in contrast, their peers who instead received twenty weeks of oral language, vocabulary, and comprehension instruction did not make significant literacy gains. In a follow-up analysis, Hulme, Bowyer-Crane, Carroll, Duff, and Snowling (2012) demonstrated that the literacy skills advantage of the phonology-with-reading group, which was still evident six months after the intervention, was fully accounted for by the children's posttraining phoneme awareness and letter knowledge skills. Together, such longitudinal and intervention studies provide strong evidence of the causal role of these skills as drivers of early alphabetic literacy.

Although the dual foundation model is well supported by English-language data, it is now clear that a third skill, the ability to rapidly name aloud visually presented stimuli such as pictures, colors, digits, or letters—referred to as *rapid automatized naming* (RAN)—is also a reliable predictor of reading and spelling (e.g., Denckla & Rudel, 1976; Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004). The *nonalphanumeric* (colors, pictures) versions of the task are typically used with children who do not yet know the names of letters and digits to an automatized (fast and accurate) level (e.g., Caravolas et al., 2012; Kirby, Parrila, & Pfeiffer, 2003; Lervåg & Hulme, 2009), while the *alphanumeric* versions (digits, letters) are used with those who have reliably acquired knowledge of these. Both versions have been found to correlate concurrently and longitudinally with literacy abilities (although alphanumeric RAN has often been found to be the stronger predictor) (Kirby, Georgiou, Martinussen, & Parrila, 2010; Norton & Wolf, 2012). This, and the finding that reading ability does not predict growth in RAN skills (Lervåg & Hulme, 2009), suggests that RAN tasks do not simply measure components of reading itself (e.g., that speed of naming letters can be viewed as one basic aspect of reading ability), but rather that they estimate skills more broadly related to literacy development.

The RAN construct has been studied extensively to understand what cognitive functions it estimates and what role it plays in typical and impaired reading development (see reviews in Kirby et al., 2010; Norton & Wolf, 2012). Two main views have been put forward. One sees RAN as an aspect of phonological processing, while the other relates it to orthographic processing. (The latter term is defined in different ways by different authors [e.g., Burt, 2006]; Kirby et al. [2010] defined orthographic processing as "when groups of letters or entire words are processed as single units rather than as a sequence of grapheme-phoneme correspondences" [p. 343].) For example, Torgesen, Wagner, Rashotte, Burgess, and Hecht (1997) explained RAN as a proxy measure for the speed of access to and retrieval of *phonological* information from memory. In contrast, Bowers and Wolf (1993) proposed RAN to estimate the speed of activation and formation of *orthographic* representations; that is, representations of letters and word spellings. A more recent hypothesis, supported by behavioral and neuroimaging research, is that RAN tasks tap a mechanism that enables the formation of rapid associations between orthographic and phonological representations, which overlap anatomically with object identification and object naming brain regions (e.g., Price & McCrory, 2005; Vaessen, Gerretsen, & Blomert, 2009; Wimmer & Schurz, 2010). Lervåg and Hulme (2009) have speculated that, fundamentally, RAN tasks assess the integrity and the efficiency of the representational systems for visual object identities (which include letters and word spellings), their retrieval, and their pronunciations. However, this and similar hypotheses still await direct empirical validation.

Incorporating the latter hypothesis into the dual foundation framework, we suggest that, during the foundational stage of literacy development, when letter knowledge and phoneme awareness drive the acquisition of the alphabetic principle, RAN may estimate the development of efficient associations between children's representations of letters (eventually letter strings) and their corresponding phonemes and as such represents a third component that is essential for reading and spelling (see Figure 22.1). The findings of several longitudinal English-language studies confirm that RAN accounts for unique variation in reading and spelling attainments, over and above contributions from phoneme awareness and letter knowledge (Georgiou, Torppa, Manolitsis, Lyytinen, & Parrila, 2012; Kirby et al., 2003; Schatschneider et al., 2004), and are consistent with the triple foundation model. Moreover, when additional skills, including IQ, verbal short-term memory, vocabulary, morphological awareness, executive function, and visual attention have been considered over and above the three core predictors, they have not systematically accounted for additional unique variance in reading and spelling outcomes. This suggests that the most reliable, proximal components of early reading and spelling are indeed phoneme awareness, letter knowledge, and RAN. An important question for the present chapter is whether this model generalizes across alphabetic orthographies and to what extent it may be influenced by variations in orthographic consistency.

Foundations of Literacy in Other Alphabetic Orthographies

In this section we discuss research in a variety of languages according to their relative orthographic consistency (as defined earlier). This approach does not negate the importance of other attributes that may differentiate orthographic systems. A dominant view in the cross-linguistic literature is that the impact of letter-sound consistency is profound, affecting not only the rate of learning but also aspects of the cognitive architecture underlying growth in early reading and spelling (e.g., Share, 2008; Wimmer, Mayringer, & Landerl, 2000). It has been argued that, whereas phoneme awareness is an important predictor of reading in English, its role is less important in consistent orthographies mainly because the highly consistent letter–sound mappings, acquired through systematic phonics teaching methods (which are typical in most European countries), are sufficient to boost and override any preexisting individual variations in phoneme awareness within the first year or two of schooling. Investigations in languages with relatively consistent orthographies, such as German and Dutch, have reported a weaker and more time-limited role for phoneme awareness in accounting for variance in reading than typically reported for English (de Jong & van der Leij, 1999, 2002; Wesseling & Reitsma, 2000; Wimmer et al., 2000). On the other hand, RAN is thought to play a more important and enduring role in reading in consistent orthographies. This is because individual differences in reading ability manifest primarily through reading fluency (not accuracy, which tends to reach ceiling by the end of first or second grade); RAN is assumed to tap the substrate that constrains reading fluency (Wimmer et al., 2000). Wimmer et al. (2002; also Wolf & Bowers, 1999) further proposed that RAN underlies the mechanism that enables learners to build up well-specified orthographic representations. Accordingly, for individuals who have slow activation of orthographic–phonological correspondences, this results in poorly acquired orthographic representations of units larger than single letters, and the deficit manifests especially in spelling difficulties. Thus, in this view, phoneme awareness and RAN are both expected to be robust predictors of spelling attainment.

In contrast to the preceding views, the triple foundation model predicts that the core cognitive skills required for children to establish the alphabetic principle and early reading and spelling skills should play the same role and should have similar relative weightings within languages, regardless of orthographic consistency. This prediction is based on the assumption that languages and their orthographies should not influence the development of the cognitive precursor skills of literacy among learners (e.g., Samuelsson et al., 2005) and that reading and spelling in any alphabetic orthography entails the same basic processes of mapping letters and letter strings to phonological units and hence relies on the same underlying cognitive architecture (see also Norris & Kinoshita, 2012; Perfetti & Dunlap, 2008; Seidenberg, 2011). The model describes the foundation phase and therefore does not make strong predictions about the development of later-developing literacy skills.

Numerous cross-linguistic investigations in languages with alphabetic orthographies of varying degrees of consistency have focused on the issues of relative weighting, timing, and persistence of the three core predictors of reading and spelling, and have made direct or indirect comparisons with English as the benchmark language (see reviews in Share, 2008). Studies that included measures of letter knowledge prior to or at the start of formal literacy instruction have unanimously found it to be an important independent predictor of reading and spelling attainment over and above measures such as IQ, phoneme awareness, and RAN. Its effects have been observed to persist, directly or indirectly, through the second and third grades in numerous languages spanning the orthographic consistency spectrum, including English, French, Greek, Norwegian, Dutch, German, and Finnish (Bruck, Genesee, & Caravolas, 1997; Caravolas et al., 2001; de Jong & van der Leij, 1999; Furnes & Samuelsson, 2011; Georgiou et al., 2012; Lervåg, Bråten, & Hulme, 2009; Muter et al., 2004; Wagner, Torgesen, & Rashotte, 1994; Wimmer, Landerl, Linortner, & Hummer,1991). Thus, the ease with which children can learn the names and sounds of the letters of their alphabet typically accounts for a significant proportion of variation in early reading and spelling achievements.

However, findings about the roles of phoneme awareness and RAN have not been systematically replicated in various cross-linguistic investigations and have not consistently endorsed the dominant view. For example, studies of primary-school-age pupils learning the relatively consistent Dutch (Patel, Snowing, & de Jong, 2004), Norwegian, Swedish (Furnes & Samuelsson, 2009), and Czech (Caravolas et al., 2005) orthographies reported significant associations between phoneme awareness and reading (and spelling) that persisted into mid-to-late primary school years, on a par with their English-learning counterparts. In particular, Caravolas et al. (2005) demonstrated that when parallel and adequately difficult phoneme awareness measures were used, they accounted for identical proportions of the variance in reading and spelling in Czech and English. Moreover, the evidence of a stronger effect of RAN in lieu of phoneme awareness on reading in consistent orthographies (e.g., Mann & Wimmer, 2002; Wimmer et al., 2002) has proven equivocal. Several recent multilanguage studies are directly relevant to this issue; we turn to these next.

A European study reported by Ziegler et al. (2010) compared the concurrent relationships between phoneme awareness, RAN, and verbal short-term memory as predictors of word and nonword reading speed and accuracy among second-grade children. Five language groups with orthographies ranging widely in orthographic consistency (in order from least to most consistent according to Borgwaldt et al.'s [2005] entropy estimates: French, Portuguese, Dutch, Hungarian, and Finnish) were represented. Controlling for IQ, phoneme awareness emerged as the main predictor in all languages and across all four reading measures, and it was in all cases stronger than RAN. Moreover, RAN did not predict reading accuracy in any analysis, and, its contribution to reading speed failed to reach significance in the most consistent orthography of the set (Finnish). These results are at odds with the dominant hypothesis that phoneme awareness is more important for the inconsistent English orthography, while RAN is most predictive in consistent orthographies (e.g., de Jong & van der Leij, 2002; Wimmer et al., 2000, 2002). Shortcomings of this study, however, were that the test batteries seemed to be only loosely matched, so it is not certain whether between-language variations were in part due to test form variations; no reliabilities were reported for any test, and only single measures were obtained for each skill (single measures offer less reliable estimates of specific constructs than multiple measures). Letter knowledge was not assessed, precluding a full test of the triple foundation model.

A second cross-sectional study by the same group, this time comparing Portuguese, Dutch, and Hungarian children in first to fourth grades, investigated predictors of reading fluency (Vaessen et al., 2010). In this study, the test batteries were more robust, with multiple reliable measures of letter knowledge, phoneme awareness, and RAN. Regression analyses revealed comparable patterns of results across the three languages such that letter knowledge was a strong predictor in the first two grades and then waned in importance, while RAN gained in importance over grades. Measures of phoneme awareness accuracy and speed remained significant and stable predictors over grades, although they fluctuated somewhat, such that accuracy was more important in less consistent orthographies, while speed was more important for more consistent orthographies. The main limitation of this study was that it was cross-sectional.

In one of the very few longitudinal studies, Georgiou et al. (2012) compared English, Greek, and Finnish children who were followed from kindergarten to second grade. All groups were assessed on measures of letter knowledge, phonological awareness, and RAN in kindergarten, and on nonword reading accuracy, text reading fluency, and spelling at the end of second grade. Multigroup path analyses revealed that letter knowledge was a significant and stable predictor in each language and for each literacy skill. Phoneme awareness only predicted nonword decoding in English, while RAN predicted nonword decoding in English, spelling in Greek and English, and reading fluency in Greek. Surprisingly, neither phoneme awareness nor RAN contributed to reading or spelling development in Finnish. A limitation of the study was the two-year gap between time 1 and time 2 testing; learners undergo many rapid changes in the first two years of schooling, especially in consistent orthographies. Thus, the study of Georgiou et al. may have missed important transitions in development. Also, this study obtained only single estimates of each predictor and outcome measure, and reading and spelling ability were not estimated at time 1, precluding controls of autoregressors.

The methodological variations and shortcomings of previous studies, many of which considered single languages, have resulted in a mixed picture. Nevertheless, the emerging consensus seems to be that phoneme awareness is an important predictor of reading and spelling in all alphabetic orthographies (e.g., Caravolas et al., 2005; Vaessen et al., 2010; Ziegler et al., 2010), although some studies show its role to weaken more quickly in consistent orthographies (Lervåg & Hulme, 2009; Vaessen et al., 2010; Ziegler et al., 2010). The role of RAN may increase as reading skills increase. Yet in line with the triple foundation model, the preceding studies provide little evidence to suggest that this applies more to consistent than inconsistent orthographies. But most fundamentally, for a causal theory of literacy development, well-controlled, cross-linguistic, longitudinal studies are needed of children assessed prior to or at the start of formal literacy instruction, and again within the first year or two of schooling. That is surely the window within which to capture the components of the emerging reading and spelling system and to examine the relationships among them.

This design was applied in a recent study by Caravolas and colleagues (Caravolas et al., 2012) in a four-language, longitudinal comparison of the roles of phoneme awareness, letter knowledge, and RAN, along with measures of vocabulary, nonverbal ability, and verbal memory (measured prior to [Spanish, Czech, Slovak] or just after [English] the onset of reading instruction) as predictors of variations in reading and spelling skills measured approximately ten months later in mid-grade one. All core skills were estimated from highly reliable, composite measures. Relative to English, the Spanish, Czech, and Slovak orthographies have highly consistent letter–sound mappings, as confirmed by context-free estimates for each language (e.g., Kessler & Caravolas, 2011). Additional information gathered via questionnaires and other literacy measures in each of the four languages, confirmed that schooling practices, including preschool language- and literacy-related experiences, primary literacy teaching methods, and relevant cultural norms were very similar across groups. Multigroup (multilanguage) structural equation models predicting reading and spelling, respectively, yielded very clear results: Prior reading/spelling skills, phoneme awareness, letter knowledge, and RAN were all unique longitudinal predictors of the growth of reading/spelling skills in the first ten months spanning the transition into formal education in all four languages. Both models yielded excellent fits to the data, providing strong support for the idea that these three measures are of relatively equal importance as predictors of progress in learning to read and to spell across the four languages studied. Finally, both models accounted for over two-thirds of the variance in reading and spelling skills. These findings are similar to those of another recent cross-linguistic comparison of the patterns of growth in reading and spelling in English and the consistent orthographies of Swedish and Norwegian (Furnes & Samuelsson, 2011).

To summarize, the growing body of cross-linguistic research on early literacy development suggests that phoneme awareness, letter knowledge, and RAN are universal foundation skills of alphabetic literacy. Studies that have tracked learners at similar moments in their literacy development have revealed remarkably similar cross-linguistic patterns. That is, as the reading/spelling system is being established, all three core predictors play a role from the beginning of literacy development, their relative importance seems to be similar in each language, and their effects persist at least into the middle of grade one (Caravolas et al., 2012; Furnes & Samuelsson, 2011). However, these studies cannot directly assess the hypothesis that RAN taps a mechanism for the efficient mappings between letter knowledge and phoneme awareness. This could be tested in future studies, for example by modeling RAN as a moderator for the longitudinal, reciprocal relationship between letter knowledge and phoneme awareness, which has been observed in earlier studies (Caravolas et al., 2001; Muter et al., 2004). In such a model, good RAN skills should lead to a stronger longitudinal relationship between these two constructs. The studies just described also cannot speak to the possibility that orthographic consistency begins to have language-specific effects on the relative importance of the predictors later in reading/spelling development; this issue would be better addressed by longitudinal studies of the growth of literacy. To our knowledge, only one direct cross-linguistic study of early growth in reading (described in what follows) has been published.

Early Growth of Alphabetic Literacy Skills

A clear and undisputed finding of cross-linguistic literacy research is that children learning to read in English attain the early milestones of reading ability more slowly than learners of consistent alphabetic orthographies. Seymour et al. (2003) demonstrated that children learning languages with relatively consistent orthographies, including Finnish and Spanish, reached mastery in fluent and accurate (>80%) word reading by the end of grade one, while learners of English attained the lowest levels of accuracy (34%). This pattern has been replicated in many smaller-scale studies (see Share, 2008; Ziegler & Goswami, 2005). However, few studies have directly compared the growth of reading skill longitudinally across languages using growth curve modeling techniques. Thus it is not known whether the typical differences between English and more consistent orthographies in children's rate of early reading development are also underpinned by different patterns of growth or by different predictors of growth in reading skill.

Growth curve modeling techniques can be informative about the rate at which skills are acquired, the shape of that growth—including growth spurts, periods of steady growth, and of growth deceleration—and about the growth trajectories of specific groups or individuals (e.g., Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Lervåg & Hulme, 2010). The few studies focusing on the earliest phases of learning report nonlinear growth patterns in English (Compton, 2003; Skibbe, Grimm, Bowles, & Morrison, 2012; Torgesen et al., 1999), as well as in Finnish (e.g., Leppänen, Niemi, Aunola, & Nurmi, 2004; Parrila, Aunola, Leskinen, Nurmi, & Kirby, 2005) and Dutch (Verhoeven & van Leeuwe, 2011).

Studies of early reading development in English (Skibbe et al., 2012) and Finnish (Leppänen et al., 2004; Parrila et al., 2005) show that although some growth in reading may occur among preschoolers, not surprisingly, a particularly rapid period of growth occurs when children begin to receive formal literacy instruction. Among US-English pupils, this seems to begin in kindergarten and continues into first grade, while among learners of the consistent Finnish orthography it seems confined mainly to first grade (in response to the start of systematic literacy instruction). These differences are compatible with the typical differences observed in cross-sectional comparisons between learners of English and learners of relatively consistent orthographies (e.g., Seymour et al., 2003).

In the only (to our knowledge) direct, cross-linguistic study of early reading development, Caravolas et al. (2013) compared growth of word reading in groups of children learning English, Czech, and Spanish (as in Caravolas et al., 2012). All groups were assessed twice yearly, from just before (Czech, Spanish) or coincidentally with (English) the start of formal schooling until the end of grade two. Parallel versions of a test of silent word reading efficiency were created for each language, and the groups were matched on reading ability at the start of the study. The analyses revealed that growth could be best described in two distinct phases. As expected, all three groups experienced a growth spurt at the start of formal reading instruction. However, while the English children had faster initial growth (when their Czech and Spanish counterparts were still in kindergarten), they exhibited a relatively slow and steady rate of growth over the next two years. In contrast, the Czech and Spanish children underwent a large growth spurt with the start of schooling, which was confined to the first grade, and during which they shot ahead of their English counterparts; their growth then stabilized and decelerated in second grade. As in Caravolas et al.'s earlier study (2012), the groups had experienced similar preschool and early grade school experiences as regards literacy practices and these were not likely to have influenced the preceding outcomes (see Caravolas et al., 2012, online supplement).

Further analyses examined which, if any, of the core foundation skills of phoneme awareness, letter knowledge, and RAN predicted growth curve components; that is, the initial word reading level (i.e., intercept), the rate of early growth (i.e., slope), and the acceleration of later growth (i.e., quadratic component). Consistent with their previous findings (Caravolas et al., 2012), all three foundation skills predicted individual differences in initial reading levels in all three languages. In addition, in all language groups, phoneme awareness and letter knowledge (but not RAN) predicted individual differences in the early rate of growth, while variations in RAN alone predicted variations in the rate of acceleration over the course of grade one. Growth during grade two was only predicted by earlier reading growth. The preceding cross-linguistic study is broadly consistent with previous single-language studies in confirming that exposure to formal literacy instruction coincides with an accelerated rate of growth, which, however, is steeper and shorter in the consistent orthographies than in English.

Summary

The accruing body of cross-linguistic research sheds light on the commonalities and differences in early literacy development in alphabetic orthographies. In line with the triple foundation model, the studies we have reviewed show that, when learners are compared during the same window of literacy development and are assessed on the same measures, their growth in literacy depends on the same skills in a similar way. The predictors of growth have the same relative importance within languages, suggesting the existence of the same cognitive architecture. While phoneme awareness and letter knowledge predict children's level of attainment in reading and spelling in the very early stages of learning, RAN seems more strongly related to the rate of change in reading development to the end of grade one. Despite the differential rates of growth in English relative to more consistent orthographies, it seems to be the case that from second grade onward, single-word reading is best predicted by prior reading ability—thus, all component skills now have only indirect influence on its further development.

We concur with Furnes and Samuelsson (2011) that in early development the similarities outnumber the differences across alphabetic orthographies and the same core skills seem to launch the alphabetic reading and spelling systems. What is more, the statistical models of the patterns of development in this phase seem to generalize across the orthographic consistency continuum. We take these conclusions one step further to suggest that the triple foundation model of alphabetic literacy provides an accurate explanation of the proximal cognitive causes of early reading and spelling development in alphabetic orthographies. Importantly, despite the markedly different rates of growth that are associated with letter–sound and sound–letter consistency, the evidence suggests that early growth in reading and spelling seems to have the same foundations across alphabetic orthographies.

Learning to Read and Spell in Nonalphabetic Writing Systems

In the following section we review the most relevant research on the cognitive processes that underlie literacy development in logographies (Chinese) and alphasyllabaries (Korean, Kannada). We summarize evidence from cross-sectional and, where available, longitudinal studies that have considered the relative importance of the learner variables in the proposed triple foundation model, namely phonological awareness, symbol set knowledge, and rapid naming speed. Their influence is considered in the context of two relevant orthographic variables: size of the symbol inventory and visual complexity of the symbols; these were not considered earlier given that the visually simple symbol inventories of alphabetic orthographies rarely exceed thirty to forty letters. How well do the data from nonalphabetic orthographies fit the triple foundation model account? We begin by briefly considering the case of Chinese, where most published research on literacy acquisition in nonalphabetic orthographies has been conducted.

Chinese Logography

Unlike English and other alphabetic orthographies, where letters represent phoneme-sized units, the basic writing unit in Chinese, the *hanzi*, stands for a syllable and a morpheme simultaneously. Chinese is an extensive orthography requiring knowledge of 3,000 to 4,000 characters for full literacy (Li, Shu, McBride-Chang, Liu, & Peng, 2012). Chinese characters are visually complex, with a high number of strokes confined to square-shaped forms. Most Chinese characters (80% to 90% in modern Chinese; Shu, Chen, Anderson, Wu, & Xuan, 2003) are made of two components (called *radicals*). *Semantic radicals* cue meaning to some extent, but provide no information about a composite character's pronunciation; this can be sometimes derived from the *phonetic radical*, but this component at best provides partial phonetic information (Shu et al., 2003; Zhou, 1978).

With respect to phoneme awareness as a component skill of learning to read and write words in Chinese, on balance it does not emerge as a strong early predictor. McBride-Chang, Bialystok, Chong, and Li (2004) found that Chinese character recognition in kindergarten and grade one was best predicted by syllable awareness among children, whether they were learning to read Chinese by a phonemic coding system, *pinyin* (Xian, Mainland China), or the look and say method (Hong Kong). Awareness of individual phonemes did not predict unique variance in Chinese character recognition in either group of children. It was, however, a significant predictor of English word recognition, over and above syllable awareness, in a group of Canadian beginning readers.

Beyond phonology, morphological awareness—the awareness of and access to the smallest units of meaning in words—has been assessed as another reliable metalinguistic predictor of reading in Chinese. Due to its unique word-compounding structure, and to the abundance of homophones, which render phonological information unreliable for Chinese character identification (Shu, McBride-Chang, Wu, & Liu, 2006), it has been suggested that the hanzi orthographic unit elicits a need for good awareness of the corresponding morphemic units in spoken language, perhaps over and above awareness of phonological units such as syllables and phonemes (e.g., Perfetti & Dunlap, 2008; Ziegler & Goswami, 2005). Several studies by McBride-Chang and colleagues convey this most convincingly. For example, McBride-Chang, Shu, Zou, Wat, and Wagner (2003) found that morpheme awareness (a construct estimated from a measure of children's sensitivity to homophone morphemes and a test assessing children's word-compounding skills) accounted for significant unique variation in Chinese character reading. The finding held true among kindergarteners as well as second graders even after controlling for age and other reading-related abilities (verbal skills, visual skills, phonological awareness, and rapid naming skills). In a subsequent multilanguage study with second graders from Hong Kong, Beijing, Korea, and the United States, McBride-Chang et al. (2005) investigated whether language- and script-related differences (e.g., differences in the mapping units of the written language, literacy instruction differences between Chinese cultures) affected the strength of association of morphological and phonological awareness with word recognition. Although both constructs were similarly associated with vocabulary knowledge for all groups, their relative contribution to word recognition differed. First, while morphological awareness was necessary for the prediction of Chinese and Korean word recognition, it did not fit well in the model predicting English word recognition. Phonological awareness, on the other hand, was a significant predictor of Korean and English word recognition but was not essential for a good fit of the Chinese model (Beijing, Hong Kong).

The second component of the triple foundation model should equate to a skill similar to letter knowledge in alphabetic systems—that is, to the knowledge of the functional symbols of Chinese orthography, the hanzi characters, and presumably of the attributes and constraints on their visual and linguistic composition (e.g., permissible sequences and positions of strokes, radicals, and compound characters). However, the earlier literature on Chinese literacy development rarely isolated character knowledge as a separate foundational skill, and frequently grouped it alongside visual processing skills (a term covering a broad range of skills such as visual memory ability and visual discrimination ability). While it is clear that visual skills correlate with reading ability in Chinese (Ho & Bryant, 1997; Huang & Hanley, 1995), a unique association over and above other predictors has not always been found. A potential explanation for these discrepant findings is that, over and above important methodological variations, visual processing skills may be predictive of reading performance only during earlier stages of reading development because Chinese reading development may progress from a visual (logographic) to a phonological phase (Ho & Bryant, 1997; Siok & Fletcher, 2001). More recent longitudinal investigations of learning to read and write Chinese (e.g., Tong, McBride-Chang, Shu, & Wong, 2009; Yeung et al., 2011) have identified character knowledge (e.g., knowledge of hanzis' internal structure, or the functions of their radical components) as a separable component of reading and have considered its relative importance together with phonological awareness, morphological awareness, and RAN. These studies demonstrated that, while phonological awareness measured by syllable and phoneme deletion was not uniquely related to reading with other variables statistically controlled, orthographic character knowledge, RAN, and morphological awareness were all associated with reading measured concurrently and longitudinally. It should be noted that phonological and morphological awareness measures were moderately correlated with each other, which may explain why syllable awareness was not uniquely associated with reading. Similar to reading, word writing to dictation was concurrently predicted both from morphological awareness and orthographic (character) knowledge. In longitudinal predictions, when the autoregressor (earlier word writing) was controlled, only RAN accounted for unique variance in Chinese word writing.

Turning to the third component of the triple foundation model, RAN seems to emerge the most systematically as an early predictor of reading success as well as of word writing in Chinese (e.g., Georgiou, Parrila, & Liao, 2008; McBride-Chang & Ho, 2000; Tong et al., 2009). McBride-Chang and Kail (2002), for example, compared the influence of RAN (among other cognitive processes) on Chinese character recognition and English word recognition and showed that paths to reading from RAN were significant for US and Hong Kong beginning readers (see also McBride-Chang & Ho, 2000). The work of Georgiou et al. (2008) with older (fourth-grade) children speaking English (Canadian), Chinese (Taiwanese), or Greek (Cypriot) provides further evidence of the robust RAN–reading relationship across languages. Interestingly, the latter study showed no significant difference across languages in the strength of this association when similar tasks and outcome measures were used, although sample size limitations warrant a cautious interpretation.

Alphasyllabaries

Unlike the logographic approach of Chinese, the Korean language has adopted an alphasyllabic writing system called hangul (detailed descriptions of the Korean alphasyllabary are provided by Taylor & Taylor, 1995, and Kessler & Treiman, this volume) which represents phonemes arranged into syllable-sized characters. The characters can be decomposed into their constituent letters, of which there are twenty-four in the alphabet. However, early literacy training emphasizes syllable recognition, and anecdotal evidence indicates that Korean letter names and sound-letter relations are not explicitly taught in Korean (Cho, McBride-Chang, & Park, 2008; Kim, 2007). The letters are arranged in a nonlinear fashion, left-to-right and top-to-bottom in a square structure, or block, creating clear syllable boundaries in print. While there are only twenty-four letters in the Korean alphabet, the hangul characters number in the thousands. Korean has relatively high grapheme–phoneme consistency, although it preserves some morphophonemic features.

The relatively small number of predictor and longitudinal studies of early reading and writing development in Korean means that the findings must be interpreted with caution as regards their fit to the triple foundation model. To date, the most consistent finding is that awareness of the phonological units that mirror the mapping units of the Korean script (syllables and phonemes) predicts both reading and word writing development. For example, in a cross-sectional study by Cho and McBride-Chang (2005), kindergarten and second-grade children's performance on a phoneme deletion and a syllable deletion task contributed strongly and uniquely to concurrent word reading scores across age samples. Additional support for the role of both levels of phonological awareness in Korean hangul reading comes from Cho et al.'s (2008) study of 4- and 5-year-old kindergarteners. Interestingly, while regular word recognition was uniquely predicted by children's phoneme and syllable awareness, irregular word recognition was best explained by an oral morphological awareness task (lexical compounding) above and beyond the other variables. This suggests that different sublexical units may be invoked for regular and irregular word recognition processes in Korean.

The findings suggest that at, least for Korean, morphological awareness in the form of lexical compounding may be important for early irregular word reading, separable from the effects of phonological awareness. It may be that knowledge of the structure of the morphemes in one's language is particularly helpful for early irregular word recognition in Korean because such knowledge facilitates children's use of analogies to map a morpheme from a known word to a new word. Morphemic knowledge also may affect phonological transformations across words. For example, some morphemes may change in pronunciation when combined with other derivational forms (e.g., in English, *know/knowledge*; *photo/photography*). These transformational processes involving morphology are unnecessary for regular Korean hangul words, because phonological processes are all that is needed to pronounce the word. However, additional semantic knowledge may be essential for learning exception words.

Two studies by Kim (2009, 2010) speak directly to the hypothesis that letter-name knowledge is an essential building block of literacy acquisition in Korean. While both phoneme and syllable awareness made unique contributions to the literacy skills of 4- and 5-year-old Korean preschoolers (word reading, pseudoword reading, spelling)—replicating the findings of Cho et al. (2005, 2008)—hangul letter-name knowledge explained significantly higher amounts of variability in all three literacy outcomes, above and beyond phonological awareness and age. Turning to spelling skill, individual differences in the mastery of Korean letter names were shown to be highly predictive of individual differences in conventional spelling skill in Korean, above and beyond the remaining critical variables, namely, morphological, phonological, and orthographic awareness (Kim, 2010). Rapid automatized naming was included in the studies of Cho and colleagues, and it accounted for unique variance in reading in the later (2008) but not the earlier (2005) study, even though it correlated with performance on phonological awareness tasks in the latter.

Further preliminary evidence relevant to the triple foundation model comes from Kannada, an alphasyllabic script of South Asia that represents phonemes in syllable blocks (*aksharas*) (for details on Kannada, see Nag, Treiman, & Snowling, 2010). This language is spoken by 40,000,000 people in the state of Karnataka in South India. Briefly, the visually complex nature of Kannada aksharas (combination of base symbols with diacritic markers to the top or bottom of the base), their nonlinear spatial layout (e.g., postconsonantal vowels can be placed either to the right, top, or bottom of the initial consonant), and the extensive orthographic registry of over 400 hundred aksharas are three major aspects in which Kannada differs from alphabetic systems (Nag, 2007). Similar to Korean hangul, Kannada is consistent (regular and transparent). Nag and Snowling (2012) examined the concurrent association between Kannada reading accuracy/rate and phonological (phoneme and syllable) awareness, akshara knowledge, and RAN in a sample of 9- to 12-year-old children in fourth to sixth grades. Reading accuracy was best predicted by children's askhara knowledge—which explained approximately half of the variance in reading accuracy—with further contributions from both syllable and phoneme awareness, as well as RAN. Reading rate in Nag and Snowling's (2012) study was not significantly predicted by syllable awareness (as anticipated according to the triple foundation model), but the contribution of phoneme awareness and RAN were highly significant.

Summary

Research on nonalphabetic orthographies is dominated by studies of Chinese. This body of work suggests that the triple foundation model, broadly specified, accounts well for the critical basis of Chinese literacy. Rapid automatized naming clearly emerges as a reliable predictor skill, and more recent research also highlights the importance of character knowledge. As regards the metalinguistic component (corresponding to phoneme awareness in alphabetic orthographies), the construct needs to be broadened to better fit the dominant mapping units of the written language, namely morphemes. Thus, awareness of morphemes seems to be the crucial metalinguistic component for Chinese, although some studies suggest that phonological awareness—in particular syllable awareness—may also play a role, albeit weaker and perhaps less persistent (e.g., McBride-Chang et al., 2004, 2005). In the interesting alphasyllabic systems, the findings replicated those of the alphabetic orthographies. Korean letter knowledge, and Kannada akshara knowledge, as well as RAN were found to be robust predictors of early reading and spelling. Interestingly, the metalinguistic skills that emerged as important predictors included both phoneme awareness and syllable awareness—mirroring the corresponding duality (phonemic and syllabic) of the constituents of the writing units in these languages. Thus, the studies from these languages suggest that the triple foundation model provides an adequate account of the underpinnings of early literacy, with the proviso of a broadening of the phoneme awareness component to a phoneme and syllable awareness component. A suggestion arising from several studies of the alphasyllabaries and recent studies of Chinese was that character knowledge may be the most critical skill contributing to the triple foundation. This finding, if replicated in direct tests of this hypothesis, would further suggest that an orthography-related variable influencing the initial learning architecture may be the size of the symbol set to be learned in the extensive orthographies (Nag, 2007). However, it remains to be seen whether the potentially heavier weighting of character knowledge in the logographic systems is proportionately similar to the relative weighting of letter knowledge in alphabetic orthographies, where it is also probably the most stable and robust predictor of early literacy.

It is also clear that what are broadly termed visual skills also play a role in learning to read and write, and these are probably more important in logographic and alphasyllabic systems, which contain visually complex logographs and syllabographs (Nag & Snowling, 2012; see also Perfetti & Dunlap, 2008). However, our review showed that the evidence regarding visual skills is rather mixed and may in some studies be conflated with character knowledge. We suggest that visual skills, broadly defined, may function much like broad language skills (e.g., vocabulary knowledge, syntactic skills, and listening comprehension) in that they are more distally related to early reading and writing skills. As such, they are often found to correlate with but rarely emerge as unique (core) predictors of early word reading and spelling.

Conclusion

Studies of early development of alphabetic literacy point to three key cognitive abilities that learners need to bring to the reading and spelling acquisition tasks: knowledge of the letters of the alphabet, phoneme awareness, and rapid naming of visually presented stimuli. We refer to this as the triple foundation model, building on Byrne's (1998) dual foundation model. In this review, we explored the evidence within and across writing systems of the extent to which this model generalizes as a universal architecture for the initial word reading and writing learning process. Within alphabetic orthographies, where a relatively large body of research has been carried out, we found strong support for the generality of the model. Not only were letter knowledge, phoneme awareness, and RAN consistently found to predict early reading and spelling across languages, as was the case in English, but also in well-controlled studies (e.g., Caravolas et al., 2012, 2013; Furnes & Samuelsson, 2011) the models of the predictive relationships with the literacy measures were found to be essentially the same across languages across the consistency spectrum.

Although this chapter has reviewed research evidence from a variety of languages and orthographies, in the grand scheme of the world's languages, they represent but a small sample (Share, 2008). Nevertheless, the findings across the languages and writing systems are quite consistent to date. They suggest that all children must bring a very similar set of three core cognitive skills to the task of learning to read and write words, regardless of the orthography they are to learn. We propose that these three foundation skills may generalize across writing systems and orthographies as suggested by the triple foundation model. Thus, individual differences in the initial growth of word reading and writing skills are determined by variations in the ability to learn the functional symbol set of their orthography (for example the hanzi characters of Chinese, the akshara characters of Kannada, or the alphabet letters of English), the ability to have conscious awareness of the sublexical units used to encode the language in print (e.g., morphemes, syllables, phonemes), and the ability to connect these two representations rapidly and efficiently, as estimated by measures of RAN. In future research we look forward to direct and full tests of this hypothesis with the aim of teasing apart language-specific phenomena from those that are language general, including in orthographies that have not yet been examined.

References

Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly, 26,* 49–66.

Borgwaldt, S. R., Hellwig, F. M., & De Groot, A. M. (2005). Onset entropy matters—letter-to-phoneme mappings in seven languages. *Reading and Writing: An Interdisciplinary Journal, 18,* 211–229.

Bowers, P. G., & Wolf, M. (1993). Theoretical links between naming speed, precise timing mechanisms and orthographic skill in dyslexia. *Reading and Writing: An Interdisciplinary Journal, 5,* 69–85.

Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., . . . Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry, 49,* 422–432.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read—A causal connection. *Nature, 301,* 419–421.

Bruck, M., Genesee, F., & Caravolas, M. (1997). A cross-linguistic study of early literacy acquisition. In B. A. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 145–162). Mahwah, NJ: Erlbaum.

Burt, J. S. (2006). What is orthographic processing and how does it relate to word identification in reading? *Journal of Research in Reading, 29,* 400–417.

Byrne, B. (1998). *The foundation of literacy: The child's acquisition of the alphabetic principle*. Hove, England: Psychology Press.

Caravolas, M. (2005). The nature and causes of dyslexia in different languages. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 336–356). Oxford, England: Blackwell.

Caravolas, M., & Bruck, M. (1993). The effect of oral and written language input on children's phonological awareness: A cross-linguistic study. *Journal of Experimental Child Psychology, 55,* 1–30.

Caravolas, M., Hulme, C., & Snowling, M. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language, 45,* 751–774.

Caravolas, M., Kessler, B., Hulme, C., & Snowling, M. (2005). Effects of orthographic consistency, frequency, and letter knowledge on children's vowel spelling development. *Journal of Experimental Child Psychology, 92,* 307–321.

Caravolas, M., & Landerl, K. (2010). The influences of syllable structure and reading ability on the development of phoneme awareness: A longitudinal, cross-linguistic study. *Scientific Studies of Reading, 14,* 464–484.

Caravolas, M., Lervåg, A., Defior, S., Málková, G. S., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science, 24,* 1398–1407.

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litaviský, M., Onochie-Quintanilla, E., . . . Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science, 23,* 678–686.

Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology, 89,* 631–644.

Cho, J.-R., & McBride-Chang, C. (2005). Correlates of Korean hangul acquisition among kindergartners and second graders. *Scientific Studies of Reading, 9,* 3–16.

Cho, J.-R., McBride-Chang, C., & Park, S. (2008). Phonological awareness and morphological awareness: Differential associations to regular and irregular word recognition in early Korean Hangul readers. *Reading and Writing: An Interdisciplinary Journal, 21,* 255–274.

Compton, D. L. (2003). Modeling the relationship between growth in rapid naming speed and growth in decoding skill in first-grade children. *Journal of Educational Psychology, 95,* 225–239.

Defior, S., Martos, F., & Cary, L. U. Z. (2002). Differences in reading acquisition development in two shallow orthographies: Portuguese and Spanish. *Applied Psycholinguistics, 23,* 135–148.

de Jong, P. F., & van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology, 91,* 450–476.

de Jong, P. F., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading, 6,* 51–77.

Denckla, M. B., & Rudel, R. G. (1976). Rapid "automatized"naming (RAN): Dyslexia differentiated from other learning disabilities. *Neuropsychologia, 14,* 471–479.

Durgunoğlu, A. Y., & Öney, B. (1999). A cross-linguistic comparison of phonological awareness and word recognition. *Reading and Writing: An Interdisciplinary Journal, 11,* 281–299.
Ehri, L. C. (1997). Learning to read and learning to spell are one and the same, almost. In C. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory and practice* (pp. 237–269). Mahwah, NJ: Erlbaum.
Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading, 21,* 5–19.
Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88,* 3–17.
Furnes, B., & Samuelsson, S. (2009). Preschool cognitive and language skills predicting kindergarten and grade 1 reading and spelling: A cross-linguistic comparison. *Journal of Research in Reading, 32,* 275–292.
Furnes, B., & Samuelsson, S. (2011). Phonological awareness and rapid automatized naming predicting early development in reading and spelling: Results from a cross-linguistic longitudinal study. *Learning and Individual Differences, 21,* 85–95.
Georgiou, G. K., Parrila, R., & Liao, C. H. (2008). Rapid naming speed and reading across languages that vary in orthographic consistency. *Reading and Writing: An Interdisciplinary Journal, 21,* 885–903.
Georgiou, G. K., Torppa, M., Manolitsis, G., Lyytinen, H., & Parrila, R. (2012). Longitudinal predictors of reading and spelling across languages varying in orthographic consistency. *Reading and Writing: An Interdisciplinary Journal, 25,* 321–346.
Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65,* 41–57.
Hayes, H., Treiman, R., & Kessler, B. (2006). Children use vowels to help them spell consonants. *Journal of Experimental Child Psychology, 94,* 27–42.
Ho, C. S.-H., & Bryant, P. (1997). Phonological skills are important in learning to read Chinese. *Developmental Psychology, 33,* 946–951.
Huang, H. S., & Hanley, J. R. (1995). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition, 54,* 73–98.
Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological Science, 23,* 572–577.
Hulme, C., Caravolas, M., Málková, G., & Brigstocke, S. (2005). Phoneme isolation ability is not simply a consequence of letter-sound knowledge. *Cognition, 97,* B1–B11.
Kessler, B., & Caravolas, M. (2011). *Weslalex: West Slavic lexicon of child-directed printed words*. http://spell.psychology.wustl.edu/weslalex/
Kessler, B., & Treiman, R. (2001). Relationships between sounds and letters in English monosyllables. *Journal of Memory and Language, 44,* 592–617.
Kim, Y.-S. (2007). Phonological awareness and literacy skills in Korean: An examination of the unique role of body-coda units. *Applied Psycholinguistics, 28,* 67–93.
Kim, Y.-S. (2009). The foundation of literacy skills in Korean: The relative contribution of letter-name knowledge and phonological awareness and their interrelationship in Korean. *Reading and Writing: An Interdisciplinary Journal, 22,* 907–931.
Kim, Y. S. (2010). Componential skills in early spelling development in Korean. *Scientific Studies of Reading, 14,* 137–158.
Kirby, J. R., Georgiou, G. K., Martinussen, R., & Parrila, R. (2010). Naming speed and reading: From prediction to instruction. *Reading Research Quarterly, 45,* 341–362.
Kirby, J. R., Parrila, R. K. & Pfeiffer, S. L. (2003). Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology, 95,* 453–464.
Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A German-English comparison. *Cognition, 63,* 315–334.
Leppänen, U., Niemi, P., Aunola, K., & Nurmi, J. E. (2004). Development of reading skills among preschool and primary school pupils. *Reading Research Quarterly, 39,* 72–93.
Lervåg, A., Bråten, I., & Hulme, C. (2009). The cognitive and linguistic foundations of early reading development: A Norwegian latent variable longitudinal study. *Developmental Psychology, 45,* 764–781.
Lervåg, A., & Hulme, C. (2009). Rapid automatized naming (RAN) taps a mechanism that places constraints on the development of early reading fluency. *Psychological Science, 20,* 1040–1048.
Lervåg, A., & Hulme, C. (2010). Predicting the growth of early spelling skills: Are there heterogeneous developmental trajectories? *Scientific Studies of Reading, 14,* 485–513.
Li, H., Shu, H., McBride-Chang, C., Liu, H., & Peng, H. (2012). Chinese children's character recognition: Visuo-orthographic, phonological processing and morphological skills. *Journal of Research in Reading, 35,* 287–307.
Mann, V., & Wimmer, H. (2002). Phoneme awareness and pathways into literacy: A comparison of German and American children. *Reading and Writing: An Interdisciplinary Journal, 15,* 653–682.
McBride-Chang, C., Bialystok, E., Chong, K. K. Y., & Li, Y. (2004). Levels of phonological awareness in three cultures. *Journal of Experimental Child Psychology, 89,* 93–111.
McBride-Chang, C., Cho, J.-R., Liu, H., Wagner, R. K., Shu, H., Zhou, A., . . . Muse, A. (2005). Changing models across cultures: Associations of phonological awareness and morphological structure awareness with vocabulary and word recognition in second graders from Beijing, Hong Kong, Korea, and the United States. *Journal of Experimental Child Psychology, 92,* 140–160.
McBride-Chang, C., & Ho, C. S.-H. (2000). Developmental issues in Chinese children's character acquisition. *Journal of Educational Psychology, 92,* 50–55.
McBride-Chang, C., & Kail, R. (2002). Cross-cultural similarities in the predictors of reading acquisition. *Child Development, 73,* 1392–1407.
McBride-Chang, C., Shu, H., Zhou, A. B., Wat, C. P., & Wagner, R. K. (2003). Morphological knowledge uniquely predicts young children's Chinese character recognition. *Journal of Educational Psychology, 95,* 743–751.
Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology, 40,* 665–681.

Nag, S. (2007). Early reading in Kannada: The pace of acquisition of orthographic knowledge and phonemic awareness. *Journal of Research in Reading, 30,* 1–17.

Nag, S., & Snowling, M. (2012). Reading in an alphasyllabary: Implications for a language-universal theory of learning to read. *Scientific Studies of Reading, 16,* 404–423.

Nag, S., Treiman, R., & Snowling, M. J. (2010). Learning to spell in an alphasyllabary: The case of Kannada. *Writing Systems Research, 2,* 41–52.

Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: Why there's nothing special about the perception of orthography. *Psychological Review, 119,* 517–545.

Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology, 63,* 427–452.

Pacton, S., Fayol, M., & Perruchet, P. (2002). The acquisition of untaught orthographic regularities in French. In L. Verhoeven, C. Erlbro & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 121–137). Dordrecht, the Netherlands: Kluwer.

Parrila, R., Aunola, K., Leskinen, E., Nurmi, J. E., & Kirby, J. R. (2005). Development of individual differences in reading: Results from longitudinal studies in English and Finnish. *Journal of Educational Psychology, 97,* 299–319.

Patel, T. K., Snowling, M. J., & de Jong, P. F. (2004). A cross-linguistic comparison of children learning to read in English and Dutch. *Journal of Educational Psychology, 96,* 785–797.

Peereman, R., & Content, A. (1999). Lexop: A lexical database with orthography-phonology statistics for French monosyllabic words. *Behavior Research Methods, Instruments, and Computers, 31,* 376–379.

Peereman, R., Lété, B., & Sprenger-Charolles, L. (2007). Manulex-infra: Distributional characteristics of grapheme-phoneme mappings, and infralexical and lexical units in child-directed written material. *Behavior Research Methods, 39,* 579–589.

Perfetti, C. A., & Dunlap, S. (2008). Learning to read: General principles and writing system variations. In K. Koda & A. Zehler (Eds.), *Learning to read across languages* (pp. 13–38). Mahwah, NJ: Erlbaum.

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The lexical constituency model: Some implications of research on Chinese for general theories of reading. *Psychological Review, 112,* 43–59.

Price, C. J., & McCrory, E. (2005). Functional brain imaging studies of skilled reading and developmental dyslexia. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 473–496). Oxford, England: Blackwell.

Rittle-Johnson, B., & Siegler, R. S. (1999). Learning to spell: Variability, choice, and change in children's strategy use. *Child Development, 70,* 332–348.

Samara, A., & Caravolas, M. (2014). Statistical learning of novel graphotactic constraints in children and adults. *Journal of Experimental Child Psychology, 121,* 137–155.

Samuelsson, S., Byrne, B., Quain, P., Wadsworth, S., Corley, R., DeFries, J., . . . Olson, R. (2005). Environmental and genetic influences on prereading skills in Australia, Scandinavia, and the United States. *Journal of Educational Psychology, 97,* 705–722.

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology, 96,* 265–282.

Seidenberg, M. S. (2011). Reading in different writing systems: One architecture, multiple solutions. In P. McCardle, J. Ren, & O. Tzeng (Eds.), *Dyslexia across languages: Orthography and the gene–brain–behavior link* (pp. 146–168). Baltimore, MD: Brookes.

Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94,* 143–174.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134,* 584–615.

Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development, 74,* 27–47.

Shu, H., McBride-Chang, C., Wu, S., & Liu, H. (2006). Understanding Chinese developmental dyslexia: Morphological awareness as a core cognitive construct. *Journal of Educational Psychology, 98,* 122–133.

Siok, W. T., & Fletcher, P. (2001). The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition. *Developmental Psychology, 37,* 886–899.

Skibbe, L. E., Grimm, K. E., Bowles, R. P., & Morrison, F. J. (2012). Literacy growth in the academic year versus summer from preschool through second grade: Differential effects of schooling across four skills. *Scientific Studies of Reading, 16,* 141–165.

Taylor, I., & Taylor, M. (1995). *Writing and literacy in Chinese, Korean, and Japanese*. Amsterdam, the Netherlands: Benjamins.

Tong, X., McBride-Chang, C., Shu, H., & Wong, A. M. (2009). Morphological awareness, orthographic knowledge, and spelling errors: Keys to understanding early Chinese literacy acquisition. *Scientific Studies of Reading, 13,* 426–452.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Burgess, S., & Hecht, S. (1997). Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second-to fifth-grade children. *Scientific Studies of Reading, 1,* 161–185.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in children with phonological processing difficulties: Group and individual responses to instruction. *Journal of Educational Psychology, 81,* 579–593.

Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. New York, NY: Oxford University Press.

Treiman, R., Berch, D., & Weatherston, S. (1993). Children's use of phoneme-grapheme correspondences in spelling: Roles of position and stress. *Journal of Educational Psychology, 85,* 466–477.

Vaessen, A., Bertrand, D., Tóth, D., Csépe, V., Faísca, L., Reis, A., & Blomert, L. (2010). Cognitive development of fluent word reading does not qualitatively differ between transparent and opaque orthographies. *Journal of Educational Psychology, 102,* 827–842.

Vaessen, A., Gerretsen, P., & Blomert, L. (2009). Naming problems do not reflect a second independent core deficit in dyslexia: Double deficits explored. *Journal of Experimental Child Psychology, 103,* 202–221.

Verhoeven, L., & van Leeuwe, J. (2011). Role of gender and linguistic diversity in word decoding development. *Learning and Individual Differences, 21,* 359–367.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30,* 73–87.

Wesseling, R., & Reitsma, P. (2000). The transient role of explicit phonological recoding for reading acquisition. *Reading and Writing: An Interdisciplinary Journal, 13,* 313–336.

Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development: Word recognition in English and German children. *Cognition, 51,* 91–103.

Wimmer, H., Landerl, K., Linortner, R., & Hummer, P. (1991). The relationship of phonemic awareness to reading acquisition: More consequence than precondition but still important. *Cognition, 40,* 219–249.

Wimmer, H., Mayringer, H., & Landerl, K. (2000). The double-deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of Educational Psychology, 92,* 668–680.

Wimmer, H., & Schurz, M. (2010). Dyslexia in regular orthographies: Manifestation and causation. *Dyslexia, 16,* 283–299.

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91,* 415–438.

Yeung, P. S., Ho, C. S. H., Chik, P. P. M., Lo, L. Y., Luan, H., Chan, D. W. O., & Chung, K. K. H. (2011). Reading and spelling Chinese among beginning readers: What skills make a difference? *Scientific Studies of Reading, 15,* 285–313.

Zhou, Y. G. (1978). To what extent are the "phonetics" of present-day Chinese characters still phonetic? *Zhongguo Yuwen, 146,* 172–177.

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., . . . Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science, 21,* 551–559.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin, 131,* 3–29.

Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments, & Computers, 28,* 504–515.

CHAPTER

23 Children's Reading Comprehension and Comprehension Difficulties

Jane V. Oakhill, Molly S. Berenhaus, *and* Kate Cain

Abstract

This chapter considers the normal development of children's reading comprehension, as well as individual differences and specific difficulties related to children's reading comprehension. Most of the studies in this area have been carried out with children who are learning to read in an alphabetic orthography, and this chapter reflects that bias. The chapter outlines the development of various processes that are related to reading comprehension in the early school years. The chapter also considers the relationship between these processes and reading comprehension ability. The chapter concludes with a consideration of the implications for research and practice during the early school years.

Key Words: reading comprehension, vocabulary, inference, comprehension monitoring, text structure, understanding

Reading comprehension requires the use of many different skills, strategies, knowledge bases, and cognitive processes. In most cases, beginning readers will already possess competent oral language skills and will be able to apply them to reading comprehension. Although skills in spoken language understanding serve as a foundation for reading comprehension, they do not in themselves guarantee success: A substantial minority of young readers have problems with reading comprehension despite the fact that they develop good word reading skills. Because their word reading appears to be intact, these children have a specific comprehension problem (in both reading and listening comprehension). In this chapter we discuss both reading comprehension development and the problems experienced by poor comprehenders. Longer reviews of the development of comprehension and comprehension-related skills can be found in Oakhill and Cain (2007) and Cain and Oakhill (2007). The chapter ends with a brief overview of what can be done to support reading comprehension in early readers and to improve reading comprehension in those children who have such difficulties.

What Is Successful Text Comprehension?

We first consider what constitutes successful text comprehension, and we will use the criteria outlined here as a framework within which the relevance and importance of different components we outline later in the chapter can be understood. Understanding a text can be viewed as a constructive process that results in a coherent and integrated representation of the state of affairs described. This representation is often referred to as a *mental model* (e.g., Johnson-Laird, 1983) or a *situation model* (Kintsch, 1998) (see O'Brien and Cook, this volume, for a review of models of adult text comprehension). This text representation will be very similar whether the text is read aloud to the comprehender (listening comprehension) or read by the comprehender (reading comprehension), although of course reading comprehension requires the additional skill of being able to recognize the written words.

In order to construct such a representation, the reader needs to engage in a number of processes. Readers need to access the meanings of individual words, integrate the meanings of the sentences and paragraphs, and identify the key ideas or themes of the text. To understand a story, for instance, the main characters and their motives, and the plot of the story, need to be determined. Inferences must be made to fill in information that is left implicit in the text. Sentences need to be integrated across the text, and the information in the text needs to be integrated with relevant general knowledge. In addition, skilled readers reflect on what they are reading (e.g., what the main points are and whether or not their emerging model of the text makes sense). By monitoring their comprehension in this way, readers can identify when they need to make an inference to fill in missing details and can take corrective action, such as rereading, when comprehension fails. In efficient text comprehension, it is likely that many of these processes go on in parallel. The focus in this chapter is on the development of the aspects of language that are associated with the construction of this representation of a text's meaning, namely, vocabulary and syntax, inference-making, comprehension monitoring, and understanding and use of text structure and the cognitive processes that support them, such as working memory.

The Development of Reading Comprehension and Problems With Comprehension

In this section we outline the main component skills, strategies, and types of knowledge that are important for reading comprehension and discuss how they develop. As mentioned earlier, children are typically very proficient language users even before they start learning to read. Thus, once word-decoding skills have been mastered, most children should be able to use their existing spoken language comprehension skills to understand texts. Indeed, once children have learned to decode words, there is a strong association between listening comprehension and reading comprehension (measured, for instance, by ability to answer questions after listening to the passages from a reading comprehension test, rather than reading them). This correlation tends to increase with age, as children become more proficient decoders (Catts, Hogan, & Adlof, 2005). The *simple view* of reading (Gough & Tunmer, 1986; Hoover & Gough, 1990) proposes that reading comprehension is the product of word decoding and listening comprehension. Nevertheless, despite the importance of oral language skills for reading, there are substantial differences between everyday spoken interactions and text comprehension, because, for example, the context of spoken language provides many cues to meaning (e.g., facial expression, tone of voice, and intonation patterns of the speaker). Such cues are not present when a text is being read. Moreover, the language of written texts is more formal and more complex than that of oral communication (Garton & Pratt, 1998). In particular, written language typically makes use of syntactic constructions and vocabulary that may not be familiar to young children (see Scott, 2004), and a text cannot be interrogated in the way that a speaker can. Thus, beginner readers in particular might have comprehension problems that are specific to written texts (even when they listen to them) because they are not familiar with the language of books, particularly if they have not been read to regularly before they start school. Typical comparisons of reading and listening comprehension use texts from a standardized test (i.e., texts that conform to the conventions of written, not spoken, language), which are either read to the children or which they are required to read themselves. Such comparisons show very similar levels of reading and listening performance in comprehension beyond the initial stages of word decoding (Cain, Oakhill, & Bryant, 2000; Stothard & Hulme, 1992).

A further problem for beginning readers is that they are likely to find word decoding quite demanding and may not have enough spare cognitive capacity for essential comprehension processes. Thus as children get older and their decoding skills improve, they are likely to have more resources to devote to comprehension (see also the chapter by Goldman & Snow, this volume, which addresses the challenges of postprimary reading). Indeed, it has been shown that although early reading comprehension skills may be limited by word reading, in the later primary school years other factors come into play (e.g., Curtis, 1980). Many of the comprehension processes discussed in this chapter are common to comprehension of both written and spoken texts.

The simple view emphasizes the important contribution of general language comprehension for overall reading comprehension. In this vein, a number of studies (e.g., Cutting, Materek, Cole, Levine, & Mahone, 2009; Goff, Pratt, & Ong, 2005) have explored the relation between language skills (generally a composite measure comprising, for example, vocabulary and syntax, sometimes together with inference skills or listening comprehension) and reading comprehension development or difficulties. However, it is important to understand, for both theoretical and practical reasons, how these different aspects of language contribute to reading comprehension in children and how they support each other. In terms of practical implications, it is important to understand how different aspects of language and other cognitive processes support comprehension during development and which of these processes might need specific instruction. However, in order to make such recommendations for instruction, it is crucial to understand more about the relation between different processes and reading comprehension.

Many different abilities will correlate with comprehension skill, but not all will be causally implicated in comprehension development and improvement. In order to identify which processes are causally linked to comprehension, three designs have been used: longitudinal studies (for a justification of the use of longitudinal studies to infer causality, see de Jong & van der Leij, 2002; Oakhill & Cain, 2012), training studies, and comprehension-age match studies. The logic behind training studies is fairly obvious: If a skill believed to be causally implicated in reading comprehension is trained, and if there are concomitant increases in comprehension skill relative to a control group, then it can be concluded that the improvement in the trained skill caused the improvement in comprehension. The logic behind comprehension-age match studies is more complex. In such studies, a group of poor comprehenders (for their age) is compared not only with a same-age group of good comprehenders but also with a younger group of average comprehenders who have the same absolute level of comprehension as the older poor comprehenders. The crucial comparison in this design is between the poor comprehenders and the younger average (for their age) comprehenders. If the younger children do better on some language-related task than the poor comprehenders (with the same absolute comprehension level), then this difference cannot occur as a result of better absolute comprehension in one group. Thus the causal link in the opposite direction (from performance on the language task to better comprehension ability) is likely, though not proven, and the nature of the link can be investigated further using other methods (such as training studies).

In the following sections, we consider the evidence in relation to specific processes. For each process we consider first the development of the process and then whether good and poor comprehenders differ on the process. Where there is evidence, we also consider whether there is evidence for a causal link between the process and reading comprehension.

Word- and Sentence-Level Skills

Word- and sentence-level meaning serves as a fundamental basis for the construction of text meaning. Some studies have explored the relation between word-level semantic (typically vocabulary) skills and reading comprehension; others have explored primarily syntactic skills, and others have explored both. This section covers evidence from studies with these different foci.

VOCABULARY KNOWLEDGE

Vocabulary knowledge is one of the best concurrent predictors of reading comprehension ability (Carroll, 1993; Thorndike, 1973). For instance, Thorndike found correlations of between .66 and .75 between vocabulary knowledge and reading comprehension. Several longitudinal studies suggest that vocabulary is causally implicated in the development of reading comprehension. This work shows that early vocabulary knowledge predicts later reading comprehension scores during the early school years (Bast & Reitsma, 1998; de Jong & van der Leij, 2002; Roth, Speece, & Cooper, 2002; Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997). However, the link between vocabulary and reading comprehension is not unidirectional, and there is good evidence that it is reciprocal throughout development. For example, studies that have investigated the other direction of causality have shown that reading comprehension is a better predictor of later vocabulary than vice versa, at least early in the early stages of learning to read (7–8 years) (Eldredge, Quinn, & Butterfield, 1990). Presumably this occurs because the meanings of new words are inferred and the meanings of partially known words refined through the use of context. The notion that new vocabulary is primarily acquired through reading goes back to Huey (1908/1968) and Thorndike (1917), and a body of work shows that written text is an important source of vocabulary acquisition once children become relatively fluent readers (Cunningham, 2005; Echols, West, Stanovich, & Zehr, 1996; Nagy & Scott, 2000). One reason for this influence is that vocabulary in books tends to be more difficult and demanding than that used in everyday conversation (Cunningham & Stanovich, 1998).

Vocabulary knowledge (and assessments of vocabulary) can be differentiated into *breadth* and *depth.* Roughly speaking, breadth refers to the number of words known and can be assessed by the ability to produce word definitions or to select synonyms, and depth refers to what is known about a particular word and refers to more detailed knowledge about the meanings of words. Depth of vocabulary knowledge is an important predictor of comprehension skill even when breadth of vocabulary has been taken into account (Ouellette, 2006). A likely reason for this relation is that the associative links between words, which are (among other things) the product of having a greater depth of vocabulary knowledge, will aid reading comprehension by supporting inference-making. In particular, depth of vocabulary knowledge is likely to be more important than breadth in supporting inference-making, because rich and well-connected semantic representations of words will permit the rapid activation of not only a word's meaning but also those of related concepts. This activation can then provide the basis for many of the inferences that are crucial for the construction of a coherent representation of a text. Indeed, a recent study (Cain & Oakhill, in press) showed that depth, but not breadth, of vocabulary knowledge was an important predictor of global coherence inferences (inferences that link up ideas and themes in the text overall) and that this relation held even when word reading skill and literal memory for the text had been taken into account in statistical analyses.

In relation to problems with comprehension it is tempting to reason that poor comprehenders will be characterized by having inadequate vocabularies, because although they might be able to decode words, they do not understand their meanings. However, this is not necessarily the case. Good and poor comprehenders, who differ on standardized measures of reading comprehension as well as tests of specific comprehension skills such as inference-making, can be matched for knowledge of both written and spoken word meanings (Cain, Oakhill, & Lemmon, 2004). Thus, although vocabulary knowledge and reading comprehension are typically highly correlated, problems with reading comprehension can arise even in the presence of a good level of vocabulary. However, the tests of vocabulary used for matching purposes are typically measures of breadth, whereas recent research suggests that depth of vocabulary knowledge might be more important for reading comprehension (Ouellette, 2006; Tannenbaum, Torgesen, & Wagner, 2006).

Training studies provide direct tests of causality: If performance in reading comprehension improves after training in a component skill such as vocabulary, there is good evidence that the component skill is causally implicated in reading comprehension skill. However, there is little evidence for a direct causal link between vocabulary knowledge and comprehension skill, an indication that access to word meanings may be necessary but not sufficient for comprehension. For instance, training studies can be effective in improving vocabulary knowledge (e.g., Jenkins, Pany, & Schreck, 1978) but generally do not result in increases in comprehension skill. One exception is a study by Beck, Perfetti, and McKeown (1982) that did show transfer of vocabulary training to reading comprehension. They argued that in order for vocabulary instruction to have effects on reading comprehension not only must it be very intensive but also it is necessary to increase both the number of words learned and the fluency with which these new meanings can be accessed.

The idea that the automaticity with which meanings can be accessed is important in reading comprehension is not new. Early models of reading (e.g., Laberge & Samuels, 1974) emphasized the importance of fluency and automaticity of access to word meanings. In correlational studies, good and poor comprehenders have been shown to differ on measures of semantic fluency, such as the ability to rapidly produce a number of instances of a category (Nation & Snowling, 1998). Recently, Oakhill, Cain, McCarthy, and Field (2012) found a strong and specific link between speed of semantic access in vocabulary tasks (synonym and hypernym judgment tasks) and reading comprehension (measured by a standardized task).[1] The link was specific because it was not entirely mediated by word-reading skill or by knowledge about words (assessed by a synonym and hypernym production task), nor was it related to a simple association between comprehension skill and generally faster response times in a control task that required non-semantic identity judgments (i.e., the ability to say whether two words were orthographically identical or not).

These findings suggest that it is not sufficient to know the meanings of words to understand a text, Knowledge of the interrelations between words and rapid access to the semantic representations of words are both also important for comprehension. Comprehension happens in real time, and if appropriate meanings and associations of words are not accessed very rapidly, the reader will often have moved on in the text and the opportunity for semantic information to support inference and integration of the text will have been missed.

SYNTACTIC SKILLS

Children's syntactic skills are usually quite highly developed by the time they start reading, but their syntactic competence continues to develop, albeit in more subtle ways, for several years (Chomsky, 1969; Garton & Pratt, 1998). Children's grammatical skills have typically been assessed with measures of syntactic awareness, a metalinguistic skill, rather than measures of grammatical knowledge. Grammatical knowledge is needed to extract meaning from syntax alone (e.g., the ability to understand the difference between active and passive constructions), and this knowledge may be implicit. In contrast, syntactic awareness is typified by explicit knowledge of syntactic forms, and this awareness can include deliberate and controlled reflection on language. This type of knowledge would be used in judgments about syntactic well-formedness.

It is clear that syntactic knowledge will be needed in order to derive meaning from sentences in text, and it will be fairly crucial since sentences are the building blocks of the meaning of a text. However, research that has investigated the relation between syntactic awareness and reading comprehension has produced rather mixed results. Whereas some studies suggest a direct relation between syntactic knowledge and reading comprehension (Willows & Ryan, 1981), others do not find a relation once vocabulary has been controlled for (Bowey & Patel, 1988). One difficulty in interpreting the results from work in this area is that the various measures of syntactic awareness are differently related to vocabulary and memory (Cain, 2007). Indeed, other studies have shown that working memory may limit the ability to parse syntactically complex sentences in particular (Gottardo, Stanovich, & Siegel, 1996; Smith, Marcaruso, Shankweiler, & Crain, 1989). We discuss the relations between processing capacity and comprehension in more depth in a later section.

There are few longitudinal studies of syntactic skills and comprehension, but those that exist show rather weak relations. In such studies, syntactic awareness is at best a weak predictor of reading comprehension when initial levels of reading comprehension are controlled (Demont & Gombert, 1996; Oakhill & Cain, 2012). Longitudinal studies that have focused on shorter-term follow-ups of young readers find that syntactic awareness at the end of the first year of school predicts reading comprehension one year later (Muter, Hulme, Snowling, & Stevenson, 2004; Tunmer, 1989). However, the findings from those studies are limited because earlier reading comprehension ability was not controlled for (because of the age of the readers). Thus, although there is some evidence that syntactic skills predict the later development of reading comprehension, such skills do not seem to play a major role, particularly when compared with other skills and abilities that are discussed later.

Since understanding of both written and spoken text depends on understanding of the sentences within that text, it could be expected that syntactic knowledge differs between good and poor comprehenders, and some studies show such a difference (Stothard & Hulme, 1992). However, other studies show little evidence of such a relation (Cain, Patson, & Andrews, 2005; Yuill & Oakhill, 1991). These discrepancies are particularly surprising, since the studies all used the same standardized assessment of grammar, the Test for Reception of Grammar (Bishop, 1983). It is possible that the discrepancies might be attributed to differences in the criteria for group selection, or that not all children with comprehension difficulties have the same skill profiles. Studies that use measures of syntactic awareness consistently report difficulties for poor comprehenders (e.g., Gaux & Gombert, 1999; Nation & Snowling, 2000). However, as with vocabulary, there is little evidence for a direct causal relation between syntactic skills and reading comprehension. For instance, Layton, Robinson, and Lawson (1998) were successful in training syntactic awareness in 8- to 10-year-olds, using sentences appropriate to the children's reading level. However, the improvements in syntactic awareness did not result in improvements in reading comprehension in either good or poor comprehenders.

In conclusion, although studies have shown that the development of syntactic skills (in particular, syntactic awareness) is linked to reading comprehension and that good and poor comprehenders differ on some measures of syntax, the findings tend to show that syntactic skills are not as strong a predictor of reading comprehension development and individual differences in reading comprehension as other discourse-level skills that we discuss in the remainder of this chapter.

Discourse-Level Skills

Successful reading comprehension depends not only on understanding the words and sentences. Readers need to connect information from different parts of the text and often need to make inferences in order to form a coherent representation of the text as a whole. To do so, they need to continually update their text representation and monitor the progress of their comprehension. These are both processes that are demanding of working memory resources. Text integration processes will be helped by adequate understanding of cohesive devices (conjunctions, such as *because* and *therefore*, and anaphoric devices, such as pronouns). Understanding of the text as a whole will be supported by knowledge of how texts are typically structured.

WORKING MEMORY AND MEMORY UPDATING

Many of the skills involved in forming a coherent mental model, such as comprehension monitoring and integration and inference-making, are dependent on the storage and coordination of information in memory. For reading and listening comprehension, these draw heavily on two components of Baddeley and Hitch's (1974) classic model of working memory: the phonological loop (a short-term store of verbal information) and the central executive (which manipulates information from short-term memory and long-term stores). In particular, the central executive process of working memory updating is considered essential for successful reading comprehension (Carretti, Cornoldi, De Beni, & Romanò, 2005).

Working memory updating refers to the modification of content to accommodate new input. For text comprehension, this would involve updating of the mental model. For example, if readers inferred that a concept introduced in a text was a butterfly (because it flew up and away from a rose bush), but later changed their interpretation because the text qualified the concept to be a bird (because it had feathers and sang), they would need to update their mental model and replace the incorrect inference (of a butterfly) with the correctly inferred concept (a bird) (Radvansky & Copeland, 2001). A classic task to measure updating is a modification of a word span task, in which participants are presented with a set of items and asked to recall the *x* smallest ones (Belacchi, Caretti, & Cornoldi, 2010; Carretti et al., 2005). For example, in the following list of items, the two physically smallest items are *pen* and *pea*: *pen, dog, shoe, chair, pea*. A participant asked to recall the two smallest items would need to inhibit the item *shoe* on hearing *pea* and update the list of to-be-remembered items (see also Radvansky & Copeland, 2001, for a discussion of updating in text comprehension). Comprehenders constantly have to modify and update their mental model, sometimes excluding previously encoded information that is found to be no longer accurate or relevant.

Although the main components of working memory are in place from an early age, substantial gains in capacity on both short-term storage and working memory tasks (those tapping the central executive) are evident across childhood (Gathercole, Pickering, Ambridge, & Wearing, 2004). In relation to reading comprehension, measures of working memory that tap the central executive processes and those that involve the manipulation and storage of verbal information are more strongly related to reading comprehension in children and adults than memory tasks that require only passive storage of information or the manipulation of visuo spatial information (Carretti, Borella, Cornoldi, & De Beni, 2009; Daneman & Merikle, 1996). Independent measures of working memory are related to discourse-level skills, specifically inference and integration, comprehension monitoring, and knowledge and use of text structure (Cain, Oakhill, & Bryant, 2004).

Children with poor reading comprehension are not typically impaired on measures of short-term storage, as assessed by their ability to store and recall a set of words or digits (Cain, 2006; Carretti et al., 2009; Oakhill, Yuill, & Parkin, 1986; Stothard & Hulme, 1992; but see Nation, Adams, Bowyer-Crane, & Snowling, 1999, for poor performance on specific word types). In contrast, poor comprehenders do less well than same-age peers when the task involves the storage and manipulation of verbal information, tasks that tap the central executive processes of working memory. Poor performance is evident across a range of tasks and materials, including digits, words, and sentences (e.g., Cain, 2006; de Beni & Palladino, 2000; Oakhill, Yuill, & Garnham, 2011; Yuill, Oakhill, & Parkin, 1989). This dissociation in the relation between reading comprehension and tasks that tap short-term storage versus storage and processing is confirmed in a meta-analysis conducted by Carretti et al. (2009). Note that poor comprehenders do not do poorly on tasks that tap visual-spatial skills even if both storage and processing are required. Thus, poor comprehenders' difficulty with working memory tasks may depend on the similarity between the working memory task and reading.

These working memory problems in poor comprehenders map onto their understanding of text under conditions of higher working memory demands. Poor comprehenders have been found to perform more poorly on assessments of inference-making and comprehension monitoring when their working memory is taxed by separating the information to be integrated or compared across several sentences, whereas good comprehenders are considerably less affected by the working memory demands of the comprehension tasks (Cain, Oakhill, & Bryant, 2004; Oakhill, Hartt, & Samols, 2005).

As noted earlier, some researchers have developed tasks designed to specifically tap participants' ability to regulate the contents of working memory. Poor comprehenders are less likely to successfully inhibit information that is no longer relevant and therefore update their memory representation (Cain, 2006; Carretti et al., 2005; de Beni & Palladino, 2000). For example, Carretti et al. found that 8- to 11-year-old poor comprehenders not only recalled fewer words in a working memory task than good comprehenders but also made more errors that involved selecting items that no longer fit the to-be-remembered criterion (i.e., intrusions of words or pictures that were relevant only for a certain amount of time during the execution of the task).

In conclusion, certain working memory tasks can successfully differentiate between good and poor comprehenders: specifically, those tasks that involve verbal stimuli and those that involve complex operations. In addition, these difficulties with working memory may account for poor comprehenders' difficulties on some reading comprehension tasks.

Inference and Integration

Skilled adult readers almost always make the required text-connecting (local and global coherence) inferences quickly and effortlessly, but younger children and poor comprehenders may have difficulties with inference-making for various reasons. Although developmental studies have demonstrated that younger children (6- to 7-year-olds) are able to make inferences, they are less likely than older children and adults to do so spontaneously, and may only do so when prompted or questioned (Casteel & Simpson, 1991; Omanson, Warren, & Trabasso, 1978; Paris & Lindauer, 1976). A number of studies have shown that the ability to make various kinds of inferences increases with age (Ackerman, 1986, 1988; Paris & Lindauer, 1976; Paris, Lindauer, & Cox, 1977). Ackerman (1986) suggested that age-related differences in spontaneous inference-making might be the result of younger children failing to establish referential coherence and being less aware of the need for inferences.

A study by Barnes, Dennis, and Haefele-Kalvaitis (1996) directly addressed the developmental relation between inference skills and background knowledge. The authors trained children aged between 6 and 15 years on a novel knowledge base (i.e., facts about a made-up planet), which they had to learn to criterion (perfect). They were then presented with a multiepisode story and were asked questions, some of which required them to integrate their newly learned knowledge with information in the text to generate inferences. Even though all the children had the relevant knowledge and could access it, this did not attenuate the age-related differences in success at inference-making. Our own work in this area (Oakhill & Cain, 2012) has shown that inference skills contribute to later comprehension skill between 7 and 11 years, over and above the contributions of vocabulary, verbal IQ, and earlier comprehension skill (the *autoregressive effect*). This pattern suggests a possible causal link between inference skill and reading comprehension during development (see also Cain & Oakhill, 1999).

Studies that have investigated individual differences have found that poor comprehenders generate fewer constructive inferences than good comprehenders. An instance is inferences that require the combination of information from two different sentences in a text. Consider: "The boy was chasing the girl. The girl ran into the playground." A plausible inference would be: "The boy ran into the playground" (Oakhill, 1982). Poor comprehenders make fewer such inferences, but this does not seem to be simply because they are poorer at remembering the text. They are able to recall literal details from a text just as well as good comprehenders (Oakhill, 1982), and inference-making difficulties are still apparent even when the text is available to refer to (Oakhill, 1984). Additional support for the contention that poor comprehenders have difficulties with inference-making comes from an investigation of good and poor comprehenders' performance on different types of inference questions. Bowyer-Crane and Snowling (2005) found that poor comprehenders had difficulties in making knowledge-based and elaborative inferences, but performed comparably to normal readers on questions that required attention to literal information or use of cohesive devices.

The relation between general knowledge and the inference problems of poor comprehenders has also been investigated using Barnes et al.'s (1996) paradigm, mentioned earlier, which keeps the knowledge base constant while investigating group differences in inference skill. The findings showed that even when knowledge was controlled for in this very strict manner, less skilled comprehenders generated fewer inferences than did their more skilled counterparts (Cain, Oakhill, Barnes, & Bryant, 2001). Of course, one reason for group differences may be the speed with which children activate and access the relevant information. As noted earlier speed of access to vocabulary is related to reading comprehension, and Barnes et al. (1996) found that speed of access to critical facts influenced the likelihood that they would be used to generate inferences. This is a critical question for future research investigating the sources of poor comprehenders' difficulties with inference-making.

There is also evidence that knowing how to use background knowledge effectively is critical to inference-making. Elbro and Buch-Iversen (2013) found evidence that comprehension problems may be caused by a reader not knowing how to integrate background knowledge with the text. The study focused on global coherence inferences, which require the integration of background knowledge with information from the text to help form a coherent mental model (e.g., "The bank did not give Ole a loan for a new boat. He began to look for a spare-time job." Question: "Why did Ole want a spare-time job?"). Training that focused on the contribution of background knowledge for text comprehension improved 9- to 10-year-olds' ability to make gap-filling inferences. Thus, inference-making difficulties can be explained partially by an inability to use background knowledge appropriately.

Research using designs that can address causality (see earlier discussion) also indicate a causal role for inference skills. Poor comprehenders generate fewer inferences than younger children matched for absolute comprehension level, as measured by a standardized reading test (a comprehension-age match group design), suggesting a causal link between inference skills and reading comprehension outcomes (Cain & Oakhill, 1999). Longitudinal work supports this conclusion, demonstrating that inference skills predict subsequent reading comprehension over time in addition to vocabulary knowledge and verbal IQ (among other variables; Oakhill & Cain, 2012).

COHESIVE DEVICES

One aspect of language skill that may be particularly important in the construction of a mental model is the understanding of cohesive devices. These are linguistic ties such as anaphors and connectives that can aid the integration of successive clauses within and across sentences in a text. Anaphoric pronouns, for example, refer to protagonists or concepts introduced earlier in a text, and therefore require integration between elements in a text to be understood. An example is: "Michael told Jane about his new telescope. *She* thought it was really cool and wanted *one* too." Connectives are another cohesive device that specify how two clauses or sentences are related. For example, *before* and *after* are temporal connectives that signal the chronological order of events, and *because* and *so* specify the direction of causality. Connectives are found in, for example, "Tim was running late, *so/because* he called his girlfriend."

Children with poor reading comprehension have difficulties with anaphor comprehension. Problems with anaphors are likely to be related to poor comprehenders' difficulties with inference and integration. For example, 7-year-old poor comprehenders are less able to take advantage of the gender of a pronoun (e.g., *she* vs. *he*) and use this as a cue to integrate clauses within a sentence (Megherbi & Ehrlich, 2005; see also Oakhill & Yuill, 1986; Yuill & Oakhill, 1988). Studies of adults indicate that the ability to link a pronoun to an appropriate antecedent in the text is related to working memory capacity (Daneman & Carpenter, 1983). The same may be true for children. Difficulties with anaphor comprehension arise when there is intervening text between the anaphor and its antecedent (Ehrlich & Rémond, 1997; Yuill & Oakhill, 1988), and these difficulties are much more pronounced for poor comprehenders.

Children with poor reading comprehension also have problems with both use and comprehension of connectives on a variety of tasks. In a narrative production task, they are less likely than their peers to use specific connectives to indicate causality between events (Cain, 2003). They are also less likely to supply an appropriate connective to link clauses in a cloze task, in which specific connectives have been deleted from sentences in a text, for example "All of the female giraffes stayed under the shady trees the sun was shining" (Cain et al., 2005). As with anaphors, these problems may be intricately related to their broader difficulties with inference and integration because of the role that connectives play in signaling cohesion and coherence (Sanders & Maat, 2006).

COMPREHENSION MONITORING

Comprehension monitoring encompasses several different abilities, but can be characterized as the ability to reflect on what has just been read (i.e., a metacognitive skill). Thus monitoring understanding might include consideration of whether the text made sense, whether it was enjoyable, what was learned from the text, and what the main points were. In reading comprehension research, comprehension monitoring is often measured using an inconsistency-detection task, which measures the reader's ability to identify an inconsistency between two pieces of information in a text. For example, "Moles *cannot see very well*, but their hearing and sense of smell are good. Moles are easily able to find food for their young because *their eyesight is so good*" (inconsistency italicized). Whether the inconsistency has been detected can be measured explicitly by assessing detection errors or implicitly by use of reading times or eye tracking. Comprehension monitoring is likely to be closely related to reading comprehension because readers can only detect inconsistencies when they are actively engaged in the constructive process of reading. Thus, comprehension monitoring skill is likely to overlap with other processes necessary for creating and maintaining a coherent representation of the text (i.e., constructing a coherent mental model, updating that model, and activating information from that model).

In general, younger children are less likely to realize when a text does not make sense and less likely to know what to do about it if they do realize it (for a review, see Baker & Brown, 1984; Markman, 1981). Seminal studies by Markman demonstrated that 6- to 7-year-olds failed to realize that there were serious inadequacies in instructions for how to play a game or perform a magic trick (Markman, 1977) or that there were contradictions within a text of the kind described earlier (Markman, 1979). Performance on these tasks improved with age and was, to a certain extent, enhanced by specific instructions that directed children's attention to the nature of the problem. But even then, detection of missing or erroneous information was not perfect even in sixth graders (see also Baker, 1984). One hypothesis for younger children's failure on these comprehension monitoring tasks is that the demands on their cognitive resources affect their ability to monitor for sense (Baker, 1984; Ruffman, 1996). Indeed, comprehension monitoring errors have been shown to arise simply because children failed to remember the inconsistent premises (Vosniadou, Pearson, & Rogers, 1988). Information processing capabilities increase with age (for a summary, see Oakhill, 1988), and it is likely that children's competence in comprehension monitoring shows a concomitant increase.

In summary, children develop the ability to reflect on their understanding during the primary-school years. Younger children's problems might result at least in part from their lack of knowledge of appropriate standards with which to evaluate their comprehension or their difficulties in building a coherent representation of the text as a whole. However, the precise causal relation between comprehension monitoring and comprehension remains unclear. For instance, Markman (1981) suggests that the ability to think about one's own comprehension is fundamental to comprehension itself, and others similarly have argued that comprehension monitoring, and metalinguistic awareness more generally, are the driving forces behind the development of reading comprehension (see Donaldson, 1978; Vygotsky, 1962). There are others, however, who have suggested that comprehension is fundamental to monitoring and that it is comprehension ability itself that underpins the ability to monitor for meaning (Perfetti, Marron, & Folz, 1996).

To understand the nature of the relation between monitoring and comprehension we need longitudinal research to explore the pattern of relations over time. Few studies have done so. One exception was by Chaney (1998), who showed that early metalinguistic skills at the sentence level (a composite of two tests of structural awareness at the sentence level, which included correction of syntactic errors) predicted reading ability (a combined measure of word reading and comprehension) four years later, over and above the effects of general language ability. In our own longitudinal study (Oakhill & Cain, 2012), we found that comprehension monitoring at age 7 to 8 significantly predicted reading comprehension four years later, even when the autoregressive effect of comprehension had been taken into account, providing evidence for a causal link between earlier comprehension monitoring and later reading comprehension (see de Jong & van der Leij, 2002). However, studies are needed that test the opposite direction of causality, namely the influence of early reading comprehension skills on later comprehension monitoring and metalinguistic skills.

The inconsistency-detection paradigm described earlier has been used extensively to explore the nature and extent of comprehension monitoring differences between good and poor comprehenders. For instance, Ehrlich and colleagues have explored comprehension monitoring by comparing good and poor comprehenders' (12- to 15-year-olds') ability to detect inconsistent anaphors in expository texts. For example, a noun phrase anaphor might have a meaning that is contradictory to its antecedent. Ehrlich (1996) used this manipulation: In the consistent version, a noun phrase was repeated, such as "The protection of existing reserves . . . This protection" whereas in the contradictory version, the second (anaphoric) *protection* was replaced by *wastage*. She found that the good comprehenders were more likely to detect the problematic anaphors than were the poor comprehenders. In a follow-up study, using a reading time paradigm, Ehrlich, Rémond, and Tardieu (1999) showed that good comprehenders spent more time reading sections of text with inconsistent anaphors than did poor comprehenders. The good comprehenders were also more likely to look back to preceding text when they encountered an inconsistent anaphor. Thus, the good comprehenders were not only more likely to spot inconsistencies but also engaged in additional processes to try to make sense of the text.

Oakhill et al. (2005) compared the inconsistency detection abilities of 9- to 10-year-old good and poor comprehenders when the inconsistencies were close in the text (in adjacent sentences) and when they were more distant (separated by several sentences). They found that although poor comprehenders detected fewer inconsistencies in both conditions, there was an interaction between group and condition, such that the difference between groups was significant only in the distant condition. Thus, poor comprehenders are able to do the task but have particular difficulties when it requires the comparison of information across a number of sentences in the text.

UNDERSTANDING STORY STRUCTURE

An important aspect of children's developing understanding of how ideas in stories and other texts are related is their developing knowledge of how texts are structured. These features of text can be supportive of comprehension in that they can evoke relevant background knowledge and schemas, which can provide a framework for understanding and can guide inference generation and constructive processing. Much of the work in this area has focused on narratives for the reasons described earlier. Some of the indicators of text structure are explicit (e.g., titles, subheadings, summaries), and others are implicit (e.g., knowledge that the main character(s) are usually introduced at the beginning of a story). For example, if a story is titled "Pip's first day at school," the reader will have some idea of what the story is going to be about and can start the story with some reasonable assumptions about Pip's experiences. More broadly, even simple stories have a beginning (introduction), a middle (some sort of crisis or major event), and an end (the resolution of what happens in the middle), and the action in the text is largely driven by the main character's goals and motives. Again, this sort of information leads to certain expectations and a preliminary framework for the story.

In many cases, children may develop their knowledge of story structure by being exposed to well-structured stories, and some have argued that narrative discourse acts to bridge the transition between oral language use and reading comprehension (e.g., Westby, 1991). Thus the developmental pattern is thought to progress from conversational discourse to narrative (oral) discourse to literacy. Indeed, narrative oral discourse and written narrative share many features, including more complex syntax and more abstract vocabulary, so it is likely that children's appreciation of and exposure to narrative discourse will have an impact on their reading comprehension development.

One way to assess children's understanding of narrative structure is to get them to tell stories themselves orally and to assess their productions. In general, children's narratives become more coherent as they get older (see Baker & Stein, 1981, for a review of children's developing sensitivity to narrative structure and knowledge of what makes a good story). Children also expect certain types of information to occur in stories and if they are asked to retell short stories that have been orally presented and from which crucial information is missing, they will often add that information in their retellings, so that the retold story conforms to the story as they expected it to be. Similarly, if a story is told with the events out of order, children often restore it to a more normal order when they retell it (e.g., Stein, 1979), and older children are more likely to make these changes to stories.

Some of our own work has shown a strong contribution of story structure understanding to comprehension development (Oakhill & Cain, 2012). In that study, understanding of story structure (as measured by a task in which written sentences had to be reordered to form a sensible story) was a good predictor of later reading comprehension skill between 7 and 11 years, over and above vocabulary skill and general verbal ability.

Van den Broek and colleagues have also explored the development of children's ability to understand the causal structure of texts, and how that ability influences later reading comprehension (for a review, see van den Broek, 1997). Van den Broek characterizes this development in terms of three main trends: sensitivity to the causal structure; an increased focus on internal (to the character) events such as goals, with a concomitant decreased focus on external events such as actions; and the representation of between-episode rather than just within-episode connections. These studies show that even younger children are able to appreciate the causal structure of stories but are also more likely to allocate attention to nonstructural features, including things like how vivid an event was (irrespective of its narrative importance). However, a focus on structural features increased with age. In relation to the second aspect of development (focus on goals, rather than actions), younger children tend to focus on observable concrete actions rather than internal causes such as characters' goals. In relation to the third aspect (making cross-episode connections), younger children are quite good at connecting information within episodes but often fail to connect events across different episodes in the text (Trabasso & Nickels, 1992), so that they tend to miss out on the overall theme of the text and fail to construct an integrated representation of the text overall. Broadly, children's abilities in all three of these areas improve with age (Bourg, Bauer, & van den Broek, 1997).

There have also been studies of the way in which understanding of spoken narratives in prereaders maps onto later reading comprehension. Thus, Paris and Paris (2003) have used picture sequences to assess narrative comprehension in prereaders, and van den Broek, Lorch, and Thurlow (1996) have used video presentations of stories. These alternative assessments would seem to be valid since narratives have similar structures regardless of the way in which they are presented, and there is evidence that the development of children's inference skills is consistent across different media (e.g., van den Broek, 1989).

Using a longitudinal design, van den Broek and colleagues (e.g., Kendeou, van den Broek, White, & Lynch, 2007) have demonstrated not only that comprehension of different media (aurally presented and televised stories) are related within an age group but also that comprehension in 4- and 6-year-olds predicts narrative comprehension two years later. These effects were not simply attributable to general language skills: Performance on the narrative comprehension tasks was related to other aspects of language comprehension, such as oral vocabulary, but not to language skills that support word reading (e.g., phonological awareness, and letter knowledge). Taken together, these results show that narrative language skills develop before children start to learn to read and that there are commonalities in comprehension processes and abilities across different presentation media (written stories, picture sequences, and videos). However, expressive language skills, which would be important since both early comprehension skills and later reading comprehension were measured in part by children's spoken (i.e., expressive) summaries, were not controlled for.

In summary, there is good evidence that children's knowledge about story structures is nascent before reading starts and develops with age and experience of narratives and that this experience supports developing text comprehension. There is some evidence that children's ability to understand story structure across a variety of different media is causally implicated in the development of reading comprehension skill.

Children's recall of stories they hear is related to general reading ability (Smiley, Oakley, Worthen, Campione, & Brown, 1977). In relation to reading comprehension more specifically, Trabasso and Nickels (1992) suggested that children's understanding and production of stories is guided by their knowledge of story structure, and Perfetti (1994) suggested that comprehension failure might at least in part be caused by inadequate knowledge about text structures.

Some studies of our own have also addressed differences in story structure understanding between good and poor comprehenders who were selected using a standardized assessment. For instance, Yuill and Oakhill (1991) showed that poor comprehenders were considerably worse than a comparison group of good comprehenders at selecting the main point of a story from a choice of four options. The difference was apparent whether the stories were read aloud to the children or were presented as a series of pictures. Poor comprehenders also have a poor understanding of the sorts of information provided by particular story features, such as story titles, beginnings, and endings. For example, Cain (1996) interviewed good and poor comprehenders about these features of stories. When asked about story titles, for example "What can the title of a story tell us about that story?," most good comprehenders could provide appropriate examples of the type of information contained in a story title, such as "tells you what it's about and who's in it." However, fewer than 25% of the poor comprehenders were able to provide sensible responses; they were more likely to respond that a title "tells you whether you like the story or not" and some poor comprehenders reported that titles "do not tell the reader anything at all." These group differences were apparent even when children were provided with specific examples (see also Cain & Oakhill, 2006, for converging evidence).

Cain and Oakhill (1996) also used a story production task to compare the structural coherence of oral stories produced by good and poor comprehenders. When prompted by a simple topic idea, such as "the holiday," good comprehenders produced stories that were better organized overall and that were more likely to have a central main point and to comprise a series of causally related events. A further study showed that the poor comprehenders showed some benefit from a topic prompt (title) that provided some goal for the story, for example "How the Pirates Lost Their Treasure," relative to a simple topic prompt, such as "Pirates." Similar findings came from a study by Cragg and Nation (2006), which showed that poor comprehenders produced more poorly structured stories in a written production task than did good comprehenders.

There have been some attempts to train story structure understanding, with some limited positive results. Stevens, van Meter, and Warcholak (2010) reported a study in which the teachers of 5- and 6-year-olds were provided with lessons to help them teach the children about narrative structure while the children were listening to stories during the daily story time. The training continued for a year, during which time the trained children were taught, using a questioning and discussion technique, how to identify important story components such as the main characters, the setting of the story, the main problem, and how it was resolved. The children who received instruction in story structures were able to recall more ideas from new stories and answered more questions about structural elements of those stories (e.g., "who is the main character?") than were the children in a comparison group.

Conclusions and Implications

As children develop their language skills, they progress from conversational discourse to narrative discourse to a particular literate language form. However, the development of reading comprehension skill is not entirely parasitic on language skills. Memory abilities, in particular working memory, updating in memory, and efficient retrieval of information from long-term memory, are important in reading comprehension. In addition, strategy knowledge (such as how to read to obtain particular information and which parts of the text to focus on) is an important predictor of reading comprehension (Willson & Rupley, 1997).

In this review, we have outlined the evidence for a number of skills and processes that are important in reading comprehension. However, the relative importance of different skills is likely to shift during the course of comprehension development (see Scarborough, 1998). Beyond the initial stages of reading, nonphonological language skills become increasingly important in accounting for variance in reading comprehension (for a meta-analysis, see Gough, Hoover, & Peterson, 1996).

Although a number of different skills and abilities correlate—more or less impressively—with reading comprehension, most of the findings have been correlational, and so causal links between ability on particular skills and better reading comprehension cannot be inferred. Where there are studies that imply such causality—and, in particular, training studies—we have mentioned them. Broadly speaking, there is evidence that some of the higher-order comprehension skills, such as inference and integration, comprehension monitoring, and story-structure understanding, have a causal role in developing reading comprehension. However, there is also likely to be a link in the opposite direction. That is because, once children have some level of reading comprehension skill, their reading experience will help them acquire comprehension-related and other skills (see Stanovich, 1986).

The studies of preschool children by van den Broek and colleagues suggest that comprehension skills develop simultaneously with basic language skills and that these comprehension skills have their roots in early narrative comprehension.

Their findings have implications for supporting early reading comprehension. A clear implication is that oral language skills such as vocabulary, syntax, inference-making, and comprehension monitoring should be taught alongside decoding skills in the early school years. Not only are oral language skills linked to the code-related skills that help word reading to develop but also they provide the foundation for the development of the more advanced language skills needed for comprehension.

There is already substantial evidence for effects of early phonemic awareness training on later reading, but there is little work on early awareness of syntactic/narrative skills and later comprehension. Clearly, more work is needed to explore the types of early intervention that will improve young children's appreciation of narrative structure, but van den Broek's work suggests that early interventions could make use of televised or orally presented stories, and indeed Palincsar and Brown (1984) showed that comprehension skills of prereaders could be successfully improved with orally presented text.

Other studies reviewed in this chapter demonstrate that children with specific reading comprehension deficits experience difficulties on a range of language and literacy skills and their difficulties extend to the comprehension of spoken language. However, although a large number of skills are correlates of poor comprehension, only some have been found to be causally implicated in reading and listening comprehension difficulties because the understanding of the skills that are causally implicated in comprehension is still developing. A number of intervention studies reviewed here, have shown that training in many of the processes of comprehension is effective in improving reading comprehension, and studies that have trained some combinations of these processes (e.g., Clarke, Snowling, Truelove, & Hulme, 2010; Carretti, Caldarola, Tencati, & Cornoldi, 2013) have proved effective in improving performance on a standardized test (see also Connor & Al Otaiba, this volume). In the future, more comprehensive models of reading and listening comprehension development are expected to lead to more effective interventions to help children with specific comprehension difficulties.

Note

1 Hypernyms are terms for superordinate categories, for example, *rain/weather*, *dog/mammal*, and *chair/furniture*.

References

Ackerman, B. P. (1986). Referential and causal coherence in the story comprehension of children and adults. *Journal of Experimental Child Psychology, 41,* 336–366.

Ackerman, B. P. (1988). Reason inferences in the story comprehension of children and adults. *Child Development, 59,* 1426–1442.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The Psychology of Learning and Motivation, 8,* 47–90.

Baker, L. (1984). Spontaneous versus instructed use of multiple standards for evaluating comprehension: Effects of age, reading proficiency, and type of standard. *Journal of Experimental Child Psychology, 38,* 289–311.

Baker, L., & Brown, A. I. (1984). Metacognitive skills and reading. In P. D. Pearson (Ed.), *Handbook of reading research* (Vol. 1, pp. 353–394). New York, NY: Plenum Press.

Baker, L., & Stein, N. (1981). The development of prose comprehension skills. In C. M. Santa & B. L. Hayes (Eds.), *Children's prose comprehension: Research and practice* (pp. 7–43). Newark, DE: International Reading Association.

Barnes, M. A., Dennis, M., & Haefele-Kalvaitis, J. (1996). The effects of knowledge availability and knowledge accessibility on coherence and elaborative inferencing in children from six to fifteen years of age. *Journal of Experimental Child Psychology, 61,* 216–241.

Bast, J., & Reitsma, P. (1998). Analyzing the development of individual differences in terms of Matthew effects in reading: Results from a Dutch longitudinal study. *Developmental Psychology, 34,* 1373–1399.

Beck, I. L., Perfetti, C. A., & McKeown, G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74,* 506–521.

Belacchi, C., Carretti, B., & Cornoldi, C. (2010). The role of working memory and updating in Coloured Raven Matrices performance in typically developing children. *European Journal of Cognitive Psychology, 7,* 1010–1020.

Bishop, D. V. M. (1983). *Test for the reception of grammar.* Manchester, England: Chapel Press.

Bourg, T., Bauer, P., & van den Broek, P. (1997). Building the bridges: The development of event comprehension and representation. In P. van den Broek, P. Bauer, & T. Bourg (Eds.), *Developmental spans in event comprehension and representation: Bridging fictional and actual events* (pp. 385–407). Hillsdale, NJ: Erlbaum.

Bowey, J. A., & Patel, R. K. (1988). Metalinguistic ability and early reading achievement. *Applied Psycholinguistics, 9,* 367–383.

Bowyer-Crane, C., & Snowling, M. J. (2005). Assessing children's inference generation: What do tests of reading comprehension measure? *British Journal of Educational Psychology, 75,* 189–201.

Cain, K. (1996). Story knowledge and comprehension skill. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and remediation* (pp. 167–192). Mahwah, NJ: Erlbaum.

Cain, K. (2003). Text comprehension and its relation to coherence and cohesion in children's fictional narratives. *British Journal of Developmental Psychology, 21,* 335–351.

Cain, K. (2006). Individual differences in children's memory and reading comprehension: An investigation of semantic and inhibitory deficits. *Memory, 14,* 553–569.

Cain, K. (2007). Syntactic awareness and reading ability: Is there any evidence for a special relationship? *Applied Psycholinguistics, 28,* 679–694.

Cain, K., & Oakhill, J. V. (1996). The nature of the relationship between comprehension skill and the ability to tell a story. *British Journal of Developmental Psychology, 14,* 187–201.

Cain, K., & Oakhill, J. V. (1999). Inference making and its relation to comprehension failure. *Reading and Writing: An Interdisciplinary Journal, 11,* 489–503.

Cain, K., & Oakhill, J. V. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology, 76,* 683–696.

Cain, K., & Oakhill, J. V. (2007). Reading comprehension difficulties: Correlates, causes, and consequences. In K. Cain & J. Oakhill (Eds.), *Children's comprehension problems in oral and written language: A cognitive perspective* (pp. 41–75). New York, NY: Guilford Press.

Cain, K., & Oakhill, J. V. (in press). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'Année Psychologique/Topics in Cognitive Psychology.*

Cain, K., Oakhill, J. V, Barnes, M., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition, 29,* 850–859.

Cain, K., Oakhill, J. V., & Bryant, P. E. (2000). Investigating the causes of reading comprehension failure: The comprehension-age match design. *Reading and Writing: An Interdisciplinary Journal, 12,* 31–40.

Cain, K., Oakhill, J. V., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96,* 671–681.

Cain, K., Oakhill, J. V., & Bryant, P. E. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96,* 31–42.

Cain, K., Patson, N., & Andrews, L. (2005). Age- and ability-related differences in young readers' use of conjunctions. *Journal of Child Language, 32,* 877–892.

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences, 19,* 246–251.

Carretti, B., Caldarola, N., Tencati, C., & Cornoldi, C. (2013). Improving reading comprehension in reading and listening settings: The effect of two training programmes focusing on metacognition and working memory. *British Journal of Educational Psychology, 84,* 1–17.

Carretti, B., Cornoldi, C., De Beni, R., & Romanò, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. *Journal of Experimental Child Psychology, 91,* 45–66.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY: Cambridge University Press.

Casteel, M. A., & Simpson, G. B. (1991). Textual coherence and the development of inferential generation skills. *Journal of Research in Reading, 14,* 116–129.

Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In H. W. Catts & A. G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 25–40). Hillsdale, NJ: Erlbaum.

Chaney, C. (1998). Preschool language and metalinguistic skills are links to reading success. *Applied Psycholinguistics, 19,* 433–446.

Chomsky, C. (1969). *The acquisition of syntax in children from 5 to 10.* Cambridge, MA: MIT Press

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science, 21,* 1106–1116.

Cragg, L., & Nation, K. (2006). Exploring written narrative in children with poor reading comprehension. *Educational Psychology, 26,* 55–72.

Cunningham, A. E. (2005). Vocabulary growth through independent reading and reading aloud to children. In E. H. Hiebert & M. L. Kamhi (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 45–68). Hillsdale, NJ: Erlbaum.

Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. *American Educator, 22,* 8–15.

Curtis, M. E. (1980). Development of components of reading skills. *Journal of Educational Psychology, 72,* 656–669.

Cutting, L. E., Materek, A., Cole, C. A. S., Levine, T. M., & Mahone, M. E. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Annals of Dyslexia, 59,* 34–54.

Daneman, M., & Carpenter, P. (1983). Individual differences in integrating information between sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 561–584.

Daneman, M., & Merickle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3,* 422–433.

De Beni, R., & Palladino, P. (2000). Intrusion errors in working memory tasks. Are they related to reading comprehension ability? *Learning and Individual Differences, 12,* 131–143.

De Jong, P. F., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading, 6,* 51–77.

Demont, E., & Gombert, J. E. (1996). Phonological awareness as a predictor of recoding skills and syntactic awareness as a predictor of comprehension skills. *British Journal of Educational Psychology, 66,* 315–332.

Donaldson, M. (1978). *Children's minds*. Glasgow, England: Collins.

Echols, L. D., West, R. F., Stanovich, K. E., & Zehr, K. S. (1996). Using children's literacy activities to predict growth in verbal cognitive skills: A longitudinal investigation. *Journal of Educational Psychology, 88,* 296–304.

Ehrlich, M. F. (1996). Metacognitive monitoring in the processing of anaphoric devices in skilled and less skilled comprehenders. In C. Cornoldi & J. V. Oakhill (Eds.), *Reading comprehension difficulties: Processes and remediation* (pp. 221–249). Hillsdale, NJ: Erlbaum.

Ehrlich, M.-F., & Rémond, M. (1997). Skilled and less skilled comprehenders: French children's processing of anaphoric devices in written texts. *British Journal of Developmental Psychology, 15,* 291–309.

Ehrlich, M. F., Rémond, M., & Tardieu, H. (1999). Processing of anaphoric devices in young skilled and less skilled comprehenders: Differences in metacognitive monitoring. *Reading and Writing: An Interdisciplinary Journal, 11,* 29–63.

Elbro, C., & Buch-Iversen, I. (2013). Activation of background knowledge for inference making: Effects on reading comprehension. *Scientific Studies of Reading, 17,* 435–452.

Eldredge, J. L., Quinn, B., & Butterfield, D. D. (1990). Causal relationships between phonics, reading comprehension, and vocabulary achievement in the second grade. *Journal of Educational Research, 83,* 201–214.

Gaux, C., & Gombert, J. E. (1999). Implicit and explicit syntactic knowledge and reading in pre-adolescents. *British Journal of Developmental Psychology, 17,* 169–188.

Garton, A., & Pratt, C. (1998). *Learning to be literate: The development of spoken and written language* (2nd ed.). Oxford, England: Blackwell.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40,* 177–190.

Goff, D., Pratt, C., & Ong, B. (2005). The relations between children's reading comprehension, working memory, language skills and components of reading decoding in a normal sample. *Reading and Writing: An Interdisciplinary Journal, 18,* 583–616.

Gottardo, A., Stanovich, K. E., & Siegel, L. S. (1996). The relationships between phonological sensitivity, syntactic processing and verbal working memory in the reading performance of third-grade children. *Journal of Experimental Child Psychology, 63,* 563–582.

Gough, P. B., Hoover, W. A., & Peterson, C. L. (1996). Some observations on a simple view of reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and interventions* (pp. 1–13). Hillsdale, NJ: Erlbaum.

Gough, P. B., & Tunmer, W. (1986). Decoding, reading and reading disability. *Remedial and Special Education, 7,* 6–10.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2,* 127–160.

Huey, E. (1968). *The psychology and pedagogy of reading.* Cambridge, MA: MIT Press (Original work published 1908).

Jenkins, J. R., Pany, D., & Schreck, J. (1978). Vocabulary and reading comprehension: Instructional effects. *Technical Report No. 100*. Urbana-Champaign, IL: Center for the Study of Reading.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. (2007). Comprehension in preschool and early elementary children: Skill development and strategy interventions. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 27–45). Hillsdale, NJ: Erlbaum.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.

Laberge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293–323.

Layton, A., Robinson, J., & Lawson, M. (1998). The relationship between syntactic awareness and reading performance. *Journal of Research in Reading, 21,* 5–23.

Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development, 48,* 986–992.

Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development, 50,* 643–655.

Markman, E. M. (1981). Comprehension monitoring. In W. P. Dickson (Ed.), *Children's oral communication skills* (pp. 61–84). London, England: Academic Press.

Megherbi, H., & Ehrlich, M. F. (2005). Language impairment in less skilled comprehenders: The on-line processing of anaphoric pronouns in a listening situation. *Reading and Writing: An Interdisciplinary Journal, 18,* 715–753.

Muter, V., Hulme, C., Snowling, M., & Stevenson, J. (2004). Phonemes, rimes, vocabulary and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology, 40,* 665–681.

Nagy, W. E., & Scott, J. (2000). Vocabulary processes. In M. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269–284). Hillsdale, NJ: Erlbaum.

Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology, 73,* 139–158.

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39,* 85–101.

Nation, K., & Snowling, M. J. (2000). Factors influencing syntactic awareness in normal readers and poor comprehenders. *Applied Psycholinguistics, 21,* 229–241.

Oakhill, J. V. (1982). Constructive processes in skilled and less-skilled comprehenders' memory for sentences. *British Journal of Psychology, 73,* 13–20.

Oakhill, J. V. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology, 54,* 31–39.

Oakhill, J. V. (1988). The development of children's reasoning ability: Information-processing approaches. In K. Richardson and S. Sheldon (Eds.), *Cognitive development to adolescence: A reader* (pp. 169–188). Hillsdale, NJ: Erlbaum.

Oakhill, J. V., & Cain, K. (2007). Introduction to comprehension development. In K. Cain & J. V. Oakhill (Eds.), *Children's comprehension problems in oral and written language* (pp. 41–75). New York, NY: Guilford Press.

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading, 16,* 91–121.

Oakhill, J. V., Cain, K., McCarthy, D., & Field, Z. (2012). Making the link between vocabulary knowledge and comprehension skill. In A. Britt, S. Goldman, & J.-F. Rouet (Eds.), *From words to reading for understanding* (pp. 101–114). New York, NY: Routledge.

Oakhill, J. V., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing: An Interdisciplinary Journal, 18,* 657–686.

Oakhill, J. V., & Yuill, N. M. (1986). Pronoun resolution in skilled and less skilled comprehenders: Effects of memory load and inferential complexity. *Language and Speech, 29,* 25–37.

Oakhill, J. V., Yuill, N. M., & Garnham, A. (2011). The differential relations between verbal, numerical and spatial working memory abilities and children's reading comprehension.

International Electronic Journal of Elementary Education, 4, 83–106.
Oakhill, J. V., Yuill, N. M., & Parkin, A. (1986). On the nature of the difference between skilled and less-skilled comprehenders. *Journal of Research in Reading, 9,* 80–91.
Omanson, R. C., Warren, W. M., & Trabasso, T. (1978). Goals, inferences, comprehension and recall of stories by children. *Discourse Processes, 1,* 337–354.
Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98,* 554–566.
Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1,* 117–175.
Paris, S. G., & Lindauer, B. K. (1976). The role of inference in children's comprehension and memory for sentences. *Cognitive Psychology, 8,* 217–227.
Paris, S. G., Lindauer, B. K., & Cox, G. L. (1977). The development of inferential comprehension. *Child Development, 48,* 1728–1733.
Paris, A. H., & Paris, S. G. (2003). Assessing narrative comprehension in young children. *Reading Research Quarterly, 38,* 36–76.
Perfetti, C. A. (1994). Psycholinguistics and reading ability. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 849–894). San Diego, CA: Academic Press.
Perfetti, C. A., Marron, M. A., & Foltz, P. W. (1996). Sources of comprehension failure: Theoretical perspectives and case studies. In C. Cornoldi & J. V. Oakhill (Eds.), *Reading comprehension difficulties: Processes and remediation* (pp. 137–165). Hillsdale, NJ: Erlbaum.
Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *Journal of Educational Research, 95,* 259–272.
Radvansky, G. A., & Copeland, D. E. (2001). Working memory and situation model updating. *Memory & Cognition, 29,* 1073–1080.
Ruffman, T. (1996). Reassessing children's comprehension-monitoring skills. In C. Cornoldi and J. V. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 33–67). Hillsdale, NJ: Erlbaum.
Sanders, T., & Maat, P. (2006). Cohesion and coherence: Linguistic approaches. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (2nd ed., pp. 591–595). London, England: Elsevier.
Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In P. Accardo, A. Capute, & B. Shapiro (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75–119). Timonium, MD: York Press.
Scott, C. M. (2004). Syntactic contributions to literacy learning. In C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of learning and literacy: Development and disorders* (pp. 340–362). New York, NY: Guilford Press.
Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J., & Brown, A. L. (1977). Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. *Journal of Educational Psychology, 69,* 381–387.
Smith, S. T., Macaruso, P., Shankweiler, D., & Crain, S. (1989). Syntactic comprehension in young poor readers. *Applied Psycholinguistics, 10,* 420–454.
Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360–406.
Stein, N. L. (1979). How children understand stories: A developmental analysis. In L. G. Katz (Ed.), *Current topics in early childhood education* (Vol. 2, pp. 261–291). Norwood, NJ: Ablex.
Stevens, R. J., van Meter, P., & Warcholak, N. D. (2010). The effects of explicitly teaching story structure to primary grade children. *Journal of Literacy Research, 42,* 159–198.
Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children: The role of language comprehension and working memory skills. *Reading and Writing: An Interdisciplinary Journal, 4,* 245–256.
Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading, 10,* 381–398.
Thorndike, E. L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology, 8,* 323–332.
Thorndike, R. L. (1973). *Reading comprehension education in fifteen countries*. New York, NY: Wiley.
Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Burgess, S., & Hecht, S. (1997). Contributions of phonological awareness and raid automatic naming ability to the growth of word-reading skills in second- to fifth-grade children. *Scientific Studies of Reading, 1,* 161–185.
Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes, 15,* 249–275.
Tunmer, W. E. (1989). The role of language-related factors in reading disability. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the puzzle* (pp. 91–131). Ann Arbor: University of Michigan Press.
Van den Broek, P. W. (1989). The effects of causal structure on the comprehension of narratives: Implications for education. *Reading Psychology: An International Quarterly, 10,* 19–44.
Van den Broek, P. (1997). Discovering the cements of the universe: The development of event comprehension from childhood to adulthood. In P. van den Broek, P. Bauer, & T. Bourg (Eds.), *Developmental spans in event comprehension: Bridging fictional and actual events* (pp. 321–342). Hillsdale, NJ: Erlbaum.
Van den Broek, P., Lorch, E. P., & Thurlow, R. (1996). Children's and adult's memory for television stories: The role of causal factors, story-grammar categories, and hierarchical level. *Child Development, 67,* 3010–3028.
Vosniadou, S., Pearson, P. D., & Rogers, T. (1988). What causes children's failures to detect inconsistencies in text? Representation versus comparison difficulties. *Journal of Educational Psychology, 80,* 27–39.
Vygotsky, L. S. (1962). *Thought and language.* Cambridge, MA: MIT Press.
Westby, C. E. (1991). Learning to talk—talking to learn: Oral-literate language differences. In C. Simon (Ed.), *Communication skills and classroom success: Therapy methodologies for language learning disabled students* (pp. 181–218). San Diego, CA: College-Hill Press.
Willows, D. M., & Ryan, E. B. (1981). Differential utilisation of syntactic and semantic information by skilled and less skilled readers in the intermediate grades. *Journal of Educational Psychology, 73,* 607–715.

Willson, V. L., & Rupley, W. H. (1997). A structural equation model for reading comprehension based on background, phonemic and strategy knowledge. *Scientific Studies of Reading, 1,* 45–63.
Yuill, N., & Oakhill, J. (1991). *Children's problems in text comprehension: An experimental investigation*. New York, NY: Cambridge University Press.
Yuill, N. M., & Oakhill, J. V. (1988). Understanding of anaphoric relations in skilled and less skilled comprehenders. *British Journal of Psychology, 79,* 173–186.
Yuill, N. M., Oakhill, J. V., & Parkin, A. J. (1989). Working memory, comprehension skill and the resolution of text anomaly. *British Journal of Psychology, 80,* 351–361.

CHAPTER

24 Development of Dyslexia

Bruce F. Pennington *and* Robin L. Peterson

Abstract

This chapter explains the reciprocal relation between the biology and psychology of reading by reviewing what is known about how dyslexia develops, beginning with etiology (genes and environment and their interplay) and moving across levels of analysis to reading itself. Current research supports the view that early changes in brain development lead to reductions in white matter connectivity in the left hemisphere, which in turn affect the development of cognitive processes necessary for reading development.

Key Words: dyslexia, genes, white matter, etiology, development, brain development

How can we achieve a complete understanding of how typical and atypical reading develops? The key premise of this chapter is that such an understanding will require multiple levels of analysis, beginning with early genetic and environmental influences on brain development, next considering how those changes in brain development affect the neural networks involved in cognitive processes important for learning to read, then saying how those cognitive processes influence reading development itself, and finally considering how later environmental influences, including the language and writing system being learned by the developing reader, affect this developmental process. Because research on behaviorally defined disorders like dyslexia must begin with the behavioral phenotype and then move to deeper levels of analysis, we will cover these levels of analysis in roughly reverse order.

Before doing that, it is important to clarify some basic issues and terminology. Dyslexia is an interesting example of the intersection between an evolved behavior (language) and a cultural invention (literacy). While it is less likely that there are genes for reading or other relatively recent cultural inventions (consider agriculture, banking, and football), there are genetic influences on evolved cognitive and behavioral traits necessary for proficiency in such cultural inventions. Hence, even though reading is a cultural invention, there still is a biology of reading development.

Since the term "etiology" is sometimes used in different ways, it is important to be precise at the outset about what we mean by this term. Etiology as used here refers to initial or *distal* causes of individual differences within a species, the early factors that change the trajectory of development in some domain of function so as to produce different outcomes among individuals in a population. So various health outcomes, both favorable and unfavorable (e.g., longevity and physical fitness, but also heart disease, cancer, obesity, and cystic fibrosis) all have etiologies, as do various psychological traits (e.g., intelligence, personality, and the various aspects of reading skill discussed in this book) and psychological disorders (e.g., intellectual disability, anxiety, and dyslexia). Some disorders, like cystic fibrosis, are categorical (you either have the disorder or you do not); these categorical disorders often have a discrete etiology, like a mutation in a single gene, as is true for cystic fibrosis, phenylketonuria

(PKU), and Huntingon's dementia (HD). Many other disorders, and especially behaviorally defined disorders, are not categorical, but just extremes on a continuous distribution that ranges from optimal outcomes to poor outcomes, with the underlying mechanisms being similar across the whole distribution. For instance, many cases of intellectual disability (formerly called "mental retardation") are mainly defined by a cutoff on the distribution of intelligence, just as is reading disability or dyslexia (even though there are forms of intellectual disability that have a known genetic etiology, like untreated PKU, Down syndrome, or Fragile X syndrome.) For these noncategorical disorders the etiology is often complex, due to many etiological factors acting together.

Etiology consists of genetic and environmental risk and protective factors (and their interplay) that act in development to produce outcome differences among members of a population. If the outcome in question involves behavior, then these etiological factors generally act on brain development in some fashion or another, because our brains produce our behavior. The resulting changes in the anatomy, physiology, and cognitive processes of the developing brain constitute the *proximal* causes of behavior. So the proximal causes of behavior found in brain mechanisms are not what we mean by the term "etiology."

Nonetheless, the identification of etiological risk factors, especially genetic ones, can be very informative about the development of individual differences because different genes act at different times on different processes in brain development. Identifying even one rare gene that influences an outcome can greatly accelerate progress in finding other genes, because there are families of genes that work together in development. As we will see, many of the candidate genes for dyslexia appear to be part of such a gene family.

Hence, etiology is about individual differences within a population or species. There are also universal species-typical behaviors (like language and social behavior in humans) that are caused by evolution, both biological and cultural, but these causes of human universals are also not what we mean by the term "etiology." Nonetheless, evolved human genes and cultural practices can be very informative about where to look for etiological factors that lead to individual differences in behaviors like language and social behavior. Hence, the etiology of behaviorally defined disorders is potentially informative about both individual differences in development and about human evolution, just as the evolution of human genes and culture can be informative about the etiology of individual differences.

It is these reciprocal relations across levels of analysis that make the study of etiology so exciting and so important for both basic and applied science. For instance, a discovery about the etiology of a rare pathology can lead to the discovery of not only other related genes but also pathogenetic and evolutionary mechanisms. As a specific example, a mutation in the FOXP2 gene was found in a family (the KE family) with a rare oral-motor coordination disorder (i.e., dyspraxia) that affected their speech and language development (Fisher, Vargha-Khadem, Watkins, Monaco, & Pembrey, 1998; Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001). Subsequent imaging studies found that (1) this gene appeared to act on the basal ganglia in the brain, an important structure in motor control (Lai, Gerrelli, Monaco, Fisher, & Copp, 2003); (2) this gene evolved recently in human evolution (Enard, 2011); and (3) an earlier form of this gene is important in audio vocal communication in birds (Scharff & Haesler, 2005). As can be seen, etiological research on this rare disorder led to breakthrough discoveries with much wider significance for our understanding of the evolution of human language.

Hence, the long-term goal of research on etiology of both typical and atypical behavior is to trace causal pathways that run from evolution to etiology to brain mechanisms to the development of behavior. Achieving this goal for a given behavior will have important implications for the neuroscience of all behavior. In this chapter, we review how close we are to achieving this long-term goal for dyslexia.

Definition of Dyslexia

As discussed in Peterson and Pennington (2012), individuals with *developmental dyslexia* have difficulties with accurate or fluent word recognition and spelling despite adequate instruction and intelligence and intact sensory abilities (Lyon, Shaywitz, & Shaywitz, 2003). The ultimate goal of reading is comprehension, which is a function of both decoding ability and oral language comprehension (Hoover & Gough, 1990). Dyslexia is defined by difficulties with decoding, whereas by comparison, listening comprehension is typically more intact. Thus while individuals with very limited decoding abilities (i.e., young children or individuals with severe dyslexia) have poor reading comprehension, individuals with milder decoding problems can still support adequate reading comprehension with intact oral language skills (Bruck, 1990, 1992, 1993). So-called poor comprehenders show the opposite profile of adequate decoding but poor understanding of what is read. Not surprisingly, poor comprehenders tend to have deficits in oral language comprehension, and this profile is sometimes considered a type of language disorder (Nation, Cocksey, Taylor, & Bishop, 2010).

Although some previous diagnostic systems have grouped dyslexia and poor reading comprehension together (e.g., *Diagnostic and Statistical Manual of Mental Disorders* [DSM], 2000), this chapter is only about dyslexia. Many researchers use the terms "dyslexia" and "reading disability" interchangeably, although as the preceding distinction makes clear, other learning disorders (i.e., language disorder) can affect reading. Research suggests that dyslexia represents the low end of a normal distribution of word reading ability (Rodgers, 1983; Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992). Thus in order to diagnose the disorder, a somewhat arbitrary cutoff must be set on a continuous variable.

Should the diagnostic threshold for dyslexia be relative to age or intelligence quotient (IQ)? The logic behind the IQ-discrepancy definitions is that the cause of poor reading might differ between low-IQ and high-IQ individuals. Specifically, it has been assumed that IQ sets a limit on achievement across domains, and so children with low IQ are likely to be poor readers because of general learning difficulties rather than a specific decoding problem. Genetic differences contribute more to high-IQ dyslexia than to low-IQ dyslexia (Wadsworth, Olson, & DeFries, 2010). A related finding is that dyslexia is more genetically based in children from higher socioeconomic status (SES) families than in children from lower SES families (Friend et al., 2009). Together, these results suggest that advantaged children with strong cognitive abilities are likely to be good readers unless they have specific genetic risk factors for poor decoding. On the other hand, there are myriad reasons why other children will struggle with reading. These include environmental influences associated with low SES, and those will account for more of the variance in poor reading in children from lower SES families than children from higher SES families. While the same risk genes are probably important across the range of SES, they contribute less to poor reading in the presence of environmental risk factors associated with lower IQ. We do not yet know which proximal environmental factors are most likely to contribute to low reading ability, though some reasonable possibilities are discussed later in the "Environmental Influences" section.

Despite this evidence for a different weighting of genetic and environmental risk factors in the etiologies underlying dyslexia in children with high versus low IQ, published work does not support the external validity of the distinction between age-referenced and IQ-referenced definitions in terms of underlying neuropsychology or appropriate treatments. Specifically, poor readers of all general ability levels have disproportionately poor skills in phonological processing (processing sounds in language), as discussed further in the "Neuropsychology of Dyslexia" section. As a group, children with dyslexia respond best to treatment emphasizing phonics-based reading instruction. Although there are individual differences in how well individuals with dyslexia respond to such intervention, these differences do not appear to be solely or even primarily a function of IQ (Jimenez, Siegel, O'Shanahan, & Ford, 2009; Silva, McGee, & Williams, 1985; Stuebing, Barth, Molfese, Weiss, & Fletcher, 2009).

The two definitions overlap, but some people with clinically significant reading problems meet only IQ-discrepancy criteria (high ability, weaker-than-expected word reading), whereas others meet only age-discrepancy criteria (low ability, poor word reading). The previous version of the DSM (Diagnostic and Statistical Manual of Mental Disorders, 2000) required that reading achievement be below the level expected for both age and IQ. The most recent revision of the DSM now requires that reading be below age expectations in every case. Although the updated definition should facilitate the identification and remediation of reading problems in children with broader cognitive difficulties, it unfortunately continues to exclude those of high ability who nonetheless have clinically impairing difficulties and could benefit from reading intervention. Indeed, as the preceding discussion makes clear, the new definition ironically means that fewer children with a stronger genetic etiology will probably be classified as dyslexic. Thus for both research and clinical purposes, we think it is more appropriate to identify children who meet either age- or IQ-discrepancy criteria as having dyslexia.

Neuropsychology of Dyslexia

Scientific progress concerning the etiology of dyslexia has been built on a fairly mature understanding of its neuropsychology. By neuropsychology, we refer to the study of specific brain-based processes (such as particular attentional, memory, or language-based skills) that are not directly observable and that are hypothesized to underlie a disorder's defining symptoms. It turns out that the neuropsychological deficits associated with a developmental disorder are often more stable and heritable than the defining symptom itself and are frequently present in family members who do not meet full diagnostic criteria for the disorder. In the case of dyslexia, relatives of affected family members can have reading skills in the normal range despite deficits on some specific phonological processing tasks. In other words, neuropsychological constructs can serve as *endophenotypes* for behaviorally defined disorders. Most of what we know about the genetics of dyslexia has depended on decades of research on its neuropsychology, which has allowed for the use of optimal endophenotypes in etiologic studies. The relationship is reciprocal, because as scientists discover links from etiology to pathogenesis, that knowledge will further constrain the neuropsychological level of analysis and will particularly help inform which brain and cognitive changes may be causal in a disorder (as opposed to associated with the disorder for other reasons). Because of the importance of neuropsychology to the study of etiology, we now briefly review what is known about the neuropsychology of dyslexia.

Much research has made clear that dyslexia is a language-based disorder whose primary underlying deficit involves problems in phonological processing (i.e., processing of sounds in oral language) which leads to later problems processing written language. In the phonological theory of dyslexia, the ability to attend to and manipulate linguistic sounds is crucial for the establishment and automatization of letter–sound correspondences, which in turn underlie accurate and fluent word recognition through the process of phonological coding. As discussed further in what follows, phonological processes are important not only for learning to read alphabetic orthographies (which represent phonemes, or individual speech sounds) but also for learning to read logographic orthographies (in which the script represents language at the morpheme/syllable level), although the phonological grain size most important for skilled reading varies across scripts (Perfetti, Zhang, & Berent, 1992). An important caveat is that the relation between phonological skills (particularly phonological awareness) and reading is bidirectional; over time, poor reading can cause poor phonological awareness (Castles, Wilson, & Coltheart, 2011; Morais, Cary, Alegria, & Bertelson, 1979). The general consensus is that the phonological deficits of dyslexia result from faulty development of phonological representations, which are characterized as poorly segmented, imprecise, or otherwise degraded (Elbro, Borstrøm, & Petersen, 1998; Manis, McBride-Chang, Seidenberg, & Keating, 1997). Evidence for this view comes from studies that demonstrate that children with dyslexia perform poorly on implicit phonological processing tasks, which do not require explicit awareness or manipulation of speech sounds. For example, compared with typically developing controls, children with dyslexia need to hear more of a word in order to recognize it or to show priming effects (Boada & Pennington, 2006).

Any neuropsychological theory of dyslexia must account for the fact that young children who go on to develop dyslexia have subtle difficulties with spoken language long before they encounter a written script. Babies who will become dyslexic show a different brain response to speech stimuli than babies who will not (Guttorm et al., 2005). As toddlers, these children lag behind their peers in vocabulary and syntax (grammar) development, and in preschool they have difficulties with phonological awareness (Scarborough, 1990; Torppa, Lyytinen, Erskine, Eklund, & Lyytinen, 2010). Research on typical language development suggests that, for all children, phonological representations start out as fairly holistic and become gradually more detailed or segmented over time. Babies probably represent most words as single entities. With language development, phonological representations begin to emphasize syllables, then subsyllabic distinctions, and ultimately individual phonemes (Fowler, 1991). Studies with adult natural illiterates (who are cognitively normal but have no formal schooling) demonstrate that phoneme-level representations do not arise automatically in language development and are likely to be a result of exposure to an alphabetic writing system (Castro-Caldas, Petersson, Reis, Stone-Elander, & Ingvar, 1998; Morais et al., 1979). Thus, difficulties in phonological development in dyslexia are probably not restricted to phonemic or segmental representations and must lie in other dimensions of the speech stream.

For many years, a single-deficit phonological theory of dyslexia was most prominent. However, mounting evidence shows that, although phonological deficits are standard in individuals with dyslexia, a single phonological deficit is probably not sufficient to cause the disorder. Other deficits could relate to phonological problems in several ways: The additional deficit could be independent of the phonological issue, with several deficits needed to cause the full clinical phenotype (Pennington, 2006); there could be phonological and nonphonological subtypes of dyslexia (Bosse, Tainturier, & Valdois, 2007; Hadzibeganovic et al., 2010); the phonological deficit could arise from a sensory or general learning problem (Buchholz & Davies, 2007; Nicolson & Fawcett, 2007); or the phonological deficit might cause the reading trouble, whereas other deficits are associated for other reasons (Ramus, 2004).

Consistent with a multiple deficit hypothesis, results of family risk designs (which follow children who are at genetic risk for dyslexia based on their family history, but who are too young to have been diagnosed with the disorder themselves) and longitudinal studies of children with early speech/language disorders have consistently found that many children develop normal-range literacy skills despite preschool phonological deficits similar in magnitude to those of children who ultimately develop dyslexia (Bishop, McDonald, Bird, & Hayiou-Thomas, 2009; Peterson, Pennington, Shriberg, & Boada, 2009; Snowling, Gallagher, & Frith, 2003). These children appear to be protected from dyslexia because of relative strengths in other cognitive skills associated with reading. Conversely, children with multiple cognitive deficits are at much higher risk for dyslexia. Across countries and languages, many cognitive-linguistic constructs consistently predict later dyslexia. Those most consistently implicated include phonological awareness, rapid serial naming (speeded naming of a matrix of familiar objects, colors, letters, or numbers), verbal short-term memory, vocabulary and other aspects of broader oral language skill, and graphomotor processing speed (McGrath et al., 2011; Pennington et al., 2012; Scarborough, 1998; Wolf & Bowers, 1999). The most powerful individual predictor varies with developmental stage. In toddlers and young children, broader language development is most strongly linked to later reading; by 4 or 5 years of age, phonological awareness is the dominant predictor; and tasks emphasizing speed (i.e., rapid serial naming and processing speed) become increasingly important as literacy development progresses, probably because they are more linked to reading fluency than to single-word reading accuracy (Pennington & Lefly, 2001; Puolakanaho et al., 2007, 2008; Scarborough, 1990; Snowling et al., 2003; Torppa et al., 2010). Longitudinal research suggests that these deficits make a causal contribution to reading problems and are not fully accounted for by comorbidities (other disorders that frequently co-occur with dyslexia) or the cumulative effects of reading difficulties.

Research has made clear for many years that dyslexia does not result from disturbances in basic visual perception (Vellutino, 1979; Ramus et al., 2003). However, there has recently been renewed interest in the possible role of visual attentional deficits in reading difficulties (Facoetti, Corradi, Ruffino, Gori, & Zorzi, 2010). Visual attention is measured through serial search, orienting/cueing paradigms, or crowding paradigms that require participants to recognize pictures amid varying degrees of visual clutter; some of these skills probably contribute to performance on nonlinguistic processing speed tasks known to be correlated with reading. A recent study demonstrated that performance on visual attention tasks in preschool significantly predicted reading ability two years later, after accounting for the influence of reading-related phonological processing skills (Franceschini, Gori, Ruffino, Pedroll, & Facoetti, 2012). Initial evidence suggests a similar pattern of results across writing systems with varying degrees of consistency in letter–sound relationships (i.e., Italian and French) (Zorzi et al., 2012). While deficits in visual attention do not easily account for the early speech-language phenotype in predyslexic children, they might represent an additional cognitive deficit that interacts with language problems to cause reading failure. Further research is needed on this question.

Cross-Linguistic Findings

Although research on dyslexia initially focused primarily on reading difficulties in English, there has recently been a good deal of attention focused on the nature of dyslexia across languages. Here, we briefly summarize what is known about how dyslexia manifests across languages showing two different types of variability: first, among alphabetic orthographies that vary in the degree of consistency of letter–sound correspondences; and second, in alphabetic versus logographic orthographies.

Children at the low end of reading ability distribution in languages with more consistent mappings between letters and sounds (e.g., Italian or Finnish) have less severe reading problems than those learning to read less consistent languages (i.e., English), at least in terms of accuracy (Landerl, Wimmer, & Frith, 1997). Difficulties with reading fluency, or speed of reading connected text, seem similar across languages (Caravolas & Samara, this volume; Caravolas, Volin, & Hulme, 2005).

Several studies have noted important universal features in normal and disordered reading across cultures, despite linguistic differences. Cognitive predictors of early reading were similar for five European orthographies (Finnish, Hungarian, Dutch, Portuguese, and French), in agreement with previous results in English. Particularly, phonological awareness was the main predictor of reading in each language, although it had more of an effect in consistent than in less consistent orthographies. Other predictors, such as rapid serial naming, vocabulary knowledge, and verbal short-term memory, made smaller contributions than did phonological awareness, except in Finnish (the most consistent writing system), in which vocabulary had at least as large an effect on reading (Ziegler et al., 2010).

Cross-cultural similarities appear to extend in large part to logographic languages as well, such as Chinese. By contrast with alphabetic writing systems, in which letters represent phonemes, the smallest written units in Chinese are characters representing monosyllabic morphemes (units of language that convey meaning). However, phonology is not irrelevant to reading in Chinese. Chinese characters have phonological elements (Kessler & Treiman, this volume), and skilled readers of the language show phonological effects on word recognition (Pollatsek, this volume). Phonological awareness is a key correlate and predictor of reading skill in Chinese just as in alphabetic orthographies. However, in contrast to alphabetic languages in which awareness of phonemes is critically important, morphological and syllabic awareness play a larger role in learning to read Chinese (see Caravolas & Samara, this volume). This finding is not surprising given the differences in how the orthographies represent language.

Brain Bases of Dyslexia

Because reading is a linguistic skill, we would expect it to involve activation of brain structures used in oral language processing and some additional structures associated with visual-object processing and establishment of visual–linguistic mappings. Indeed, functional imaging studies have consistently revealed that individuals with dyslexia show abnormal activations of a distributed left hemisphere language network (Demonet, Taylor, & Chaix, 2004; Richlan, Kronbichler, & Wimmer, 2009). Underactivation has been reported in two posterior left hemisphere regions: a temporoparietal region believed to be crucial for phonological processing and phoneme–grapheme conversion and an occipitotemporal region, including the so-called visual word form area, which is thought to participate in whole-word recognition. Abnormal activation of the left inferior frontal gyrus is also commonly reported. Structural imaging studies have revealed gray matter decreases in this same network. A recent family risk study demonstrated that these gray matter decreases predate literacy instruction and are thus not only a consequence of reading failure (Raschle, Chang, & Gaab, 2011).

That individuals with dyslexia show functional abnormalities in both posterior and anterior language networks has led to the hypothesis that dyslexia is a disconnection syndrome (a neurological syndrome produced by white matter damage). Accordingly, much research has explored white matter correlates of dyslexia by use of diffusion tensor imaging. The most consistent findings have included local white matter changes (as indexed by a technique called fractional anisotropy) in children and adults with dyslexia in left temporoparietal regions and in the left interior frontal gyrus (Deutsch et al., 2005; Dougherty et al., 2007; Klingberg et al., 2000; Rimrodt, Peterson, Denckla, Kaufmann, & Cutting, 2010). Studies have consistently reported correlations between white matter integrity (a measure of how thoroughly axons are coated with myelin) and phonological skills.

The neural correlates of dyslexia appear remarkably consistent across alphabetic languages with varying degrees of consistency (Paulesu et al., 2001; Silani et al., 2005) and even across alphabetic and logographic orthographies (Hu et al., 2010) despite the fact that the neural basis of skilled reading in Chinese and English is at least partly different (Hu et al., 2010). However, learners of consistent alphabetic orthographies are less likely to display clinically significant reading problems compared with learners of inconsistent orthographies (probably because those with reading vulnerabilities can still read accurately, even if slowly, in consistent languages). In sum, cross-cultural work suggests universality in the neurobiological and neurocognitive causes of dyslexia, but there is cross-cultural specificity in the manifestation of these underpinnings, with the same biological liability more likely to cause substantial impairment in some languages than in others.

Possible Etiological Mechanisms That Apply to Dyslexia

Before discussing the range of possible etiological mechanisms, it is important to be clear that neither genes nor environments code for behavior directly. As discussed by Oyama (1985), both sides of the nature-nurture debate share the same erroneous assumption that the instructions for behavior are preexistent either in the genome or in the environment and are imposed on the developing organism. Instead, genetic and environmental influences are inputs to a developmental process and their impact on behavioral outcome depends on their interactions with all the components of that process. Consequently, it is misleading to speak of the genome as a blueprint or to think that genes code for behaviors. A better metaphor for the genome is that it is a recipe—that is, a sequence of operations—that produces a new form. But even this metaphor is misleading, because there is no chef to follow the recipe. Genes simply code for protein structure or regulate other genes, and variations in the structure of a given protein in a particular developmental context may push behavioral outcomes in one direction or another. Thus genetic and environmental factors are best conceptualized as acting as risk (or protective) factors in the development of individual differences in behavior; their effects are probabilistic rather than deterministic. We next consider possible etiological mechanisms, beginning with behavioral genetics and then molecular genetics.

The ACE Model

Behavioral geneticists have documented moderate heritability (often around .50) for individual differences in most dimensions of human cognition and personality (Plomin, Haworth, Meaburn, Price, & Davis, 2013), including both typical reading and dyslexia. It is important to understand what the technical term "heritability" means and does not mean. Heritability refers to the proportion of variance in a given population that is attributable to genetic influences; other variance components are attributable to environmental influences, gene–environment interplay, or just error of measurement. Heritability estimates do not tell you about the cause of an individual's outcome and, because they are population-specific, they can vary across populations. The ACE model for estimating these variance components is described shortly. Like all behaviorally defined disorders, the cause of dyslexia is multifactorial and is associated with multiple genes and environmental risk factors. Both dyslexia and normal variations in reading skill are familial and moderately heritable (Pennington & Olson, 2005). Subsequent to the Pennington and Olson (2005) review, three large twin studies in the United States and the United Kingdom have confirmed these results (Christopher et al., 2013; Harlaar et al., 2005; Logan et al., 2013). Since these subsequent twin studies are longitudinal, they have been able to examine how the heritability of reading skill changes with age. For instance, Logan et al. (2013) demonstrated that the heritability of individual differences in reading skill steadily increases from .22 at 6 years to .82 at 12 years. These increases in heritability probably reflect both a narrowing of environmental influences on reading produced by a fairly standard reading curriculum once children enter formal education, and an increasing correlation between genotype and environment (i.e., G-E correlation) as children increasingly are able to pick niches that fit their level of reading skills (e.g., good readers read more on their own and become even better readers, while poor readers avoid reading.) Both of these explanations are examples of gene–environment interplay, which is discussed later.

A similar range of heritability estimates has been found for other behaviorally defined neurodevelopmental disorders, such as attention deficit hyperactivity disorder (ADHD), speech sound disorder (SSD), and language impairment (LI), all of which are comorbid with dyslexia. Because these results come from mainly middle-class twin samples in developed countries, it is important to remember that they may not generalize to other populations (but see Hensler, Schatschneider, Taylor, & Wagner, 2010, who found moderate heritability, >.50, both for dyslexia and typical reading skill in a more ethnically and economically diverse sample).

Heritability estimates are usually derived from applying a very simple variance components model to data from twin or adoption studies. This ACE model estimates main effects of genes acting additively (A), common or shared environment (C), and nonshared environment (E). Shared environmental influences are ones that are shared by siblings in the same family (e.g., the number of books in the home) but differ across families; environmental influences that are shared by all families, like light and gravity, are crucial to development but do not contribute to individual differences. Nonshared environmental influences are ones that differ among siblings in the same family (e.g., seeking out books from the school library or going to a reading tutor). The E component also includes error of measurement, and, importantly, currently unpredictable variations in development, sometimes called epigenetic noise (Molenaar, Boomsma, & Dolan, 1993). So the E component is not always necessarily environmental. Because the ACE model only includes these three main effects, it does not tell us about gene–environment interplay, which is discussed next.

Beyond the ACE Model: Gene–Environment Interplay

Going beyond the main effects of genes and environment captured by the ACE model, we can ask how genetic and environmental risk factors act together in the development of abnormal behavior, including dyslexia. As Rutter (2006) discusses, there are many kinds of interplay between genes and environments. Two broad classes of such interplay are GxE interaction and G-E correlation. In GxE interaction, the effect of independent genetic and environmental factors is synergistic rather than additive. There are three subtypes of GxE interaction: diathesis-stress, bioecological, and susceptibility. In a diathesis-stress GxE interaction, the effects of a risk genotype are increased by an environmental risk factor and vice versa. In a bioecological GxE interaction, the opposite pattern is observed: The effects of a risk genotype are stronger in a protective environment than in a risk environment. Finally, in a susceptibility GxE interaction, a susceptibility genotype leads to a worse outcome in a risk environment, but a better outcome in a protective environment, whereas a nonsusceptibility genotype is less affected by either type of environment.

Of these three types of GxE interaction, only a bioecological interaction has been found for dyslexia (Friend et al., 2009). Specifically, that study found that the heritability of dyslexia increased as parent education increases. This result suggests that the child's literacy environment is, on average, both more favorable and less variable as parent education increases, resulting in genetic risk factors playing a bigger role in a child's dyslexia. Conversely, as parent education decreases, the child's literacy environment is on average less favorable and more variable, resulting in environmental risk factors playing a bigger role in a child's dyslexia.

There is also increasing evidence for the importance of transactional processes in atypical development, in which the child and environment mutually alter each other over time. Gene–environment correlation is an example of such a transaction. Such transactions occur because children evoke different kinds of reactions from their environments (Scarr & McCartney, 1983) and select different kinds of environments for themselves. Not surprisingly, the individual characteristics that influence such reactions and selections are genetically influenced. There are three subtypes of G-E correlation: passive, evocative, and active (Scarr & McCartney, 1983). In the case of reading development, an example of a passive G-E correlation is the relation between parents' reading skill and the number of books in the home. Parents' reading skill is partly due to genes, and parents who are better readers on average have more books in their homes. Without any action on the part of their biological children, their literacy environment is correlated with their reading genotype, on average. In contrast, an evocative G-E correlation occurs when adults in a given child's environment notice their interests and talents and seek to foster them. In the case of reading development, an example of an evocative G-E correlation would be a parent or relative taking a child who likes to read to the library. Finally, an active G-E correlation occurs when children on their own initiative seek or avoid environments as a function of their genotype. Dyslexia provides a clear example of an active G-E correlation. Even before formal literacy instruction, children at genetic risk for later dyslexia who will later develop the disorder avoid being read to and spend less independent play time looking at books than their siblings who do not develop dyslexia (Scarborough, Dobrich, & Hager, 1991). As they get older, school-age children with dyslexia read dramatically fewer words per year than typically developing children (Cunningham & Stanovich, 1998), and this reduced reading experience negatively influences both their reading fluency and their oral vocabularies (Stanovich, 1986; Torgesen, 2005).

Molecular Genetics

Molecular methods—ones that rely on measuring DNA variations among individuals—test directly for genetic influences on a phenotype and now allow us to go beyond the indirect methods used in classical behavior genetics, which were just discussed. They also allow a direct test of whether behavior genetic results are valid. Briefly, molecular genetic studies of the etiologies of typical traits and disorders exploit two important facts about the genome. The first fact is that some rungs in the DNA ladder (where these rungs consist of pairs of the four chemical bases adenine (A), cytosine (C), guanine (G), and thymine (T)) differ across individuals in a species such that one individual may have the pair AG for one rung and another individual may have the pair CT (in humans, about 1 per 1,000 base pairs show differences across individuals on average; our genome has a total of about three billion pairs). Those base pairs that frequently differ across individuals are called single-nucleotide polymorphisms (SNPs). The second important fact is that the DNA segments (e.g., SNPs) on chromosomes are shuffled by recombination in the process of making individual sperm and egg cells (i.e., gametes). As a result of this shuffling, only DNA segments that are close together on the same chromosome will be inherited together, or linked. As a result, individuals in a species differ in their exact DNA sequences (except for identical twins), and some of these DNA differences lead to differences in behavior and other traits. By relating trait similarity to DNA similarity, we can eventually discover which DNA variants are important for a given trait.

Using molecular methods, dyslexia has been linked to nine risk loci (which are termed DYX1–DYX9, with DYX standing for dyslexia and the number indicating the order of discovery) through replicated linkage studies (Fisher & DeFries, 2002; McGrath, Smith, & Pennington, 2006), although not every study has replicated these results (Ludwig et al., 2008; Meaburn, Harlaar, Craig, Schalkwyk, & Plomin, 2008). For instance, Meaburn et al. (2008) used DNA pooling and over 100,000 SNPs to identify loci that distinguished a high reading versus a low reading group. Their few significant hits each accounted for very small amounts of the variance and did not include the best-replicated dyslexia loci.

More precise mapping methods have led to the identification of six candidate genes (termed C for candidate and followed by a number, again indicating the order of discovery) in some of the nine replicated risk loci (a risk locus is specified by its chromosome number out of the 23 human chromosomes, which of the two arms, short (p) or long (q), the risk locus is on, and an address on that arm indicated by a number). These six candidate genes are *DYX1C1* in the DYX1 locus on chromosome 15q21; *DCDC2* and *KIAA0319* in the DYX2 locus on chromosome 6p21; *C2Orf3* and *MRPL19* in the DYX3 locus on chromosome 2p16–p15; and *ROBO1* in the DYX5 locus on chromosome 3p12–q12. Studies of their role in brain development (Kere, 2011) in rodents has shown that *DYX1C1, DCDC2, KIAA0319*, and *ROBO1* affect prenatal processes of brain development, specifically neuronal migration (the movement of immature neurons from where they are first formed to their final destination in the brain) and the formation of connections once they reach that destination (e.g., neurite—axon and dendrite—outgrowth and guidance). More generally, these two processes of early brain development are each genetically controlled by a family or network of genes that interact with each other through molecular signals. In contrast, very little is known about the functions of the two DYX3 candidate genes. Two other studies have identified three new candidate genes for dyslexia (*MC5R, DYM*, and *NEDD4L*) (Scerri et al., 2010) on chromosome 18 and one shared with language impairment (*CMIP*) (Scerri et al., 2011), but these results need to be replicated.

In a recent review of the molecular genetics of dyslexia, Carrion et al. (2013) discuss the two molecular signaling networks already implicated in the development of dyslexia: neuronal migration and neurite outgrowth and guidance, as well as a third one, ciliary biology. Cilia are microscopic hair-like structures on the surface of cells, as in a paramecium, and their rhythmic movement turns out to play a role in the patterning of early brain development. Carrion et al. also discuss in detail the sometimes inconsistent evidence found across samples for the various candidate genes for dyslexia. This inconsistency is due partly to the fact that the mutations found in dyslexia are not in the regions of genes that directly code for the structure of proteins but in noncoding regions that affect expression levels of structural genes, sometimes with small and subtle effects, and partly to the fact that many of the samples in these studies are too small, as discussed later.

Nonetheless, the fact that these candidate genes interact with each other and act on the same molecular signaling pathways is a promising beginning for eventually discovering the many more genes that are likely to be involved and working out the early developmental biology of this disorder. We next turn to important recent developments in the molecular genetics of complex phenotypes that have implications for future research on the etiology of dyslexia.

Missing Heritability?

A potential criticism of behavior genetics (Wahlsten, 2012) is that molecular studies have identified only a very few of the many genes needed to account for the heritability found by twin studies for common traits like IQ or reading or height; this problem is called the problem of missing heritability (Manolio et al., 2009). Once microarrays (sometimes called gene or SNP chips) with large numbers of SNPs across the genome became readily available, researchers undertook genome-wide association (GWA) studies of complex phenotypic traits, like height, IQ, and common diseases. At the outset of this research, the common disease, common variant hypothesis was popular. This hypothesis predicted there would be prevalent genetic variants with large effect sizes contributing to common diseases. However, a typical result of many GWA studies was that very few SNPs produced significant associations with the phenotype being studied, and those few combined only accounted for a small proportion (at most 1%–3%) of the variance in the phenotype. This result for common diseases, like autism, schizophrenia, and diabetes, soundly rejected the common disease, common variant hypothesis.

Since these GWA studies were motivated by the fact the normal and abnormal traits being investigated had all demonstrated substantial heritabilities in behavior genetic twin studies, these disappointingly meager GWA results posed a puzzle. For instance, twin studies typically find a heritability for human height of around 0.90 and for IQ of around 0.50. Thus, the puzzle was the large gap between the small amount of genetic variance accounted for by GWA results and the large indirect estimates of this genetic variance based on previous behavior genetic twin studies. This gap was called missing heritability.

Several explanations were offered to explain this commonly observed phenomenon of missing heritability. These explanations included (1) very large number of genetic variants (i.e., alleles) with very small additive effect sizes (i.e., a highly polygenic etiology); (2) rare variants with large effects that are hard to detect with common SNPs; (3) copy number variations, which are new changes in the number of nucleotides (DNA base pairs) at a locus; (4) high levels of gene–gene interaction (called epistasis); and (5) overestimation of heritability by behavior genetic designs. Possibility (5) posed a serious threat to the validity of many decades of research in behavior genetics, and was embraced by some critics (e.g., Wahlsten, 2012), who asserted that the conclusion of moderate heritability for many human traits and disorders was fundamentally mistaken. As we will see, later empirical results have indicated that possibility (5) is quite unlikely, and that instead possibility (1) appears to hold for continuously distributed individual differences like those in height and IQ, and possibly dyslexia. Possibilities (2), (3), and (4) are more likely mechanisms to explain the missing heritability of severe developmental disorders, like autism and schizophrenia.

To understand why current GWA studies failed to find common variants affecting human traits like IQ and height and common human disorders, it is important to understand the relation between effect size and allele frequency for alleles that affect important aspects of human development. An important model for understanding this relation is called the mutation-selection model (Keller, 2008), which was proposed to explain why deleterious disorders like schizophrenia and autism persist at a fairly high rate (~1%) in the population. Both of these disorders reduce an individual's reproductive success (i.e., how many children they have), so natural selection should quickly eliminate common risk alleles with larger effect sizes. Hence, risk alleles with large effect sizes that persist in the gene pool will necessarily be rare. Thus, we need an explanation for why natural selection has not eliminated the risk alleles for common deleterious conditions like schizophrenia. The mutation-selection explanation proposes that new mutations balance the elimination of old risk alleles for such disorders, leading to a fairly stable prevalence of such disorders over time. So the mutation-selection model holds that common variants with large (i.e., detectable with current GWA studies) effect sizes on common disorders will not be found. Instead, such deleterious variants will be rare, and there will be an inverse relation between effect size and allele frequency, which are represented on a log scale in Figure 24.1. Effect size means what proportion of the variance in the phenotype is caused by the risk allele, and allele frequency is the prevalence of the risk allele in the population. As can be seen in Figure 24.1, most genetic variants affecting human traits will fall between the dotted lines. Those with big effect sizes, like the gene for PKU or HD, will be rare, whereas those affecting common, adaptive traits like height or IQ will have small effect sizes. Phenylketonuria and HD are examples of Mendelian diseases (i.e., ones caused by a mutation in a single gene, which can be recessive, as in PKU, or dominant, as is true for HD). In contrast, there will be very few common variants that affect common diseases, because selection will have eliminated them.

So the mutation-selection hypothesis explains missing heritability in GWA studies of adaptive traits with possibility (1), a highly polygenic etiology. The Meaburn et al. (2008) study of high versus low reading skill suggests the same is true for

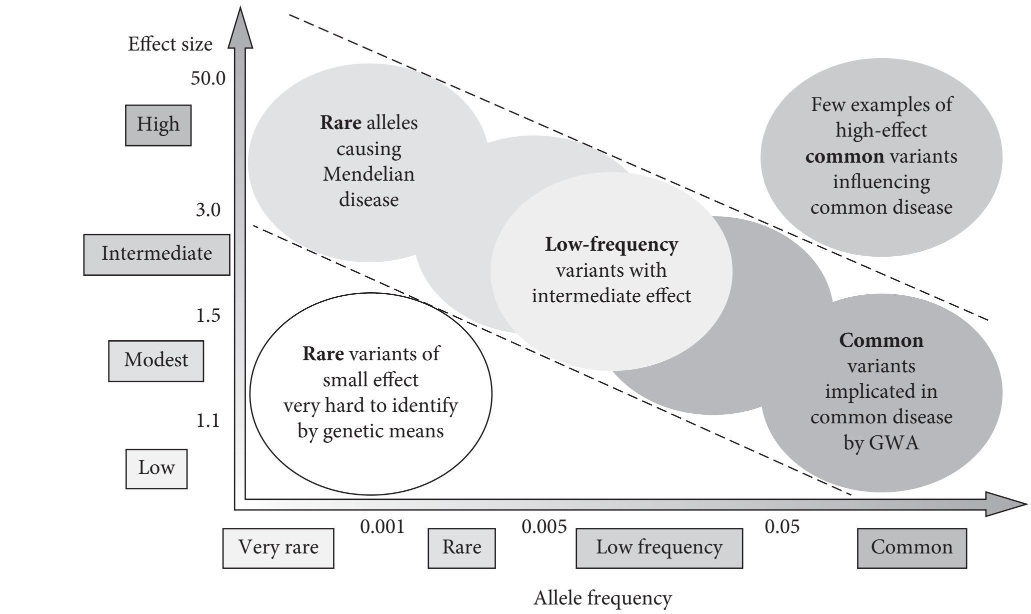


Fig. 24.1 Feasibility of identifying genetic variants by risk allele frequency and strength of genetic effect.

dyslexia because the few significant SNPs discovered accounted for very small amounts of variance in reading. If the mutation-selection hypothesis is true, then most existing GWA studies are dramatically underpowered to detect individual alleles with very small effect sizes. Genome-wide association studies have a very stringent threshold for significance (e.g., $p < 10^{-8}$), so only SNPs with a relatively large effect size (i.e., around 0.5% to 1.0% of the variance) will be detected unless the sample size is in the tens or hundreds of thousands of individuals.

To test whether previous twin studies overestimated heritabilities (i.e., the heritability is missing because it is not really there), different methods of analyzing GWA data are needed, ones that estimate the cumulative, additive effect size of all the SNPs in the database, not just the ones that cross the stringent threshold for significance. Such methods have been developed, and they exploit the fact that individuals in a GWA database, who are from different families, nonetheless vary in their degree of DNA sharing across all the SNPs in the analysis. Notice that these unrelated individuals do not share a common family environment (C) and are unlikely to share a unique environment (E), so that any phenotypic similarity is due mainly or exclusively to additive genetic similarity. Consequently, the relation between their genetic similarity and their phenotype similarity can be evaluated to give a direct, molecular estimate of heritability.

When this approach has been applied to GWA data for height (Yang et al., 2010) and IQ (Chabris et al., 2012; Davies et al., 2011; Plomin et al., 2013), SNP heritability estimates come closer to twin study heritability estimates, but there still is some missing heritability. The small amount of remaining missing heritability could be due to possibilities (2) rare variants, (3) copy number variations, or (4) epistasis in the list presented earlier. In sum, the phenomenon of missing heritability does not mean that indirect estimates of genetic influence on many typical and atypical human traits are wrong. It does mean, however, that very many alleles of very many genes may be involved in the etiologies of those traits, and that working out the many developmental pathways may be very difficult.

This discussion of missing heritability highlights things we need to know about the etiology of dyslexia. There are no published GWA studies of dyslexia, and so the missing heritability problem remains largely unexplored for both individual differences in typical reading and for dyslexia. Could reading and dyslexia provide exceptions to what is found for other phenotypes by revealing a common variant with a large effect size? That is a possibility, given that dyslexia does not clearly affect reproductive success, and so there might not be selection pressure against common risk genes for dyslexia. However, the results of Meaburn et al. (2008) argue against this possibility because that study did not find alleles for dyslexia with large effect sizes.

A relatively unexplored topic in the etiology of dyslexia is interaction between genes (called epistasis), which causes nonadditive heritability. One recent study has found such an interaction (Powers et al., 2013). These researchers examined the association of reading (and associated neuropsychological phenotypes) to SNPs within a regulatory region of DCDC2, as well as a risk region of KIAA0319. Individuals who had both risk factors showed disproportionately poorer reading than expected given the individual effect of each genetic risk factor alone.

Environmental Influences

Because the heritability of dyslexia is substantially less than 100%, we know there are environmental factors that contribute to the development of the disorder. However, there has been limited methodologically rigorous work testing which specific environments causally influence reading development. Possible candidates include the language and preliteracy environments that parents provide for their children, but unfortunately much of the research on these topics has used correlational rather than genetically sensitive designs (like twin and adoption studies). Thus parents with genetic risk for dyslexia may provide less literacy exposure to their children because of the G-E correlations discussed earlier, and so it is not clear that the environment plays a causal role in the child's reading outcome. This limitation is avoided in treatment studies that use random assignment. Results of such research has suggested that training parents in various home literacy activities promotes young children's vocabulary (a reading precursor; Lonigan & Whitehurst, 1998) and early reading skills (Sénéchal, this volume; Sylva, Scott, Totsika, Ereky-Stevens, & Crook, 2008). This work is broadly consistent with findings from twin studies demonstrating that, during the preschool years, individual differences in vocabulary and some other literacy precursors are more influenced by family environment than by genes (Byrne et al., 2009; Hayiou-Thomas, Dale, & Plomin, 2012). However, this work has also shown that over time, the relative importance of etiologic influences shifts, and by later school age, genetic influences on oral language and literacy predominate. Further work is needed to know whether the effects of home literacy environment on word reading persist beyond the beginning stages of literacy instruction.

Related research has used randomized controlled trials to study the effects of instructional type on reading development in alphabetic systems. This research has consistently shown that phonologically based instruction, which emphasizes explicit knowledge about letter–sound correspondences, is superior to other forms of literacy instruction that emphasize sight word recognition (e.g., whole-word instruction) or listening comprehension (e.g., whole-language instruction) in promoting word-level reading skills, particularly for children who are at risk for reading difficulties (Brown & Felton, 1990; Snowling & Hulme, 2011; Vellutino, Scanlon, Small, & Fanuele, 2006). Because literacy curricula vary across and sometimes within countries, instructional type can influence the risk of an individual child meeting standard diagnostic criteria for dyslexia.

What Comes Next?

Despite important progress, much remains to be done to fully understand the etiology of dyslexia. First, as discussed earlier, a GWA study of dyslexia has not been published (the closest approximation is the study by Meaburn et al., 2008), and the known loci do not account for most of the heritability of dyslexia found in twin studies—the so-called missing heritability problem. Second, although our understanding of the causes of the comorbidities of dyslexia has progressed, much remains to be done to identify loci that are shared and not shared with the comorbid disorders ADHD, LI, and SSD. Third, whether dyslexia shows any of the newly discovered genetic mechanisms found in other neurodevelopmental disorders—for example, copy number variations, parent of origin effects, and epigenetic effects—is mostly unknown. So far, it appears that copy number variations will be less important in dyslexia than in more severe neurodevelopmental disorders (Girirajan et al., 2011). Virtually nothing is known about parent of origin (i.e., the effect of a gene on the offspring depends on the sex of the parent who transmits the gene) or epigenetic influences (i.e., changes in gene expression that are not due to the DNA code itself) on dyslexia, which are two other mechanisms of inheritance (see Smith, 2011, for a review of why epigenetic influences are likely to be important in dyslexia and other language disorders). Fourth, much remains to be learned about the role of the environment in the cause of dyslexia and about gene–environment interplay. Finally, although cross-cultural research on dyslexia is robust, dyslexia has been less studied in lower SES groups than in people from more privileged backgrounds, in nonwhite ancestry groups than in white populations, and in children with a bilingual background (e.g., Hispanic American children). A universal account of normal and abnormal reading development needs to encompass these understudied groups.

Acknowledgment

This research was supported by a Learning Disability Research Center grant from NICHD (HD027802).

References

Bishop, D. V., McDonald, D., Bird, S., & Hayiou-Thomas, M. E. (2009). Children who read words accurately despite language impairment: Who are they and how do they do it? *Child Development, 80,* 593–605.

Boada, R., & Pennington, B. F. (2006). Deficient implicit phonological representations in children with dyslexia. *Journal of Experimental Child Psychology, 95,* 153–193.

Bosse, M. L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition, 104,* 198–230.

Brown, I. S., & Felton, R. H. (1990). Effects of instruction on beginning reading skills in children at risk for reading disability. *Reading and Writing: An Interdisciplinary Journal, 2,* 223–241.

Bruck, M. (1990). Word-recognition skills of adults with childhood diagnoses of dyslexia. *Developmental Psychology, 26,* 439–454.

Bruck, M. (1992). Persistence of dyslexics' phonological deficits. *Developmental Psychology, 28,* 874–886.

Bruck, M. (1993). Word recognition and component phonological processing skills of adults with childhood diagnosis of dyslexia. *Developmental Review, 13,* 258–268.

Buchholz, J., & Davies, A. (2007). Attentional blink deficits observed in dyslexia depend on task demands. *Vision Research, 47,* 1292–1302.

Byrne, B., Coventry, W. L., Olson, R. K., Samuelsson, S., Corley, R., Willcutt, E. G., . . . Defries, J. C. (2009). Genetic and environmental influences on aspects of literacy and language in early childhood: Continuity and change from preschool to grade 2. *Journal of Neurolinguistics, 22,* 219–236.

Caravolas, M., Volin, J., & Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children. *Journal of Experimental Child Psychology, 92,* 107–139.

Carrion-Castillo, A., Franke, B., & Fisher, S. E. (2013). Molecular genetics of dyslexia: An overview. *Dyslexia, 19,* 214–240.

Castles, A., Wilson, K., & Coltheart, M. (2011). Early orthographic influences on phonemic awareness tasks: Evidence from a preschool training study. *Journal of Experimental Child Psychology, 108,* 203–210.

Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., & Ingvar, M. (1998). The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain. *Brain, 121,* 1053–1063.

Chabris, C. F., Hebert, B. M., Benjamin, D. J., Beauchamp, J., Cesarini, D., van der Loos, M., . . . Laibson, D. (2012). Most reported genetic associations with general intelligence are probably false positives. *Psychological Science, 23,* 1314–1323.

Christopher, M. E., Hulslander, J., Byrne, B., Samuelsson, S., Keenan, J. M., Pennington, B. F., . . . Olson, R. K. (2013). The genetic and environmental etiologies of individual differences in early reading growth in Australia, the United States, and Scandinavia. *Journal of Experimental Child Psychology, 115,* 453–467.

Cunningham, A. E., & Stanovich, K. E. (1998). The impact of print exposure on word recognition. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (p. 235–262). Mahwah, NJ: Erlbaum.

Davies, G., Tenesa, A., Payton, A., Yang, J., Harris, S. E., Liewald, D., . . . Deary, I. J. (2011). Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Molecular Psychiatry, 16,* 996–1005.

Demonet, J. F., Taylor, M. J., & Chaix, Y. (2004). Developmental dyslexia. *Lancet, 363,* 1451–1460.

Deutsch, G. K., Dougherty, R. F., Bammer, R., Siok, W. T., Gabrieli, J. D., & Wandell, B. (2005). Children's reading performance is correlated with white matter structure measured by diffusion tensor imaging. *Cortex, 41,* 354–363.

Diagnostic and Statistical Manual of Mental Disorders. (2000). (4th-TR ed.) Arlington, VA: American Psychiatric Association.

Dougherty, R. F., Ben-Shachar, M., Deutsch, G. K., Hernandez, A., Fox, G. R., & Wandell, B. A. (2007). Temporal-callosal pathway diffusivity predicts phonological skills in children. *Proceedings of the National Academy of Sciences, 104,* 8556–8561.

Elbro, C., Borstrøm, I., & Petersen, D. (1998). Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly, 33,* 36–60.

Enard, W. (2011). FOXP2 and the role of cortico-basal ganglia circuits in speech and language evolution. *Current Opinion in Neurobiology, 21,* 415–424.

Facoetti, A., Corradi, N., Ruffino, M., Gori, S., & Zorzi, M. (2010). Visual spatial attention and speech segmentation are both impaired in preschoolers at familial risk for developmental dyslexia. *Dyslexia, 16,* 226–239.

Fisher, S. E., & DeFries, J. C. (2002). Developmental dyslexia: Genetic dissection of a complex cognitive trait. *Nature Reviews Neuroscience, 3,* 767–780.

Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P., & Pembrey, M. E. (1998). Localization of a gene implicated in a severe speech and language disorder. *Nature Genetics, 18,* 168–170.

Fowler, A. E. (1991). How early phonological development might set the stage for phoneme awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 97–117). Mahwah, NJ: Erlbaum.

Franceschini, S., Gori, S., Ruffino, M., Pedroll, K., & Facoetti, A. (2012). A causal link between spatial attention and reading acquisition. *Current Biology, 22,* 814–819.

Friend, A., DeFries, J. C., Olson, R. K., Pennington, B., Harlaar, N., Byrne, B., . . . Keenan, J. M. (2009). Heritability of high reading ability and its interaction with parental education. *Behavioral Genetics, 39,* 427–436.

Girirajan, S., Brkanac, Z., Coe, B. P., Baker, C., Vives, L., Vu, T. H., . . . Eichler, E. E. (2011). Relative burden of large CNVs on a range of neurodevelopmental phenotypes. *PLoS Genetics, 7,* e1002334.

Guttorm, T. K., Leppanen, P. H., Poikkeus, A. M., Eklund, K. M., Lyytinen, P., & Lyytinen, H. (2005). Brain event-related potentials (ERPs) measured at birth predict later language development in children with and without familial risk for dyslexia. *Cortex, 41,* 291–303.

Hadzibeganovic, T., van den Noort, M., Bosch, P., Perc, M., van Kralingen, R., Mondt, K., & Coltheart, M. (2010). Cross-linguistic neuroimaging and dyslexia: A critical view. *Cortex, 46,* 1312–1316.

Harlaar, N., Butcher, L. M., Meaburn, E., Sham, P., Craig, I. W., & Plomin, R. (2005). A behavioural genomic analysis of DNA markers associated with general cognitive ability in 7-year-olds. *Journal of Child Psychology and Psychiatry, 46,* 1097.

Hayiou-Thomas, M. E., Dale, P. S., & Plomin, R. (2012). The etiology of variation in language skills changes with development: A longitudinal twin study of language from 2 to 12 years. *Developmental Science, 15,* 233–249.

Hensler, B. S., Schatschneider, C., Taylor, J., & Wagner, R. K. (2010). Behavioral genetic approach to the study of dyslexia. *Journal of Developmental and Behavioral Pediatrics, 31,* 525–532.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2,* 127–160.

Hu, W., Lee, H. L., Zhang, Q., Liu, T., Geng, L. B., Seghier, M. L., . . . Price, C. J. (2010). Developmental dyslexia in Chinese and English populations: Dissociating the effect of dyslexia from language differences. *Brain, 133,* 1694–1706.

Jimenez, J. E., Siegel, L. S., O'Shanahan, I., & Ford, L. (2009). The relative roles of IQ and cognitive processes in reading disability. *Educational Psychology, 29,* 27–43.

Keller, M. C. (2008). The evolutionary persistence of genes that increase mental disorders risk. *Current Directions in Psychological Science, 17,* 395–399.

Kere, J. (2011). Molecular genetics and molecular biology of dyslexia. *Wiley Interdisciplinary Review of Cognitive Science, 4,* 441–448.

Klingberg, T., Hedehus, M., Temple, E., Salz, T., Gabrieli, J. D., Moseley, M. E., & Poldrack, R. A. (2000). Microstructure of temporo-parietal white matter as a basis for reading ability: Evidence from diffusion tensor magnetic resonance imaging. *Neuron, 25,* 493–500.

Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature, 413,* 519–523.

Lai, C. S. L., Gerrelli, D., Monaco, A. P., Fisher, S. E., & Copp, A. J. (2003). FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain, 126,* 2455–2462.

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A German-English comparison. *Cognition, 63,* 315–334.

Logan, J. A. R., Cutting, L., Schatschneider, C., Hart, S. A., Deater-Deckard, K., & Peterill, S. (2013). Reading development in young children: Genetic and environmental influence *Child Development, 84,* 2131–2144.

Lonigan, C. J., & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly, 13,* 263–290.

Ludwig, K. U., Schumacher, J., Schulte-Körne, G., Konig, I. R., Warnke, A., Plume, E., . . . Hoffmann, P. (2008). Investigation of the DCDC2 intron 2 deletion/compound short tandem repeat polymorphism in a large German dyslexia sample. *Psychiatry Genetics, 18,* 310–312.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53,* 1–14.

Manis, F. R., McBride-Chang, C., Seidenberg, M. S., & Keating, P. (1997). Are speech perception deficits associated with developmental dyslexia? *Journal of Experimental Child Psychology, 66,* 211–235.

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., . . . Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature, 461,* 747–753.

McGrath, L., Pennington, B. F., Shanahan, M. A., Santerre-Lemmon, L. E., Barnard, H. D., Willcutt, E. G., . . . Olson, R. K. (2011). A multiple deficit model of reading disability and attention-deficit/hyperactivity disorder: Searching for shared cognitive deficits. *Journal of Child Psychology and Psychiatry, 52,* 547–557.

McGrath, L., Smith, S. D., & Pennington, B. F. (2006). Breakthroughs in the search for dyslexia candidate genes. *Trends in Molecular Medicine, 12,* 333–341.

Meaburn, E. L., Harlaar, N., Craig, I. W., Schalkwyk, L. C., & Plomin, R. (2008). Quantitative trait locus association scan of early reading disability and ability using pooled DNA and 100K SNP microarrays in a sample of 5760 children. *Molecular Psychiatry, 13,* 729–740.

Molenaar, P. C., Boomsma, D. I., & Dolan, C. V. (1993). A third source of developmental differences. *Behavior Genetics, 23,* 519–524.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7,* 323–331.

Nation, K., Cocksey, J., Taylor, J. S., & Bishop, D. V. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *Journal of Child Psychology and Psychiatry, 51,* 1031–1039.

Nicolson, R. I., & Fawcett, A. J. (2007). Procedural learning difficulties: Re-uniting the developmental disorders? *Trends in Neurosciences, 30,* 135–141.

Oyama, S. (1985). *The ontogeny of information.* Cambridge, England: Cambridge University Press.

Paulesu, E., Demonet, J. F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., . . . Frith, U. (2001). Dyslexia: Cultural diversity and biological unity. *Science, 291,* 2165–2167.

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition, 101,* 385–413

Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development, 72,* 816–833.

Pennington, B. F., & Olson, R. K. (2005). Genetics of dyslexia. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 453–472). Oxford, England: Blackwell.

Pennington, B. F., Santerre-Lemmon, L., Rosenberg, J., MacDonald, B., Boada, R., Friend, A., . . . Olson, R. K. (2012). Individual prediction of dyslexia by single versus multiple deficit models. *Journal of Abnormal Psychology, 121,* 212–224.

Perfetti, C. A., Zhang, S., & Berent, I. (1992). Reading in English and Chinese: Evidence for a "universal" phonological principle. In R. Frost (Ed.), *Orthography, phonology,* *morphology, and meaning* (pp. 227–248). Amsterdam, the Netherlands: North-Holland.

Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *Lancet, 379,* 1997–2007.

Peterson, R. L., Pennington, B. F., Shriberg, L. D., & Boada, R. (2009). What influences literacy outcome in children with speech sound disorder? *Journal of Speech, Language and Hearing Research, 52,* 1175–1188.

Plomin, R., Haworth, C. M., Meaburn, E. L., Price, T. S., & Davis, O. S. (2013). Common DNA markers can account for more than half of the genetic influence on cognitive abilities. *Psychological Science, 24,* 562–568.

Powers, N. R., Eicher, J. D., Butter, F., Kong, Y., Miller, L. L., Ring, S. M., . . . Greuen, J. R. (2013). Alleles of a polymorphic ETV6 binding site in DCDC2 confer risk of reading and language impairment. *American Journal of Human Genetics, 93,* 19–28.

Puolakanaho, A., Ahonen, T., Aro, M., Eklund, K., Leppanen, P. H., Poikkeus, A. M., . . . Lyytinen, H. (2007). Very early phonological and language skills: Estimating individual risk of reading disability. *Journal of Child Psychology and Psychiatry, 48,* 923–931.

Puolakanaho, A., Ahonen, T., Aro, M., Eklund, K., Leppanen, P. H., Poikkeus, A. M., . . . Lyytinen, H. (2008). Developmental links of very early phonological and language skills to second grade reading outcomes: Strong to accuracy but only minor to fluency. *Journal of Learning Disabilities, 41,* 353–370.

Ramus, F. (2004). Neurobiology of dyslexia: A reinterpretation of the data. *Trends in Neurosciences, 27,* 720–726.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain, 126,* 841–865.

Raschle, N. M., Chang, M., & Gaab, N. (2011). Structural brain alterations associated with dyslexia predate reading onset. *Neuroimage, 57,* 742–749.

Richlan, F., Kronbichler, M., & Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: A quantitative meta-analysis of neuroimaging studies. *Human Brain Mapping, 30,* 3299–3308.

Rimrodt, S. L., Peterson, D. J., Denckla, M. B., Kaufmann, W. E., & Cutting, L. E. (2010). White matter microstructural differences linked to left perisylvian language network in children with dyslexia. *Cortex, 46,* 739–749.

Rodgers, B. (1983). The identification and prevalence of specific reading retardation. *British Journal of Educational Psychology, 53,* 369–373.

Rutter, M. (2006). *Genes and behavior: Nature-nurture interplay explained*. Oxford, England: Blackwell.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development, 61,* 1728–1743.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75–119). Timonium, MD: York Press.

Scarborough, H. S., Dobrich, W., & Hager, M. (1991). Preschool literacy experience and later reading achievement. *Journal of Learning Disabilities, 24,* 508–511.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype-environment effects. *Child Development, 54,* 424–435.

Scerri, T. S., Morris, A. P., Buckingham, L. L., Newbury, D. F., Miller, L. L., Monaco, A. P., . . . Paracchini, S. (2011). DCDC2, KIAA0319 and CMIP are associated with reading-related traits. *Biological Psychiatry, 70,* 237–245.

Scerri, T. S., Paracchini, S., Morris, A., MacPhie, I. L., Talcott, J., Stein, J., . . . Richardson, A. J. (2010). Identification of candidate genes for dyslexia susceptibility on chromosome 18. *PLoS One, 5,* e13712.

Scharff, C., & Haesler, S. (2005). An evolutionary perspective on FoxP2: Strictly for the birds? *Current Opinion in Neurobiology, 15,* 694–703.

Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine, 326,* 145–150.

Silani, G., Frith, U., Demonet, J. F., Fazio, F., Perani, D., Price, C., . . . Paulesu, E. (2005). Brain abnormalities underlying altered activation in dyslexia: A voxel based morphometry study. *Brain, 128,* 2453–2461.

Silva, P. A., McGee, R., & Williams, S. (1985). Some characteristics of 9-year-old boys with general reading backwardness or specific reading retardation. *Journal of Child Psychology and Psychiatry, 26,* 407–421.

Smith, S. D. (2011). Approach to epigenetic analysis in language disorders. *Journal of Neurodevelopmental Disorders, 3,* 356–364.

Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family risk of dyslexia is continuous: Individual differences in the precursors of reading skill. *Child Development, 74,* 358–373.

Snowling, M. J., & Hulme, C. (2011). Evidence-based interventions for reading and language difficulites: Creating a virtuous circle. *British Journal of Educational Psychology, 81,* 1–23.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360–406.

Stuebing, K. K., Barth, A. E., Molfese, P. J., Weiss, B., & Fletcher, J. M. (2009). IQ is not strongly related to response to reading instruction: A meta-analytic interpretation. *Exceptional Children, 76,* 31–51.

Sylva, K., Scott, S., Totsika, V., Ereky-Stevens, K., & Crook, C. (2008). Training parents to help their children read: A randomized control trial. *British Journal of Educational Psychology, 78,* 435–455.

Torgesen, J. K. (2005). Recent discoveries on remedial interventions for children with dyslexia. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 521–537). Oxford, England: Blackwell.

Torppa, M., Lyytinen, P., Erskine, J., Eklund, K., & Lyytinen, H. (2010). Language development, literacy skills, and predictive connections to reading in Finnish children with and without familial risk for dyslexia. *Journal of Learning Disabilities, 43,* 308–321.

Vellutino, F. R. (1979). *Dyslexia: Theory and research*. Cambridge, MA: MIT Press.

Vellutino, F. R., Scanlon, D. M., Small, S. G., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities, 39,* 157–169.

Wadsworth, S. J., Olson, R. K., & DeFries, J. C. (2010). Differential genetic etiology of reading difficulties as a function of IQ: An update. *Behavior Genetics, 40,* 751–758.

Wahlsten, D. (2012). The hunt for gene effects pertinent to behavioral traits and psychiatric disorders: From mouse to human. *Developmental Psychobiology, 54,* 475–492.

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91,* 415–438.

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., . . . Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics, 42,* 565–569.

Ziegler, J. C., Bertrand, D., Toth, D., Csepe, V., Reis, A., Faisca, L., . . . Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science, 21,* 551–559.

Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., . . . Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences, 109,* 11455–11459.

CHAPTER

25 How Learning to Read Influences Language and Cognition

Régine Kolinsky

Abstract

As illustrated in this handbook, a substantial body of work now exists that examines which factors and functions affect reading acquisition and reading proficiency and which brain areas are involved. The converse relationship—namely, which functions and brain areas are affected by literacy—has received far less attention, probably because reading acquisition lags speech and vision by several years and because the crucial comparison of illiterate adults with people who learned to read as adults is difficult to undertake. However, this chapter illustrates that learning to read has profound influences on the processing of spoken language and beyond the domain of language, in particular on visual nonlinguistic perception. The chapter discusses research with literate adults in these areas, including the influence of spelling knowledge on speech perception. It also covers research with illiterate adults and on people who first learned to read as adults.

Key Words: illiteracy, reading acquisition, literacy effects, language, vision, cognition, brain, plasticity

Is it the case, as declared by Frith (1998, p. 1011), that literacy is "literally changing the brain," and that "culture change(s) basic brain anatomy"? Activities linked to literacy (reading books, magazines, etc.) certainly lead to increased knowledge. But literacy per se—the ability to read and write[1]—may induce other, more fundamental changes. Here, the chief issue under discussion is whether literacy modifies cognition qualitatively beyond visual word recognition processes; that is, whether it changes the principles and organization of knowledge.

Learning to read enables the emergence of mechanisms (e.g., Grainger, Tydgat, & Isselé, 2010) and brain networks (e.g., Cohen et al., 2000) tuned to the processing of written strings, which must be connected with both the visual and the spoken language systems. Presumably, direct or indirect connections are also established with the semantic, reasoning, and executive functioning systems, so that reading acquisition might in principle influence all of these functions. Recently it has been proposed that learning to read modulates other systems not only by establishing new functional links (e.g., between orthography and phonology) but also by altering the intrinsic organization of some of these systems through a process of *neuronal recycling* (Dehaene & Cohen, 2007). According to this view, previous brain circuits involved in visual object recognition and spoken language processing must adapt to perform the new task of reading. In this chapter these potential effects of reading acquisition are examined first as regards the most studied domain of language, and then beyond language, for vision and some higher-level cognitive domains.

Obviously, the effects of reading acquisition must be distinguished from those of age and neural maturation as well as from those of formal education and culture. Estimating the proper effects of reading acquisition requires the comparison of groups that do not differ by age or cognitive performance correlated with maturation differences. As regards formal education, although school attendance and learning to read are usually associated in an individual's life, we can gain insight into the specific effects of literacy by comparing adults who remained illiterate for socioeconomic reasons with *late literates* (also called *ex-illiterates*), that is, people who first learned to read as adults in special literacy classes organized by the government, the army, or industry, many having been encouraged to do so by their employer or supervisor. Contrary to *early literates*—adults who learned to read as children and attended school for several years—illiterates and late literates never attended school in childhood. Moreover, as illiterates and late literates are from the same socioeconomic background, any performance or brain differences between them should not be contaminated by sociocultural factors. A distinctive approach is to study, in literates, the effects of script directionality or of the relation between phonological segments and their spelling.

The Effects of Reading Acquisition on Spoken Language

Script Directionality Influences Listening to Speech

The first experimental evidence of the impact of reading acquisition in a purely aural context came from a study showing that script directionality influences the perceived temporal position of an extraneous noise (a click) relative to the constituents of a spoken sentence. Previous work had shown that click location is influenced by the relative positions of the click and the sentence in the auditory space: The click is judged more often as occurring simultaneously with an earlier part of the sentence (i.e., is located earlier than its objective position) when it is perceived as being on the left of the speech in auditory space than when it is perceived as being on the right of the speech (Bertelson & Tisseyre, 1972; Fodor & Bever, 1965). Bertelson (1972) found that this effect is inverted if, rather than using English or French as in previous experiments, the experiment uses Hebrew (with Israelis), which is written from right to left. In addition, the Israeli participants in that study showed the same pattern as native speakers of French when tested in French (a language they did understand and read). Thus when presented with spoken sentences, literate people "listen from left to right versus right to left" (Bertelson, 1972) according to the directional properties of graphical representation and scanning habits linked to a specific language, which influence the spatial coding of heard speech. In the following years, several experimental studies showed that reading acquisition also changes the very nature of the representations of speech.

Reading Acquisition Induces New Explicit Representations of Speech

Texts are relatively independent of the context that characterizes effective oral communication and can be reviewed, refined, and reformatted, allowing readers to overcome the limitation on the amount of conscious reflection that can be done on spoken materials (e.g., Donald, 1993; Ong, 1982). Literacy would thus favor the decontextualization (Denny, 1991) or objectification (Olson, 1991) of language, and consequently the development of *metalinguistic abilities*, namely a reflective attitude with regard to language objects and their manipulation. For instance, literacy helps in realizing that words have no intrinsic relation to the things they stand for but are just arbitrary symbols. This is difficult to understand for young children (e.g., Berthoud-Papandropoulou, 1978) as well as for illiterate adults (Kolinsky, Cary, & Morais, 1987), who assert, for instance, that a cat has a longer name than a butterfly.

Among metalinguistic abilities, *phonological awareness* (or *metaphonological ability*) refers specifically to the understanding that spoken words can be broken down into smaller parts. This is a multilevel skill, depending on the unit that is considered. With phonemes, only alphabetic literacy plays a critical role in the development of explicit representations, or *phonemic awareness*. Indeed, it is only in alphabetic writing systems that the individual printed characters represent phonemes (see Kessler & Treiman, this volume). Neither preliterate children (e.g., Liberman, Shankweiler, Fisher, & Carter, 1974) nor adults who have never learned an alphabet (either complete illiterates, e.g., Morais, Cary, Alegria, & Bertelson, 1979, Morais, Bertelson, Cary, & Alegria, 1986, or literates in a nonalphabetic system, e.g., Read, Zhang, Nie, & Ding, 1986) are able to tell that there are three "sounds" in the word *cab*, and all are very poor at phoneme deletion (e.g., /kæb/→/æb/; in all studies around 20% average correct responses in illiterates or nonalphabetic readers vs. more than 70% in late alphabetic literates), reversal (e.g., /kæb/→/bæk/), and detection (e.g., of /k/ in /kæb/). Awareness of higher-level units such as syllables or rhymes does not depend so critically on reading, as differences are smaller than with phonemes, but is improved by it. For example, in

Morais et al. (1986), late literates scored better than illiterates in syllable deletion (85% vs. 55% correct, respectively) and rhyme detection (92% vs. 67%, respectively). Notably, the representations involved in metaphonological tasks differ from perceptual representations: The same illiterate people who perform poorly on phonemic awareness tasks can discriminate almost perfectly pairs like /ta–sa/ or /pa–ba/ (Adrián, Alegria, & Morais, 1995; Scliar-Cabral, Morais, Nepomuceno, & Kolinsky, 1997).

Orthographic Knowledge Influences Metaphonological Performance in Literates

Not surprisingly, as metaphonological representations are closely linked to reading acquisition (e.g., Adams, 1990; Ehri et al., 2001), orthographic knowledge influences performance in purely auditory metaphonological tasks. Various orthographic effects in speech processing rely on the fact that in many alphabetic writing systems, the relationships between letters and phonemes are often not one-to-one, for reasons discussed by Kessler and Treiman (this volume). In addition to inconsistency in spelling-to-sound mapping (e.g., in English ‹OUGH› can be pronounced as in *cough, through, tough*), a phenomenon that affects reading performance, there is also inconsistency in sound-to-spelling mapping, namely multiple ways to spell a specific pronunciation, as for instance the rhyme of *toast* and *ghost* (e.g., Stone, Vanhoy, & Van Orden, 1997). The latter phenomenon mainly affects auditory processing (Ziegler, Petrova, & Ferrand, 2008).

In metaphonological tasks, inconsistencies in sound-to-spelling mapping lead to several effects, including *orthographic congruency effects*, with better performance, faster responses, or both when orthography and phonology lead to the same response than when they lead to opposite, competing responses. For instance, Seidenberg and Tanenhaus (1979) reported that literate adults take less time to decide that two spoken words rhyme when their spellings are similar (e.g., *toast–roast*) than when they are dissimilar (e.g., *toast–ghost*), and conversely for negative decisions (e.g., faster decisions for *leaf–ref* than *leaf–deaf*). In addition, orthographic inconsistency of phonemes leads to *orthographic consistency effects*. Indeed, in phoneme detection (a task that involves a strong metalinguistic component, as illustrated by illiterate adults' difficulties; Morais et al., 1986), literate adult listeners more rapidly detect orthographically consistent phonemes, for which there is only one spelling in the language, than orthographically inconsistent phonemes (Frauenfelder, Seguí, & Dijkstra, 1990), which are spelled in different ways in different words (e.g., /k/ in French words, as it is realized orthographically by the letters ‹c›, ‹k›, ‹cq›, or ‹qu›). Metaphonological performance is also influenced by the complexity of the relationship between phonemes and the letters representing them: Phoneme deletion and phoneme reversal performances are better when there is a one-to-one relationship between the phonemes and their spellings (e.g., deleting /d/ from *dentist*) than when there is a complex correspondence, as when deleting /n/ from *knuckle* or /k/ from *queen* (Castles, Holmes, Neath, & Kinoshita, 2003). Even letter names affect metaphonological judgments: In phoneme counting, syllables that are letter names (e.g., /ɑr/) are judged to contain fewer "sounds" than syllables that are not letter names (Treiman & Cassar, 1997).

Explicit phonological judgments about the structure of syllables are also shaped by orthographic representations. When aurally blending two consonant-vowel-consonant (CVC) monosyllabic words into a new CVC word (cf. Treiman, 1983), Portuguese adults prefer C/VC blends when the word spellings end with a consonant, as in *bar–mel*, /bar mɛl/, but prefer CV/C blends when the word spellings end with a mute ‹e›, as in *cure–pele*, /kur pɛl/ (Ventura, Kolinsky, Brito-Mendes, & Morais, 2001). Furthermore, a study of natives of Thai, a language in which tones are lexically contrastive and orthographically marked (but not orthographically consistent), showed that the influence of spelling knowledge extends beyond sublexical units. Indeed, literate Thai listeners show an orthographic congruency effect at the suprasegmental level, with better performance when the tone and the tone marker lead to the same response than when they lead to competing responses (Pattamadilok, Kolinsky, Luksaneeyanawin, & Morais, 2008).

Thus when becoming literate, listeners change the way in which they perform metaphonological tasks and use spelling knowledge in purely aural situations. An important question is whether they use this knowledge either in addition to or instead of their phonological skills. This issue has been hotly debated, as the latter possibility may cause researchers to revisit the role of phonological awareness in reading acquisition. Indeed, according to some researchers, phonological awareness does not represent a distinct set of spoken-language skills that is directly related to reading acquisition. Instead, the association between the ability to manipulate the sounds of spoken language and literacy acquisition may reflect the fact that once individuals acquire reading and spelling skills they change the way in which they perform phonological awareness tasks, using their orthographic skills to arrive at a solution. So on this account, the association between phonological awareness and literacy acquisition arises because both are, to a greater or lesser extent, indices of orthographic skill (e.g., Castles et al., 2003; Castles & Coltheart, 2004; but see Hulme, Caravolas, Malkova, & Brigstocke, 2005, for experimental arguments against the idea that phoneme manipulation ability can only develop as a consequence of orthographic—i.e., letter–sound correspondence—knowledge). This question is connected to the issue of the automaticity of the activation of orthography by speech: Does spelling knowledge become inseparable from phonological knowledge, or is it chiefly used strategically when useful?

Some studies have reported that orthographic representations are activated even when disadvantageous to performance. For example, in the phoneme deletion task used by Castles et al. (2003), adults did not improve their performance on complex items (e.g., deleting /n/ in *knuckle*) when these items were presented in pure rather than mixed blocks. Yet in pure blocks participants could have adopted a strategy that maximizes performance by not spelling the items, given the deleterious consequences in that case. But several orthographic effects occur only when the stimuli direct participants' attention to spelling, which could potentially invoke strategic effects. This is the case, for instance, with the orthographic consistency effect in phoneme detection (cf. Frauenfelder et al., 1990), which only occurs when spelling is rendered salient by the presence of many irregularly spelled words like *kneel, cough*, and *pyjamas* (Cutler, Treiman, & van Ooijen, 2010). Similarly, the orthographic congruency effect in rhyme judgment is eliminated when nonrhyming words with similar spelling (e.g., *leaf–deaf*) are not presented or when many filler items are added (Damian & Bowers, 2010). In addition, metaphonological studies using event-related brain potentials (ERPs) showed that orthographic congruency effects emerge relatively late in the course of processing, much later than phonological effects (in rhyme judgment: Pattamadilok, Perre, & Ziegler, 2011; Yoncheva, Maurer, Zevin, & McCandliss, 2013; in initial phoneme same-different judgment: Lafontaine, Chetail, Colin, Kolinsky, & Pattamadilok, 2012).

Nevertheless reading acquisition reorganizes a large brain network that includes phonological areas. As a matter of fact, a functional magnetic resonance imaging (fMRI) study (Brennan, Cao, Pedroarena-Leal, McNorgan, & Booth, 2013) showed that in aural rhyme judgment, brain activation is greater in adults than in eight- to 12-year-old children reading an alphabet (but not in readers of Chinese; see Kessler & Treiman, this volume for further discussion of its writing system), especially for words with conflicting orthography such as *pint–mint*. This occurs not only in inferior frontal areas (typically involved in phonological awareness tasks, e.g., Burton, Small, & Blumstein, 2000; Zatorre, Meyer, Gjedde, & Evans, 1996) but also in left hemisphere phonological areas (superior temporal gyrus). As discussed in the next two sections, the impact of spelling knowledge on speech processing is, in fact, much more profound than originally suspected.

Orthographic Knowledge Influences Spoken Word Recognition

There are reliable orthographic effects in aural word recognition tasks. Ziegler and Ferrand (1998) first reported an orthographic consistency effect in auditory lexical decision ("is a spoken item a word or not?"): responses to words such as *deep*, which include rimes that can be spelled differently in other words (e.g., *heap*), are slower and less accurate than responses to words with rimes that are spelled only one way. This effect has been replicated in several languages (e.g., French: Pattamadilok, Morais, Ventura, & Kolinsky, 2007; Portuguese: Ventura, Morais, Pattamadilok, & Kolinsky, 2004; English: Ziegler et al., 2008) and tasks (semantic and gender decision: Pattamadilok, Perre, Dufau, & Ziegler, 2009; Peereman, Dufour, & Burt, 2009).

Contrary to the influence of spelling in metaphonological judgments, the orthographic effects in recognition tasks take place rapidly in the course of processing, unfolding online with the word recognition process. This conclusion is supported, among other things, by ERP data. In semantic judgment (Pattamadilok et al., 2009) and lexical decision tasks (Perre, Pattamadilok, Montant, & Ziegler, 2009; Perre & Ziegler, 2008), the ERP orthographic consistency effect is time-locked to the orthographic inconsistency (e.g., earlier with the word French *rhume*, in which the initial /ry/ has multiple spellings, than with the word *noce*, in which the final /ɔs/ is inconsistent), and it starts before the onset of the frequency effect. Thus, orthography is activated early enough to modulate the core processes of lexical access.

Yet exactly how orthographic knowledge modulates speech processing is under debate. According to the *online account* (e.g., Ziegler & Ferrand, 1998), hearing a spoken word activates its corresponding orthographic code via cross-modal links; that is, through bidirectional connections between the spoken language (phonological) and visual (orthographic) systems. More precisely, in the *bimodal interactive activation model* (Grainger & Ferrand, 1996), there are bidirectional connections at both the lexical and sublexical (e.g., rhyme) levels. Words with consistent spellings thus benefit from self-consistent feedback from orthographic to phonological representations. For words with inconsistent spellings, in contrast, there is conflict at the sublexical level between several possible spellings and hence reduced feedback from orthographic to phonological representations. Alternatively, according to the *offline account*, orthographic effects take place within the phonological system itself. They reflect learning effects that happen during the course of learning to read and permanently alter the nature of the phonological representations (e.g., Muneaux & Ziegler, 2004; Taft, 2006, 2011).

The offline account is related to the *lexical restructuring* hypothesis (e.g., Garlock, Walley, & Metsala, 2001; Metsala, 1997), according to which phonological representations undergo important changes throughout language development. With children's oral vocabulary growth, the representations of lexical items become more detailed (more phonemic) with increasing pressure to discriminate between more and more similar-sounding words (*phonological neighbors*). For instance, recognizing the spoken word *dad* will require a more detailed representation for a child who also has acquired the words *bad, pad, mad, did*, and so on, than for a child who has not. Lexical restructuring depends not only on vocabulary size but also on the words' characteristics: High-frequency words are generally acquired earlier and hence undergo restructuring earlier than low-frequency words. Among the latter, only those with many phonological neighbors need to be finely represented, leading to an interaction between word frequency and number of phonological neighbors in word recognition tasks. For instance, Metsala (1997) used a gating task, in which listeners are presented with increasingly longer segments of a spoken word while attempting to identify it. She found that 7- to 11-year-old children and adults performed better (i.e., needed less input for recognition) for high-frequency words from sparse, as opposed to dense, neighborhoods, whereas they did better for low-frequency words from dense neighborhoods. The idea that children process words in a more holistic manner and that representations become more segmental with lexical growth was supported by the fact that the smallest developmental difference was found for high-frequency words from dense neighborhoods and the greatest developmental difference for low-frequency words from sparse neighborhoods, which are supposedly more holistically represented and the latest to undergo segmental restructuring. Although reading acquisition was not explicitly mentioned in the lexical restructuring hypothesis, it has been suggested that learning about letter–sound correspondences, and hence about phonemes, may make lexical representations more detailed in readers of an alphabetic script (e.g., Goswami, 2000). Yet contrary to this idea, illiterate adults have a phonologically restructured auditory lexicon similar to the one of literates, displaying the same interaction between word frequency and number of phonological neighbors (Ventura, Kolinsky, Fernandes, Querido, & Morais, 2007). Thus, phonological restructuring of the lexicon occurs in the absence of literacy. Ventura et al.'s finding argues against the idea that developmental lexical restructuring is mostly influenced by orthographic representations, but it does not refute the more general assumption that orthography impacts the phonological system.

Until recently, data supporting the offline account were scarce and disputable, as they were collected in situations that either involve phonemic awareness (e.g., the neighbor generation task used by Muneaux & Ziegler, 2004; see discussion in Ventura et al., 2007) or that use written strings (Taft, 2006), which may generate phonological codes different from those of speech. But more recent studies have shown that the orthographic consistency effect in lexical decision takes place within the phonological system itself: The cortical generator of the ERP effect sits within the vicinity of the left auditory cortex (Perre et al., 2009), and transcranial magnetic stimulation applied to an area involved in phonological processing (left supramarginal gyrus) cancels the effect (Pattamadilok, Knierim, Duncan, & Devlin, 2010).

In addition, fMRI studies suggest that both the online and offline mechanisms exist, with their relative involvement depending on the task. A study comparing illiterate with late and early literate adults disclosed both effects (Dehaene, Pegado, et al., 2010). On the one hand, actively processing speech in lexical decision, but not passively listening to spoken sentences, activates the visual word form area (VWFA, Cohen et al., 2000), the area of the left ventral occipitotemporal cortex (in the fusiform gyrus) involved in written word processing. This activation is orthographic rather than semantic, as it occurs in early and late literates but not illiterates. In literates, the recruitment of the VWFA has also been observed in other demanding tasks requiring selective attention to and analysis of the phonology of complex speech stimuli (e.g., when making rhyme judgments on words overlaid with tones; Yoncheva, Zevin, Maurer, & McCandliss, 2010). On the other hand, in both passive listening to spoken sentences and auditory lexical decision, there is a huge increase in fMRI activation of the planum temporale in literates compared with illiterates (see similar results in literate vs. preliterate age-matched children in Monzalvo & Dehaene-Lambertz, 2013). The planum temporale, like the surrounding superior temporal cortex, probably houses relatively abstract phonemic representations, as it encodes acoustic changes that are crucial for the categorical perception of speech (e.g., Chang et al., 2010; Mesgarani, Cheung, Johnson, & Chang, 2014) and also responds during silent lip reading (Calvert et al., 1997). The increase in planum temporale activation found in literate compared with illiterate adults may therefore indicate that reading acquisition enhances this kind of abstract phonological coding.

However, literacy is probably not like a "virus" that "infects all speech processing," as proposed by Frith (1998, p. 1011). Indeed, several perceptual phenomena are immune to the influence of literacy. In addition to being able to make fine phonetic discriminations (Adrián et al., 1995; Scliar-Cabral et al., 1997), like literates, illiterate adults experience slip-of-the-ear errors involving consonantal phonemes, revealing similar implicit representations of the perceptual constituents of speech (Morais & Kolinsky, 1994; see also Morais, Castro, Scliar-Cabral, Kolinsky, & Content, 1987, for errors involving phonetic features). Yet a study investigating categorical perception of speech sounds pointed to potential fine-grained differences between illiterate and literate adults (Serniclaes, Ventura, Morais, & Kolinsky, 2005). Categorical perception of speech means that only differences between identified phonemic categories (e.g., between phonemes identified as /b/ or as /d/) can be distinguished, not within-category variants (e.g., between two physically different sounds, both identified as /b/, Liberman, Harris, Hoffman, & Griffith, 1957). Categorical perception per se is thus estimated through the relation between performance in identification (obtained, e.g., through labeling) and discrimination (e.g., same-different judgment) tasks. This relation is the same in illiterate and literate adults. Yet, literates show a steeper identification slope than illiterates. Although in Serniclaes et al.'s study this effect could be attributed to a lexical bias (one of the continuum end points was a word), similar results were reported for adults and 6- to 8-year-old children (Hoonhorst et al., 2011). There was no effect of age on the relation between identification and discrimination performances, but boundary precision increased with age and was correlated with reading level. Thus although the data do not confirm the strong hypothesis according to which perceptual categorization of speech sounds depends on reading acquisition (Burnham, 2003), they suggest that literacy helps in finely tuning phonemic boundaries and hence in increasing the precision of phoneme identification.

Reading Acquisition Influences Short-Term Memory Codes and Performance

A common view is that oral memory has been traded off against literacy (e.g., Cole, Gay, Glick, & Sharp, 1971). This view was first articulated by Plato who, in *Phaedrus*, expressed concern about what he called the inhuman nature of writing, stating that written words have a destructive effect on human memory (cf. Ong, 1982). Yet poor verbal short-term memory (STM) is usually observed in illiterate adults, who display low word and digit spans (e.g., Kosmidis, Zafiri, & Politimou, 2011; Morais et al, 1986). The origin of this effect is unclear. Although it may be partly ascribed to formal education rather than literacy, as late literates also display lower word spans than early literates (Morais et al., 1986), there seems to be an additional slight benefit specifically due to literacy. For example, late literate adults who never attended school themselves but learned to read at home with their children have better forward digit-span scores than illiterates (Kosmidis et al., 2011), whereas no effect of literacy is observed with nonverbal materials (in forward spatial span, Kosmidis et al., 2011).

As discussed by Ardila et al. (2010), illiterate participants' poor recall may reflect inefficient encoding and retrieval strategies or poor organization of the material to be learned, as recall requires considerable self-initiated activity and executive skills.

The latter view is supported by the fact that illiterates are quite good on word recognition tests (telling which ones among different spoken words were previously presented, Ardila, Ostrosky-Solis, & Mendoza, 2000). Actually, illiterates' poor STM performance could reflect the fact that they were usually tested on ordered recall. Whether literacy specifically enhances memory for order, as opposed to memory for items, remains to be investigated. However, the reverse association has been reported: In kindergarteners, order (but not item) STM capacity predicts independent variance in nonword decoding abilities at the end of first grade (Martinez Perez, Majerus, & Poncelet, 2012). As regards encoding, literacy may improve it in two ways. First, it may improve phonological storage by affording more finely tuned phonological representations. Illiterate adults spontaneously use phonological codes in STM: Like literates (e.g., Baddeley, 1966; Conrad & Hull, 1964), they display a *phonological similarity effect* in ordered recall of lists of words, with poorer performance for rhyming lists than for nonrhyming ones (Morais et al., 1986). Yet illiterate adults seem to differ from literates at a finer grain size (phonemic boundaries, cf. Serniclaes et al., 2005), and this may lead to inaccurate identification of phonemes, at least in the absence of lexical support. Consistent with this idea is the fact that, in immediate repetition, illiterate adults perform poorly on pseudowords and do not activate the same brain regions as literates, but are quite good on words, with no group difference in neural activation (e.g., Castro-Caldas, Petersson, Reis, Stone-Elander, & Ingvar, 1998). Also consistent with the notion that literacy improves phonological storage are data showing that reading at 6 years predicts growth in nonword repetition between 6 and 7 years (Nation & Hulme, 2011). Moreover, in literates, spelling knowledge helps in maintaining the representation of spoken strings in STM. This has been demonstrated in serial recall, in which orthographic representations modulate the phonological similarity effect (Pattamadilok, Lafontaine, Morais, & Kolinsky, 2010). Compared with words that share neither a phonological nor an orthographic rime, literates' performance is less affected when words rhyme but have different spellings (as in the French *laine, gêne, traîne*, etc., all ending with /ɛn/) than when they both rhyme and have the same spelling (as in the French *classe, brasse, chasse*). Thus, inter-item orthographic dissimilarity reduces the detrimental effect of phonological similarity.

Reading Acquisition Induces Anatomical Changes

In addition to establishing a functional link between phonological and orthographic representations, literacy also leads to structural changes in brain connectivity. As a matter of fact, several studies have identified structural brain differences in late and early literate adults compared with illiterates. These differences, as suggested by functional connectivity analyses, seem to reflect both increased inter hemispheric functional connectivity (shown by thickening of the splenium or the isthmus of the corpus callosum, Carreiras et al., 2009; Castro-Caldas et al., 1999; Petersson, Silva, Castro-Caldas, Ingvar, & Reis, 2007) and strengthened intra hemispheric functional coupling (at the level of the left arcuate fasciculus; Thiebaut de Schotten, Cohen, Amemiya, Braga, & Dehaene, 2014) between the visual and phonological processing areas. Together with studies that showed increases in gray matter density in several brain regions involved in reading (Carreiras et al., 2009; Castro-Caldas et al., 1999; Petersson et al., 2007), these data strongly support the idea that literacy changes basic brain anatomy.

The Influence of Reading Acquisition on Spoken Language: A Brief Summary

Literacy influences many spoken language processes—not only metaphonological skills but also word recognition and verbal memory processes and perhaps even some perceptual processes. Some of these effects seem to reflect online influence of orthographic or metaphonological representations, whereas others suggest offline restructuring of lexical representations. If both online and offline mechanisms are involved, there might be multiple types of phonological representation in the lexicon, including orthographically (or metaphonologically) restructured ones (see discussions in e.g., Ranbom & Connine, 2011; Taft, 2011). In any case, these changes are also reflected at the structural level, in terms of both white- and gray-matter density. In fact, these enhanced connections reveal a bidirectional (sound-to-sight and sight-to-sound) pathway: In literate but not illiterate people, the language network of left temporal and inferior frontal regions activates almost identically to written and spoken language (Dehaene, Pegado, et al., 2010). Thus the acquisition of reading gives us access, from vision, to the spoken language system, and conversely spoken-language processing is modified by literacy.

The deep influence of literacy on spoken language is remarkable, as reading lags speech acquisition by several years and depends on explicit teaching. Frith (1998, p. 1012) wondered whether "learning to read has an equally transforming effect on processes underlying visual perception and thinking." In the following sections it is shown that the impact of literacy definitely goes beyond auditory skills and the language domain.

The Effects of Reading Acquisition on Visual Processing

The main effect of reading acquisition is that it allows the emergence of brain structures tuned to the processing of written strings (e.g., Cohen et al., 2000), thereby creating an interface through which linguistic inputs can be interpreted through vision, as already mentioned. Reading acquisition also qualitatively alters visual processes. Letter strings benefit from flexible position coding, leading to more difficulties in differentiating sequences with transposed letters such as ‹NTDF›–‹NDTF› than sequences with replaced letters such as ‹NSBF›–‹NDTF› (for a review, see Duñabeitia, Dimitropoulou, Grainger, Hernández, & Carreiras, 2012), an effect that is not observed in illiterate adults (Duñabeitia, Orihuela, & Carreiras, 2014). In addition, letter-string processing involves a specialized system that reduces the spatial extent of *crowding* for letters in words, limiting the integration of inappropriate features from neighboring stimuli (e.g., Grainger et al., 2010). Consistently, there is less integration with a surrounding geometrical shape for letters than nonletters (Van Leeuwen & Lachmann, 2004). Indeed, facilitation is smaller for letters than nonletters when the target stimulus is surrounded by a shape with a similar global contour than when it is surrounded by a shape with a different global contour (see Figure 25.1). This effect is also observed in illiterate adults with some knowledge of letters (Fernandes, Vale, Martins, Morais, & Kolinsky, 2014). Notably, the next sections illustrate that literacy also alters nonlinguistic visual processes.

Script Directionality Influences Visual Scanning and Spatial Associations

Script direction influences visual scanning not only of text (e.g., Pollatsek, Bolozky, & Rayner, 1981) but also of nonlinguistic stimuli (for a review, see Chokron, Kazandjian, & De Agostini, 2009; see also Bramão et al., 2007, for a comparison of literate and illiterate adults). The directional habits associated with text and numbers also contribute to the spatial representation of numbers: People who read words and numbers from left to right associate small numbers with the left space and large numbers with the right space (the *SNARC effect*, Dehaene, Bossini, & Giraux, 1993), whereas people reading from right to left show the reversed effect (Shaki, Fischer, & Petrusic, 2009). The direction of the writing system even affects the axis used to represent time in terms of space, for example by modulating how people place sets of cards (e.g., egg, chick, chicken) in temporal order (Bergen & Chan Lau, 2012), as well as the visual representations of action events, with literates, but not illiterates, showing a script-dependent spatial bias (e.g., Dobel, Enriquez-Geppert, Zwitserlood, & Bölte, 2014).

Reading Acquisition Induces Neural Competition in the Left Fusiform Gyrus

Because the VWFA is involved in written word processing in literates (e.g., Cohen et al., 2000), it is worth asking what role this brain area plays prior to reading acquisition. In illiterate adults the VWFA is not inactive, but strongly responsive to nonlinguistic pictures, particularly to faces. With increasing literacy, cortical responses to faces become restricted to a somewhat smaller area in the left fusiform gyrus

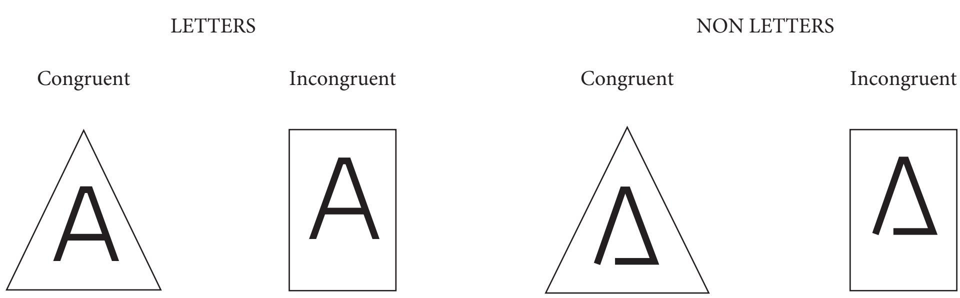


Fig. 25.1 Examples of material used by Fernandes, Vale, Martins, Morais and Kolinsky (2014; after Van Leeuwen & Lachmann, 2004): letters and nonletter shapes in congruent and incongruent surroundings.

and increase in the right fusiform gyrus (Dehaene, Pegado, et al., 2010). Thus reading acquisition induces neural competition between written words and other object categories, notably faces, leading to stronger right-hemispheric lateralization for faces in literate compared with illiterate adults. A similar shift of face responses toward the right hemisphere is observed in developmental studies using fMRI (Monzalvo, Fluss, Billard, Dehaene, & Dehaene-Lambertz, 2012), ERP recordings (Li et al., 2013), and behavioral hemifield lateralization (Dundas, Plaut, & Behrmann, 2013), suggesting that word lateralization, which emerges earlier in development, may drive later face lateralization.

Current studies aim at identifying the behavioral consequences of this process of neural competition between written strings and faces. Indeed, the stronger right-hemispheric lateralization for face processing with literacy raises the possibility that reading acquisition makes face processing more holistic in literates, as holistic (configurational) face processing is mainly implemented in the right fusiform gyrus (e.g., Rossion et al., 2000). Yet recent data on the *composite face effect* do not support this idea. The composite face effect reflects holistic face processing, as it shows that the parts of a face cannot be perceived independently from the whole face. Indeed, composite faces in which the two halves belong to two different face identities lead to a visual illusion (Young, Hellawell, & Hay, 1987). For instance, in a same-different matching task on pairs of composite faces, identical bottom face halves are perceived as being different when their top halves belong to different faces, an illusion that disappears when the bottom halves are spatially offset. Although literates may be expected to present a stronger composite face effect than illiterates, the opposite was observed: Literates were better at deciding whether the bottom halves of faces are the same or different without being distracted by the top part of the images (Ventura et al., 2013). This suggests that literacy improved an analytic strategy of attending to pictures. As a similar effect was observed with houses, it probably does not reflect the change in the lateralization of face processing but a general impact of literacy, which may bring more flexibility in reducing the influence of holistic processing when this is detrimental to the task.

Reading Acquisition in the Latin Script Pushes People to Unlearn Mirror Invariance

Most natural categories are invariant for left-right inversion, and hence lateral reversals convey little or no information about the identity of natural objects. Accordingly, there exists an intermediate stage of recognition in the ventral visual cortex where responses to pictures of objects are invariant to left-right mirror symmetry (e.g., Dehaene, Nakamura et al., 2010; Pegado, Nakamura, Cohen, & Dehaene, 2011). Yet mastering a script that includes mirror-image characters (e.g., in the Latin script, ‹p› ‹q› and ‹b› ‹d›) requires taking mirror-image contrasts into account. It pushes readers of these scripts to unlearn mirror invariance. At the brain level, this is reflected by the fact that the VWFA, which is the site of the visual system with the strongest mirror invariance for familiar objects, does not perform mirror-image generalization for words (Dehaene, Nakamura et al., 2010) or letters (Pegado et al., 2011). Behaviorally, this process of unlearning mirror invariance generalizes to nonlinguistic materials. Compared with readers of scripts that include mirror-image characters, both fully illiterate adults (e.g., Kolinsky et al., 2011) and readers of scripts that do not include mirror-image characters (Danziger & Pederson, 1998) are quite poor at discriminating mirror images of geometric shapes (e.g., ↙ and ↘) or pictures of familiar objects (Fernandes & Kolinsky, 2013). Thus, reading in a script that includes lateral mirror images boosts the ability to discriminate these contrasts even with nonlinguistic materials.

Remarkably, the ability to discriminate mirror images interferes with other visual processes. For example, using an orientation-independent, identity-based, same-different comparison task in which participants had to respond “same” to both physically identical and mirror-image stimuli, Pegado et al. (2014) showed that both early and late literate adults (reading the Latin script) performed worse when written stimuli and pictures of familiar objects were mirrored rather than strictly identical, whereas illiterates showed no cost for these mirrored pairs. Thus, interference from irrelevant mirror-image variations on identity processing is a side effect of literacy for readers of scripts including such contrasts.

These influences of literacy occur at a relatively high processing level. Indeed, illiterate adults do register mirror-image contrasts at an earlier processing level: They display the same level of illusory conjunctions (false detections of a target in very briefly presented displays; see Figure 25.2) as early literates in a situation in which, to perceive the target, the lateral mirror orientation of diagonal lines has to be registered preattentively (Kolinsky, Morais, & Verhaeghe,

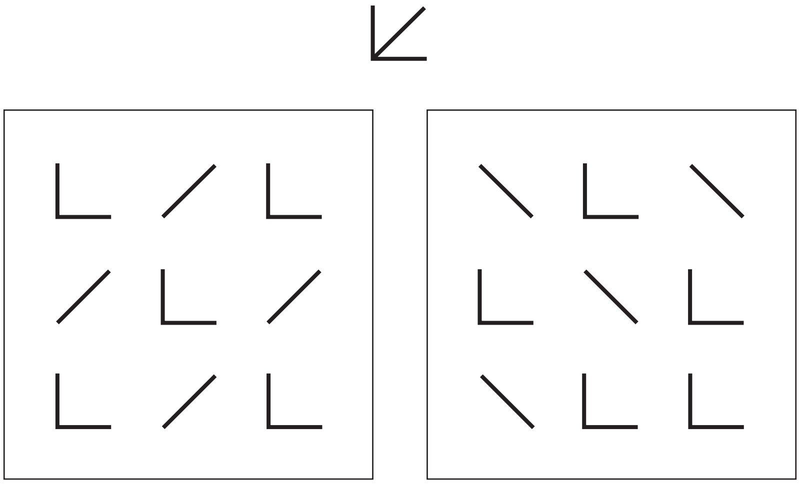


Fig. 25.2 Examples of target-absent trials used by Kolinsky, Morais, and Verhaeghe (1994) to elicit illusory conjunctions (i.e., false detections) of the target, presented on the top. On the left: experimental trial, in which the orientation of diagonal lines matches the orientation of the diagonal line of the target. On the right: control trial, in which orientation of diagonal lines is opposite.

1994). Nonetheless, other data demonstrate that literacy also alters some early vision processes, as discussed in the next section.

Reading Acquisition Alters Early Visual Processing

In the fMRI study by Dehaene, Pegado, et al. (2010), reading acquisition was shown to increase occipital responsiveness to all visual (linguistic and nonlinguistic) categories in a situation in which participants only had to perform an incidental task (detecting an occasional target star).[2] Even activation in the primary visual area V1, which is the first point of entry of visual signals into the cortex, was augmented by literacy: relative to illiterates, trained readers showed enhanced fMRI responses in V1 to written sentences and checkerboards. The latter effect was selective for horizontal over vertical checkerboards; as the spatial arrangement of the image is maintained in V1 (stimuli adjacent in the visual field are represented in adjacent positions in the visual cortex), intensive training with horizontally presented words had clearly led to a refinement of the corresponding region of the visual field.

More generally, the Latin alphabet provides an ideal stimulus for perceptual learning through extensive practice on a restricted set of visually simple shape primitives. Discriminating these shapes puts a challenge on visual resolution, which may explain the importance of visual areas sensitive to oriented bars (V1) and local contours (secondary visual cortex, V2). In agreement with this idea, areas V1/V2 of expert alphabetic readers show increased fMRI activation specific to words relative to matched scrambled controls (Szwed et al., 2011; Szwed, Qiao, Jobert, Dehaene, & Cohen, 2014). Interestingly, this effect is not observed in Chinese readers (Szwed et al., 2014), probably because Chinese characters are much more numerous and visually complex than Latin letters. Discriminating between thousands of characters that each comprises a hierarchical arrangement of many strokes seems rather to put emphasis on a higher level of visual processing, as Chinese expert readers show enhanced activations in intermediate visual areas (V3/V4, sensitive to more complex patterns than V1/V2) that are absent in alphabetic readers.

These early effects may benefit several visual tasks outside of reading. For instance, in readers of the Latin script visual integration is enhanced, as shown by early and late literates' superior capacity (compared to illiterates) in connecting local elements into an overall shape (Szwed, Ventura, Querido, Cohen, & Dehaene, 2012). The early visual changes induced by the acquisition of the Latin alphabet might also form the foundations of the more analytical strategy of attending to pictures observed in (alphabetic) literates compared with illiterates (Ventura et al., 2013). Better fine visual discrimination skills, including of nonlinguistic stimuli, have also been reported in beginning readers of modern Standard Chinese, which has been visually simplified in the 1950s and is used in mainland China, compared with beginning readers of the traditional, visually more complex Chinese script, still used for example in Hong Kong and Taiwan (McBride-Chang, Chow, Zhong, Burgess, & Hayward, 2005). This is a counterintuitive result, as one might have expected the more visually complex script to engender better fine visual discrimination. It remains to be investigated whether the greater emphasis on intermediate visual areas put by the more complex scripts (Szwed et al., 2014) induces other behavioral changes.

The Influence of Reading Acquisition on Visual Processing: A Brief Summary

In sum, reading acquisition gives individuals qualitatively new processing modes tuned to the processing of written strings; modifies scanning habits and spatial associations with numbers, time, and action events; and reorganizes the visual ventral pathway through a process of neural competition with other visual categories, principally with faces. It also alters early visual processes, enhances fine visual discrimination, and pushes readers of scripts that include mirror images to unlearn mirror invariance. The following section illustrates that in addition, literacy affects some aspects of higher-level functions.

The Effects of Reading Acquisition on Higher-Level Functions

Semantic Knowledge and Organization

Both learning to read in the classroom and activities linked to literacy (reading books, magazines, etc.) certainly increase the richness and precision of semantic knowledge. This is observed, for instance, in semantic fluency tasks in which participants are asked to generate as many words as they can that belong to a specified taxonomic category (e.g., animals). As a matter of fact, illiterate adults provide far fewer responses than early literates (e.g., Ratcliff et al., 1998), and a similar difference is observed between age-matched illiterate and literate children (Matute et al., 2012). However, this finding does not imply that literacy changes the way entities are represented in conceptual memory, including their taxonomic organization, or the mechanisms of access to stored knowledge. In semantic fluency tasks, when participants have to generate a list of words corresponding to a given taxonomic category such as animals, they tend to produce clusters of words belonging to the same subcategory (e.g., pets, insects, birds), which reflects both organization and retrieval by subcategory (e.g., Gruenewald & Lockhead, 1980). Even illiterates display such a pattern (e.g., Kosmidis, Tsapkini, Folia, Vlahou, & Kiosseoglou, 2004). Thus, contrary to the richness and precision of knowledge, taxonomic clustering and retrieval by semantic subcategory does not strongly depend on literacy. This outcome is consistent with the idea that, although unschooled illiterate people show a preference for thematic relations in categorization tasks (grouping for instance *leg* with *trousers* rather than with *arm,* e.g., Luria, 1976), they do use taxonomic organization of the items when the categories are explicitly indicated to them or simply suggested by having them sort the items into piles (e.g., Scribner & Cole, 1981).

Working Memory and Executive Functions

The use of external symbolic storage systems (books, computers, etc.) induces the need to manage multiple memory stores (both internal and external) and multiple knowledge codes (phonemic, orthographic, metalinguistic), which may modify executive functions, in particular working memory (e.g., Donald, 1993). Examining this idea is difficult, as executive functions form a set of related but clearly distinct functions (e.g., Miyake et al., 2000) and there are virtually no data on the effects of literacy on shifting between multiple tasks or criteria, deliberate inhibiting of dominant responses, and planning and organizing output sequences (but see preliminary results reviewed by Morais & Kolinsky, 2002, suggesting an effect of formal education rather than of literacy per se).

There is, however, some evidence for an effect of literacy on working memory (WM) tasks, namely on tasks that, by adding a processing demand to the requirement to remember a list of items, involve manipulation of information in addition to simple storage. Appropriately revising the items held in memory to keep track of which information is old and no longer relevant, and replacing it by newer, more relevant information is an ability closely related to the executive functions of selecting, updating, and monitoring representations (e.g., Miyake et al., 2000; for a review, see Bledowski, Kaiser, & Rahm, 2010). According to the results reported by Kosmidis et al. (2011), literacy strengthens working memory, but this skill may be further reinforced through formal education, presumably as individuals develop learning strategies. Indeed, in backward digit span (in which participants have to recall the list of items in reverse order), late literates perform similarly to illiterates, both less well than early literates. Yet a specific effect of literacy is observed on listening span (in which participants have to listen to a series of sentences, retaining the final word of each sentence for recall at the end of the series), with poorer performance in illiterates than late literates and no significant difference between late and early literates. Nevertheless, as for STM literacy effects on WM tasks may be restricted to or stronger with verbal than nonverbal materials. The reported effects of literacy on spatial WM tasks are in fact confounded with the effects of formal education, with early literates better on spatial span backward than functional illiterates who attended school for only a very short time (Kosmidis et al., 2011).

Reasoning Capacities, IQ, and Cognitive Style: Effects of Formal Education or of Literacy?

Given their context-independency and permanence, written materials and hence literacy are often considered as fostering formal thought and abstraction (e.g., Donald, 1993; Harris, 2009; Ong, 1982). Consistently, both Goody (1968) and Luria (1976) considered literacy to be a precondition for deductive reasoning, as applied in syllogisms (namely, the capacity to deduce, e.g., that Socrates is mortal from the premises *All men are mortal; Socrates is a man*), and Luria reported that illiterate adults perform poorly on reasoning tasks. In fact, illiterate adults' reasoning ability is generally masked by an empirical bias (Scribner, 1977): When presented with unfamiliar premises, they use their own experience to supplement, distort, or even reject them (e.g., Cole et al., 1971; Luria, 1976; Scribner & Cole, 1981). For example, when given the problem: *In the far North, where there is snow, all bears are white. Novaya Zemlya is in the far North. What color are the bears there?*, an illiterate participant answered: *I don't know. I've seen a black bear. I've never seen any others... Each locality has its own animals* (Luria, 1976, pp. 108–109). Yet unschooled people are quite good with syllogisms based on familiar information (Scribner & Cole, 1981) and, with unfamiliar information, illiterate adults reason accurately and appropriately justify their conclusions in terms of the supplied premises when explicitly prompted to think of these as pertaining for example to a distant planet, which allows them to set empirical considerations aside (Dias, Roazzi, & Harris, 2005). Nevertheless, the illiterate participants of Dias et al. (2005) performed less well overall than the early literates. Observations made by Scribner and Cole (1981) on the Vai people of West Africa suggest that formal education in Western-type schools (long-term tuition delivered by trained teachers and including various activities beyond literacy, such as mathematics and history) is responsible for this effect: Performance with logic problems demonstrated strong effects of this type of schooling, but neither Vai (syllabic) literacy, acquired at home through individual tuition, nor Arabic (consonantal alphabetic) literacy, acquired in Koranic schools (where tuition was restricted to reading and writing out known passages of the Koran or frequently used prayers), was found to improve performance.

A clear case of formal education influence concerns performance on tests that are designed to measure intelligence. Although IQ scores usually correlate with literacy, there is either no difference—or only a tiny one—between illiterates and late literates, both displaying far poorer scores than early literates (Verhaeghe & Kolinsky, 2006; see also the longitudinal study by Landgraf et al., 2011, on almost unschooled adults involved in a literacy course). Formal education, but not literacy, also seems responsible for differences in so-called cognitive styles. The influence of prior beliefs in reasoning and consequent cross-cultural variations led to the idea that people from different cultures use different cognitive processes when they reason. For example, Nisbett (2003) described Eastern reasoning as holistic and dialectical and Western reasoning as analytical and logical. Similarly, Easterners are said to engage in context-dependent holistic visual processes by attending to the relationship between the object and the context in which the object is located, whereas Westerners are said to engage in context-independent analytic processes by focusing on a salient object independently from the context in which it is embedded (e.g., Nisbett & Miyamoto, 2005). Ventura et al. (2008) showed that Western schooling, as part of or in addition to culture, is a crucial factor in this effect, but that literacy per se is irrelevant: Only Portuguese early literates showed a context-independent analytic processing style, whereas all other groups (Portuguese illiterates and late literates, as well as Thai illiterates and early and late literates) showed a context-dependent holistic style.

Conclusions

Much behavioral and brain-imaging evidence has now accumulated to support Frith's (1998) assertion that literacy is changing the brain, including its basic anatomy. As regards speech, apart from a few unaffected domains (phonetic discrimination, categorical perception, phonological restructuring of lexical representations, implicit phonemic codes), reading does change the way spoken language is processed. Literates do not process speech as illiterates do, and they are more deeply influenced by spelling knowledge than was initially thought. The same conclusion holds true for visual perception: Reading acquisition allows the emergence of processes and brain structures tuned to written strings, alters the way other visual categories are processed, and induces neural competition effects. In addition, literacy modifies the anatomy of the brain, including the connections between the visual (orthographic) and phonological processing areas.

Evidence for the influence of literacy on higher-level functions is far less clear. Despite interesting discussions about the extent to which new cultural tools such as reading retool our minds (e.g., Ansari, 2012; Donald, 1993; Wilson, 2010), data on the influence of literacy on executive functions and reasoning are inconclusive. Much work has still to be done to understand what may be considered the new agenda of cognitive science, namely "to understand the shared principles by which individual brains develop into diverse adult minds" (Wilson, 2010, p. 186).

Future Directions

Beyond the previously noted lack of evidence on high-level functions, the mechanisms by which reading changes brain function and structure often remain opaque. On the one hand, we need detailed models of the ways in which cultural tools affect brain function (see proposals on neural reuse in Anderson, 2010; Dehaene & Cohen, 2007). On the other hand, more data are necessary to identify the behavioral correlates of the observed brain changes (e.g., of the neural competition between written word and face processing, Dehaene, Pegado, et al., 2010) and to identify the exact brain correlates of reported behavioral effects.

Important questions also arise concerning the effects of literacy across the life span. Can adults learn to read as efficiently as children? Or is it the case, as advocated by Abadzi (2012), that adults have more difficulties than children in acquiring a new script, displaying so called neoliterate dyslexia? A related question is whether there are sensitive periods for reading-dependent effects on brain and cognition, including for neural competition. Answers to these questions have important implications for the timing and content of educational interventions, but they depend on detailed examination of the populations' characteristics. For example, the fact that neural competition between words and faces is observed only in early but not in late literates (Dehaene, Pegado, et al., 2010) may reflect either the rudimentary reading level of the latter or limited plasticity in adulthood. Likewise, we do not yet know whether late literates' rudimentary reading reflects adults' limitations, differences in number of learning years, or differences in motivation linked to personal goals.

More generally, as we begin to understand which processes and brain networks are changed by literacy, we may start thinking about how to optimize reading acquisition, particularly for children who struggle in this process despite having normal access to reading education as well as adequate intelligence and intact sensory abilities, namely developmental dyslexics. Longitudinal studies on either age-matched children (e.g., Monzalvo & Dehaene-Lambertz, 2013) or unschooled adults involved in literacy classes (e.g., Landgraf et al., 2011) as well as training studies (e.g., Brem et al., 2010) monitoring both participants' behavioral progress and brain activation changes offer promising avenues. Indeed, studying the impact of literacy should lead to better understanding of the pathogenesis (or proximal causes; see Pennington & Peterson, this volume) of developmental dyslexia. For example, dyslexics show reduced neural integration of letters and phonemes in the planum temporale as well as reduced activation in the same brain area with purely aural presentation of phonemes (Blau et al., 2010; Monzalvo et al., 2012). This has been interpreted as a proximal cause of reading failure (Blau et al., 2010). Yet because similar reduced activation in response to speech is observed in illiterate adults (Dehaene, Pegado et al., 2010) and preliterate children (Monzalvo & Dehaene-Lambertz, 2013) compared with literates, reading level might be the real cause. In the future, we should thus integrate better what we learn from studies on missing literacy with what we know on failed literacy, both in terms of pathogenesis and new remediation programs. In this respect, comparative approaches that go beyond the exclusive examination of what have been called weird (western, educated, industrialized, rich, and democratic) members of humanity (Henrich, Heine, & Norenzayan, 2010) to include the study of illiterate and late literate adults become more and more urgent, as it is increasingly hard to find representative samples of these populations.

Acknowledgments

Preparation of this chapter was supported by the Fonds de la Recherche Scientifique-FNRS under grant FRFC 2.4515.12 and by an Interuniversity Attraction Poles (IAP) grant 7/33, Belspo.

Notes

1 As most of the data reviewed in this chapter concern alphabetic writing systems, the terms *literate* and *literacy* are used to refer to alphabetic literacy, unless otherwise specified. Unless otherwise specified, the term *illiterate* refers to adults who never learned to read and write any script.

2 The ventral visual pathway that is involved in the recognition of objects, including written strings, is organized as a hierarchy of areas. From posterior (occipital) to more anterior regions, the size of the neurons' receptive fields increases in parallel with increasing sensitivity to complex patterns (from line segments to feature combinations and whole objects) and decreasing sensitivity to physical changes (e.g., in size, location, or viewpoint).

References

Abadzi, H. (2012). Can adults become fluent readers in newly learned scripts? *Education Research International*, *2012*, 1–8.

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Adrián, J. A., Alegria, J., & Morais, J. (1995). Metaphonological abilities of Spanish illiterate adults. *International Journal of Psychology, 30,* 329–353.

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences, 33,* 245–313.

Ansari, D. (2012). Culture and education: New frontiers in brain plasticity. *Trends in Cognitive Sciences, 16,* 93–95.

Ardila, A., Bertolucci, P. H., Braga, L. W., Castro-Caldas, A., Judd, T., Kosmidis, M., . . . Rosselli, M. (2010). Illiteracy: The neuropsychology of cognition without reading. *Archives of Clinical Neuropsychology, 25,* 689–712.

Ardila, A., Ostrosky-Solis, F., & Mendoza, V. U. (2000). Learning to read is much more than learning to read: A neuropsychologically-based learning to read method. *Journal of the International Neuropsychological Society, 6,* 789–801.

Baddeley, A. D. (1966). Short-term memory for sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18,* 362–365.

Bergen, B. K., & Chan Lau, T. T. (2012). Writing direction affects how people map space onto time. *Frontiers in Psychology*, *3*.

Bertelson, P. (1972). Listening from left to right versus right to left. *Perception, 1,* 161–165.

Bertelson, P., & Tisseyre, F. (1972). Lateral asymmetry in the perceived sequence of speech and nonspeech stimuli. *Perception & Psychophysics, 11,* 356–362.

Berthoud-Papandropoulou, I. (1978). An experimental study of children's ideas about language. In A. Sinclair, R. J. Jarvella, & W. J. M. Levelt (Eds.), *The child's conception of language* (pp. 55–65). Berlin, Germany: Springer.

Blau, V., Reithler, J., van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., & Blomert, L. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: A functional magnetic resonance imaging study of dyslexic children. *Brain, 133,* 868–879.

Bledowski, C., Kaiser, J., & Rahm, B. (2010). Basic operations in working memory: Contributions from functional imaging studies. *Behavioural Brain Research, 214,* 172–179.

Bramão, I., Mendonça, A., Faísca, L., Ingvar, M., Petersson, K. M., & Reis, A. (2007). The impact of reading and writing skills on a visuo-motor integration task: A comparison between illiterate and literate subjects. *Journal of the International Neuropsychological Society, 13,* 359–364.

Brem, S., Bach, S., Kucian, K., Guttorm, T. K., Martin, E., Lyytinen, H., . . . Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *Proceedings of the National Academy of Sciences, 107,* 7939–7944.

Brennan, C., Cao, F., Pedroarena-Leal, N., McNorgan, C., & Booth, J. R. (2013). Reading acquisition reorganizes the phonological awareness network only in alphabetic writing systems. *Human Brain Mapping, 34,* 3354–3368.

Burnham, D. K. (2003). Language specific speech perception and the onset of reading. *Reading and Writing: An Interdisciplinary Journal, 16,* 573–609.

Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience, 12*, 679–690.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McGuire, P. K., . . . David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science, 276,* 593–596.

Carreiras, M., Seghier, M., Baquero, S., Estévez, A., Lozano, A., Devlin, J. T., & Price, C. J. (2009). An anatomical signature for literacy. *Nature, 461,* 983–986.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition, 91,* 77–111.

Castles, A., Holmes, V. M., Neath, J., & Kinoshita, S. (2003). How does orthographic knowledge influence performance on phonological awareness tasks? *Quarterly Journal of Experimental Psychology, 56A,* 445–467.

Castro-Caldas A., Cavaleiro Miranda, P., Carmo, I., Reis, A., Leote, F., Ribeiro, C., & Ducla-Soares, E. (1999). Influence of learning to read and write on the morphology of the corpus callosum. *European Journal of Neurology, 6,* 23–28.

Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., & Ingvar, M. (1998). The illiterate brain: Learning to read and write during childhood influences the functional organization of the adult brain. *Brain, 121,* 1053–1063.

Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2010). Categorical speech representation in human superior temporal gyrus. *Nature Neuroscience, 13,* 1428–1432.

Chokron, S., Kazandjian, S., & De Agostini, M. (2009). Effect of reading direction on visuospatial organization: A critical review. In A. Gari, & K. Mylonas (Eds.), *Quod erat demonstrandum: From Herodotus' ethnographic journeys to cross-cultural research* (pp. 104–114). Athens, Greece: Pedio Books.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123,* 291–307.

Cole, M., Gay, J., Glick, J. A., & Sharp, D. W. (1971). *The cultural context of learning and thinking: An exploration in experimental anthropology*. New York, NY: Basic Books.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55,* 429–432.

Cutler, A., Treiman, R., & Van Ooijen, B. (2010). Strategic deployment of orthographic knowledge in phoneme detection. *Language and Speech, 53,* 307–320.

Damian, M. F., & Bowers, J. S. (2010). Orthographic effects in rhyme monitoring tasks: Are they automatic? *European Journal of Cognitive Psychology, 22,* 106–116.

Danziger, E., & Pederson, E. (1998). Through the looking-glass: Literacy, writing systems and mirror-image discrimination. *Written Language and Literacy, 1,* 153–164.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122,* 371–396.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56,* 384–98.

Dehaene, S., Nakamura, K., Jobert, A., Kuroki, C., Ogawa, S., & Cohen, L. (2010). Why do children make mirror errors in reading? Neural correlates of mirror invariance in the visual word form area. *Neuroimage, 49,* 1837–1848.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes, G., Jobert, A.,... Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science, 330,* 1359–1364.

Denny, J. P. (1991). Rational thought in oral culture and literate decontextualization. In D. R. Olson, & N. Torrance (Eds.), *Literacy and orality* (pp. 66–89). New York, NY: Cambridge University Press.

Dias, M., Roazzi, A., & Harris, P. L. (2005). Reasoning from unfamiliar premises: A study with unschooled adults. *Psychological Science, 16,* 550–554.

Dobel, C., Enriquez-Geppert, S., Zwitserlood, P., & Bölte, J. (2014). Literacy shapes thought: The case of event representation in different cultures. *Frontiers in Psychology*, *5,* 290.

Donald, M. W. (1993). Précis of *Origins of the modern mind: Three stages in the evolution of culture and cognition*. *Behavioral and Brain Sciences, 16,* 737–791.

Duñabeitia, J. A., Dimitropoulou, M., Grainger, J., Hernández, J. A., & Carreiras, M. (2012). Differential sensitivity of letters, numbers, and symbols to character transposition. *Journal of Cognitive Neuroscience, 24,* 1610–1624.

Duñabeitia, J. A., Orihuela, K., & Carreiras, M. (2014). Orthographic coding in illiterates and literates. *Psychological Science, 25,* 1275–1280.

Dundas, E. M., Plaut, D. C., & Behrmann, M. (2013). The joint development of hemispheric lateralization for words and faces. *Journal of Experimental Psychology: General, 142,* 348–358.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36,* 250–287.

Fernandes, T., & Kolinsky, R. (2013). From hand to eye: The role of literacy, familiarity, graspability, and vision-for-action on enantiomorphy. *Acta Psychologica, 142,* 51–61.

Fernandes, T., Vale, A. P., Martins, B., Morais, J., & Kolinsky, R. (2014). The deficit of letter processing in developmental dyslexia: Combining evidence from dyslexics, typical readers, and illiterate adults. *Developmental Science, 17,* 125–141.

Fodor, J., & Bever, T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior, 4,* 414–420.

Frauenfelder, U. H., Seguí, J., & Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception and Performance, 16,* 77–91.

Frith, U. (1998). Editorial: Literally changing the brain. *Brain, 121,* 1011–1012.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language, 45,* 468–492.

Goody, J. (1968). *Literacy in traditional societies*. Cambridge, England: Cambridge University Press.

Goswami, U. (2000). Phonological representations, reading development and dyslexia: Towards a cross-linguistic theoretical framework. *Dyslexia, 6,* 133–151.

Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language, 35,* 623–647.

Grainger, J., Tidgat, I., & Isselé, J. (2010). Crowding affects letters and symbols differently. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 673–688.

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 225–240.

Harris, R. (2009). *Rationality and the literate mind*. New York, NY: Routledge.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33,* 61–135.

Hoonhorst, I., Medina, V., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2011). Categorical perception of voicing, colors and facial expressions: A developmental study. *Speech Communication, 53,* 417–430.

Hulme, C., Caravolas, M., Malkova, G., & Brigstocke, S. (2005). Phoneme isolation ability is not simply a consequence of letter-sound knowledge. *Cognition, 97,* B1–B11.

Kolinsky, R., Cary, L., & Morais, J. (1987). Awareness of words as phonological entities: The role of literacy. *Applied Psycholinguistics, 8,* 223–232.

Kolinsky, R., Morais, J., & Verhaeghe, A. (1994). Visual separability: A study on unschooled adults. *Perception, 23,* 471–486.

Kolinsky, R., Verhaeghe, A., Fernandes, T., Mengarda, E. J., Grimm-Cabral, L., & Morais, J. (2011). Enantiomorphy through the looking-glass: Literacy effects on mirror-image discrimination. *Journal of Experimental Psychology: General, 140,* 210–238.

Kosmidis, M. H., Tsapkini, K., Folia, V., Vlahou, C. H., & Kiosseoglou, G. (2004). Semantic and phonological processing in illiteracy. *Journal of the International Neuropsychological Society, 10,* 818–827.

Kosmidis, M. H., Zafiri, M., & Politimou, N. (2011). Literacy versus formal schooling: Influence on working memory. *Archives of Clinical Neuropsychology, 26,* 575–582.

Lafontaine, H., Chetail, F., Colin, C., Kolinsky, R., & Pattamadilok, C. (2012). Role and activation time course of the phonological and orthographic information during phoneme judgment. *Neuropsychologia, 50,* 2897–2906.

Landgraf, S., Beyer, R., Pannekamp, A., Schaadt, G., Koch, D. Foth, M., & Meer, E. (2011). Dissociating improvement of attention and intelligence during written language acquisition in adults. *International Journal of Intelligence Science, 1,* 17–24.

Li, S., Lee, K., Zhao, J., Yan, Z., He, S., & Weng, X. (2013). Neural competition as a developmental process: Early hemispheric specialization for word processing delays specialization for face processing. *Neuropsychologia, 51,* 950–959.

Liberman, A. M., Harris, K. S. Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54,* 358–368.

Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18,* 201–212.

Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations*. Cambridge, MA: Harvard University Press.

Martinez Perez, T., Majerus, S., & Poncelet, M. (2012). The contribution of short-term memory for serial order to early reading acquisition: Evidence from a longitudinal study. *Journal of Experimental Child Psychology, 111,* 708–723.

Matute, E., Montiel, T., Pinto, N., Rosselli, M., Ardila, A., & Zarabozo, D. (2012). Comparing cognitive performance in illiterate and literate children. *International Review of Education, 58,* 109–127.

McBride-Chang, C., Chow, B. W.-Y., Zhong, Y., Burgess, S., & Hayward, W.G. (2005). Chinese character acquisition and visual skills in two Chinese scripts. *Reading and Writing: An Interdisciplinary Journal, 18,* 99–128.

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science, 343,* 1006–1010.

Metsala, J. L. (1997). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Memory & Cognition, 25,* 47–56.

Miyake, A., Friedman, N. P., Emerson, N. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100.

Monzalvo, K., & Dehaene-Lambertz, G. (2013). How reading acquisition changes children's spoken language network. *Brain and Language, 127,* 356–365.

Monzalvo, K., Fluss, J., Billard, C. Dehaene, S., & Dehaene-Lambertz, G. (2012). Cortical networks for vision and language in dyslexic and normal children of variable socio-economic status. *NeuroImage, 61,* 258–274.

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition, 24,* 45–64.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7,* 323–331.

Morais, J., Castro, S.-L., Scliar-Cabral, L., Kolinsky, R., & Content, A. (1987). The effects of literacy on the recognition of dichotic words. *Quarterly Journal of Experimental Psychology, 39A,* 451–465.

Morais, J., & Kolinsky, R. (1994). Perception and awareness in phonological processing the case of the phoneme. *Cognition, 50,* 287–297.

Morais, J., & Kolinsky, R. (2002). Literacy effects on language and cognition. In L. Bäckman, & C. von Hofsten (Eds.) *Psychology at the turn of the millenium: Vol. 1. Cognitive, biological, and health perspectives* (pp. 507–530). Hove, England: Psychology Press.

Muneaux, M., & Ziegler, J. C. (2004). Locus of orthographic effects in spoken word recognition: Novel insights from the neighbour generation task. *Language and Cognitive Processes, 19,* 641–660.

Nation, K., & Hulme, C. (2011). Learning to read changes children's phonological skills: Evidence from a latent variable longitudinal study of reading and nonword repetition. *Developmental Science, 14,* 649–659.

Nisbett, R. E. (2003). *The geography of thought: How Asians and Westerners think differently…and why*. New York, NY: Free Press.

Nisbett, R. E., & Miyamoto, Y. (2005). The influence of culture: Holistic versus analytic perception. *Trends in Cognitive Sciences, 9,* 467–473.

Olson, D. (1991). Literacy as metalinguistic activity. In D. R. Olson, & N. Torrance (Eds.), *Literacy and orality* (pp. 251–270). Cambridge, England: Cambridge University Press.

Ong, W. J. (1982). *Orality and literacy: The technologizing of the word*. New York, NY: Methuen.

Pattamadilok, C., Knierim, I. N., Duncan, K. J., & Devlin, J. T. (2010). How does learning to read affect speech perception? *Journal of Neuroscience, 30,* 8435–8444.

Pattamadilok, C., Kolinsky, R., Luksaneeyanawin, S., & Morais, J. (2008). Orthographic congruency effects in the suprasegmental domain: Evidence from Thai. *Quarterly Journal of Experimental Psychology, 61,* 1515–1537.

Pattamadilok, C., Lafontaine, H., Morais, J., & Kolinsky, R. (2010). Auditory word serial recall benefits from orthographic dissimilarity. *Language and Speech, 53,* 321–341.

Pattamadilok, C., Morais, J., Ventura, P., & Kolinsky, R. (2007). The locus of the orthographic consistency effect in auditory word recognition: Further evidence from French. *Language and Cognitive Processes, 22,* 700–726.

Pattamadilok, C., Perre, L., Dufau, S., & Ziegler, J. C. (2009). On-line orthographic influences on spoken language in a semantic task. *Journal of Cognitive Neuroscience, 21,* 169–179.

Pattamadilok, C., Perre, L., & Ziegler, J. C. (2011). Beyond rhyme or reason: ERPs reveal task-specific activation of orthography on spoken language. *Brain and Language, 116,* 116–124.

Peereman, R., Dufour, S., & Burt, J. S. (2009). Orthographic influences in spoken word recognition: The consistency effect in semantic and gender categorization tasks. *Psychonomic Bulletin & Review, 16,* 363–368.

Pegado, F., Nakamura, K., Braga, L., Ventura, P., Nunes, G., Jobert, A.,…Dehaene, S. (2014). Literacy breaks mirror invariance for visual stimuli: A behavioral study with adult illiterates. *Journal of Experimental Psychology: General, 143,* 887–894.

Pegado, F., Nakamura, K., Cohen, L., & Dehaene, S. (2011). Breaking the symmetry: Mirror discrimination for single letters but not for pictures in the Visual Word Form Area. *NeuroImage, 55,* 742–749.

Perre, L., Pattamadilok, C., Montant, M., & Ziegler, J. C. (2009). Orthographic effects in spoken language: On-line activation or phonological restructuring? *Brain Research, 1275,* 73–80.

Perre, L., & Ziegler, J. C. (2008). On-line activation of orthography in spoken word recognition. *Brain Research, 1188,* 132–138.

Petersson, K. M., Silva, C., Castro-Caldas, A., Ingvar, M., & Reis, A. (2007). Literacy: A cultural influence on functional

left-right differences in the inferior parietal cortex. *European Journal of Neuroscience, 26,* 791–799.
Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language, 14,* 174–180.
Ranbom, L. J., & Connine, C. M. (2011). Silent letters are activated in spoken word recognition. *Language and Cognitive Processes, 26,* 236–261.
Ratcliff, G., Ganguli, M., Chandra, V., Sharma, S., Belle, S., Seaberg, E., & Pandav, R. (1998). Effects of literacy and education on measures of word fluency. *Brain and Language, 61,* 115–122.
Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic reading. *Cognition, 24,* 31–44.
Rossion, B., Dricot, L., Devolder, A., Bodart, J. M., Crommelinck, M., De Gelder, B., & Zoontjes, R. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience, 12,* 793–802.
Scliar-Cabral, L., Morais, J., Nepomuceno, L., & Kolinsky, R. (1997). The awareness of phonemes: So close—so far away. *International Journal of Psycholinguistics, 13,* 211–240.
Scribner, S. (1977). Modes of thinking and ways of speaking: Culture and logic reconsidered. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking: Readings in cognitive science* (pp. 483–500). New York, NY: Cambridge University Press.
Scribner, S., & Cole, M. (1981). *The psychology of literacy*. Cambridge, MA: Harvard University Press.
Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 546–554.
Serniclaes, W., Ventura, P., Morais, J., & Kolinsky, R. (2005). Categorical perception of speech sounds in illiterate adults. *Cognition, 98,* B35–B44.
Shaki, S., Fischer, M. H., & Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychonomic Bulletin & Review, 16,* 328–331.
Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language, 36,* 337–359.
Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabregue, R., Amadon, A., & Cohen, L. (2011). Specialization for written words over objects in the visual cortex. *Neuroimage, 56,* 330–344.
Szwed, M., Qiao, E., Jobert, A., Dehaene, S., & Cohen, L. (2014). Effects of literacy in early visual and occipitotemporal areas of Chinese and French readers. *Journal of Cognitive Neuroscience, 26,* 459–475.
Szwed, M., Ventura, P., Querido, L., Cohen, L., & Dehaene, S. (2012). Reading acquisition enhances an early visual process of contour integration. *Developmental Science, 15,* 139–149.
Taft, M. (2006). Orthographically influenced abstract phonological representation: Evidence from non-rhotic speakers. *Journal of Psycholinguistic Research, 35,* 67–78.
Taft, M. (2011). Orthographic influences when processing spoken pseudowords: Theoretical implications. *Frontiers in Psychology, 2*, 140.
Thiebaut de Schotten, M., Cohen, L., Amemiya, E., Braga, L. W., & Dehaene, S. (2014). Learning to read improves the structure of the arcuate fasciculus. *Cerebral Cortex, 24,* 989–995.
Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition, 15,* 49–74.
Treiman, R., & Cassar, M. (1997). Can children and adults focus on sound as opposed to spelling in a phoneme counting task? *Developmental Psychology, 33,* 771–780.
Van Leeuwen, C., & Lachmann, T. (2004). Negative and positive congruence effects in letters and shapes. *Perception & Psychophysics, 66,* 908–925.
Ventura, P., Fernandes, T., Cohen, L., Morais, J., Kolinsky, R., & Dehaene, S. (2013). Literacy acquisition reduces the automatic holistic processing of faces and houses. *Neuroscience Letters, 554,* 105–109.
Ventura, P., Kolinsky, R., Brito-Mendes, C., & Morais, J. (2001). Mental representations of the syllable internal structure are influenced by orthography. *Language and Cognitive Processes, 16,* 393–418.
Ventura, P., Kolinsky, R., Fernandes, S., Querido, L., & Morais, J. (2007). Lexical restructuring in the absence of literacy. *Cognition, 105,* 334–361.
Ventura, P., Morais, J., Pattamadilok, C., & Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes, 19,* 57–95.
Ventura, P., Pattamadilok, C., Fernandes, T., Kolinsky, R., Klein, O., & Morais, J. (2008). Schooling in Western culture promotes context-free processing. *Journal of Experimental Child Psychology, 100,* 79–88
Verhaeghe, A., & Kolinsky, R. (2006). *O que os iletrados nos ensinam sobre os testes de inteligência* [What illiterate people teach us about intelligence tests]. Lisbon, Portugal: Fundação Calouste Gulbenkian- Fundação para a Ciência e a Tecnologia.
Wilson, M. (2010). The re-tooled mind: How culture re-engineers cognition. *Social Cognitive and Affective Neuroscience, 5,* 180–187.
Yoncheva, Y. N., Maurer, U., Zevin, J. D., & McCandliss, B. D (2013). Effects of rhyme and spelling patterns on auditory word ERPs depend on selective attention to phonology. *Brain and Language, 124,* 238–243.
Yoncheva, Y. N., Zevin, J. D., Maurer, U., & McCandliss, B. D. (2010). Auditory selective attention to speech modulates activity in the visual word form area. *Cerebral Cortex, 20,* 622–632.
Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16,* 747–759.
Zatorre, R., Meyer, E., Gjedde, A., & Evans, A. (1996). PET studies of phonetic processing of speech: Review, replication and reanalysis. *Cerebral Cortex, 6,* 21–30.
Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory recognition. *Psychonomic Bulletin & Review, 5,* 683–689.
Ziegler, J. C., Petrova, A., & Ferrand, L. (2008). Feedback consistency effects in visual and auditory word recognition: Where do we stand after more than a decade? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 643–661.

PART 5

Reading Instruction

CHAPTER

26

Young Children's Home Literacy Experiences

Monique Sénéchal

Abstract

The present chapter focuses on parent–child interactions that foster early literacy and oral language. The first part of the chapter presents an empirically based model of the differential links between two types of home literacy activities and child outcomes. The second part of the chapter is a synthesis of quasi-experimental research testing the impact of parent–child interactions on child early literacy. The converging correlational and quasi-experimental evidence presented is in accord with the proposed Home Literacy Model (Sénéchal, 2006; Sénéchal & LeFevre, 2002). In this model, informal literacy activities such as shared reading are rich language experiences from which children can learn about oral language. Most important, formal literacy experiences, such as parent teaching, seem to be necessary if gains in early literacy are expected from the home environment.

Key Words: early literacy, parent–child interaction, home literacy environment, reading, children, teaching

Young children can learn about literacy through their own attempts at reading and writing, through observations of role models, and through interactions with literate others (Teale & Sulzby, 1986). Parents provide resources such as books and coloring material that can facilitate children's own explorations; parents model literate acts when they read for pleasure and write notes, letters, or grocery lists; and parents provide literacy interactions when they point out letters in the text when reading picture books to their children. Understanding the impact of home literacy experiences on young children's reading success is important for two reasons. First, individual differences in children's reading skills are established early and remain fairly stable over time (e.g., Butler, Marsh, Sheppard, & Sheppard, 1985). For instance, children who have difficulty learning to read (and learning arithmetic) in grade one are more likely to have more difficulty in other school domains later on and are more likely not to complete high school or not to pursue their education beyond high school (Alexander, Entwisle, & Horsey, 1997; Entwisle, Alexander, & Olson, 2005). Second, understanding better how to optimize reading skills early in a child's life is important because the preschool and early school periods seem particularly sensitive to environmental influences (Hart & Petrill, 2009; Landry, Smith, Swank, & Guttentag, 2008).

This chapter consists of two sections on parent–child interactions that foster early literacy and language. The first part of the chapter presents an empirically based model of the links between two types of home literacy activities and child outcomes. This model is evaluated in light of recent research findings. The second part of the chapter is a synthesis of experimental and quasi-experimental research testing the impact of parent–child interactions on young children's early literacy. In both sections, the children are 4 and 5 years of age; that is, they have not yet begun their reading instruction at school. For ease of description only, the terms "preschool" and "kindergarten" are used to refer to the youngest and oldest children, respectively. As will be seen, the frequency and type of home literacy activities have been documented in North and South America, Europe, Australia, and Asia. However, most findings and descriptions presented pertain to the English language, unless otherwise specified.

A Home Literacy Model

Literacy activities are found in the homes of many young children, and this exposure to literacy may take several forms. This chapter examines parent–child interactions that could foster language and literacy. Using correlational findings from their longitudinal research, Sénéchal and LeFevre (2002) and Sénéchal (2006) elaborated a model of the association between home literacy and child outcomes. According to their home literacy model, early home literacy experiences should be considered as a function of how much focus there is on the print itself. Sénéchal and LeFevre proposed two categories of experiences; namely, informal and formal literacy activities. Informal literacy activities are those for which the meaning carried by the print is the focus of the interaction. The best example of informal literacy is when parents read books to their child, focusing on the meaning carried by the printed text. In contrast, formal literacy experiences are those where the focus of the interaction is on the features of print; that is, on letters, their use, their combinations as well as attempts to read and print words. The term "formal" is used to refer to interactions about the form of print, and it does not presuppose that the parent–child interactions are structured. In fact, these interactions can be playful, informative, or didactic. During shared book reading, formal literacy experiences occur when the parent points out letters in the text or shows how to read specific words in the text. Hence, the difference between informal and formal literacy activities is whether the focus of the interaction is the meaning carried by the print or the form of the print itself.

According to the home literacy model, the two types of home literacy experiences are differentially related to oral language and early literacy. Specifically, informal literacy activities promote the development of oral language skills, whereas formal literacy activities promote the acquisition of early literacy skills such as alphabet knowledge and initial attempts to print and read words (see Ehri, this volume). Sénéchal and LeFevre (2002) argued that it was important to consider children's phoneme awareness—a key predictor of reading—as distinct from oral language and early literacy. Doing so is necessary to test whether any association between home literacy and phoneme awareness is mediated by children's oral- and written-language skills. Each of these predictions is examined in the remainder of this section.

In the home literacy model, the frequency and variety of parent–child shared reading have been used to index informal literacy experiences. During shared reading, parent and child can enjoy the language and content of children's books as well as the accompanying illustrations. When asked to rate the importance of a variety of reasons to read books to their 4- and 5-year-old children, parents endorsed most strongly statements that they read to their children for enjoyment and to share quality time with their child (Audet, Evans, Mitchell, & Reynolds, 2008). During shared reading, children can also learn. Sénéchal, LeFevre, Hudson, and Lawson (1996) described three characteristics of shared book reading that can foster learning about the world and about language. First, the language used in books is more complex than that typically used during conversation (Hayes & Ahrens, 1988). Also, the language used by mothers is more complex during shared reading than during free play or remembering events (Crain-Thoreson, Dhalin, & Powell, 2001). As such, children may be exposed to new syntactic, grammatical, and lexical forms during shared reading episodes. The second feature of shared reading is that a child has the undivided attention of an adult who can define, explain, and question to facilitate the child's understanding or reinforce new knowledge. Certainly, the abundant literature on dialogic reading, where the adult engages the child in verbal exchanges, has shown the value of shared reading for expressive vocabulary acquisition (for a meta-analysis, see Mol, Bus, de Jong, & Smeets, 2008). A third characteristic of shared reading is that books can be read often, thus providing repeated exposure to new knowledge. Some of our work on shared reading has shown the advantage of repeated exposure for comprehending new words (Sénéchal, 1997). Because of these features, shared book reading is the single most studied aspect of children's home literacy environment. In this chapter, shared reading refers exclusively to parents reading books to their children.

For many researchers and educators, shared reading should also present frequent occasions for parents and children to discuss the printed text. Observations of parent–child interactions, however, show that parents seldom comment on print during these interactions. For example, in a study by Hindman, Connor, Jewkes, and Morrison (2008), the majority of remarks (85%) made by American parents during shared reading with their 4-year-olds were meaning-related (e.g., labeling, summarizing, discussing new words), whereas only 15% were code-related (e.g., teaching names or sounds of letters, decoding words). In addition, Hindman, Skibbe, and Foster (2013) showed that only 1% of mothers pointed out letters or sounds during shared reading, whereas 85% labeled illustrations. Similar findings were reported in other observational research (Audet et al., 2008; Deckner, Adamson, & Bakeman, 2006; Stadler & McEvoy, 2003). Observational research has also shown that 4- and 5-year-olds tend to look at the illustrations, not the written words, during shared reading (Evans & Saint-Aubin, 2005), unless their attention is drawn to the print (Justice, Pullen, & Pence, 2008). If shared reading is not used frequently to stimulate early literacy and if children do not look at print readily, then one wonders during which activities parents stimulate their child's early literacy knowledge (Sénéchal & LeFevre, 2002). A recent study by Martini and Sénéchal might resolve this apparent paradox.

Martini and Sénéchal (2012) documented the activities and contexts that parents used to help their child learn about literacy. Examples of learning contexts include using familiar household items, street signs, games, letters from school, mail, and newspapers, as well as children's books. They found that parents generally reported using a wide variety of learning contexts: Of the eighteen contexts presented, parents selected on average fourteen different contexts that they used at least some of the time. Moreover, Martini and Sénéchal found that parents who reported teaching about literacy more frequently tended to use a greater number of learning contexts. Martini and Sénéchal concluded that parents focus on naturally occurring activities to impart knowledge about the alphabet, printing, and reading words. The reported frequency of teaching along with the numerous contexts used might also suggest that these teaching moments are not very long in duration. Importantly, frequent teaching moments that are varied and short in duration might indicate parents' sensitivity to the attention span and interest of their young child as well as the difficulty of the task. Learning the alphabet, for example, requires learning to discriminate the different forms of letters, learning that letters are symbols that represent individual speech sounds, and learning that letters have names and sounds that may or may not be the same.

Although the Martini and Sénéchal (2012) findings are based on questionnaires, additional support for the view that parents stimulate literacy knowledge in a variety of contexts comes from analyses of everyday parent–child conversations. These analyses demonstrated that parents sometimes talked to their children about letters, asking questions about letter shapes and letter–word associations (Robins & Treiman, 2009; Robins, Treiman, & Rosales, 2014; Robins, Treiman, Rosales, & Otake, 2012).

Given the variability of activities and contexts that parents seem to use for teaching about literacy, researchers have relied on more general questions to survey parents on the frequency of formal literacy experiences at home. The questions pertain to parents' attempts to impart literacy knowledge to their children. Researchers have used many different terms—*parent coaching* (Edwards, 1991), *parent mediation* (Aram & Levin, 2004), or *parent scaffolding* (Evans, Moretti, Shaw, & Fox, 2003)—to indicate the didactic support parents can provide for learning about literacy. In this chapter, the terms "parent teaching" and "tutoring" are used interchangeably. These terms are preferred because they tend to be more easily understood. However, these terms do not presuppose that parents use structured activities but rather that they use naturally occurring occasions to stimulate literacy knowledge. Again, these interactions focusing on the form and use of print can be playful (e.g., asking the child to find all ‹O›s on a page as quickly as possible), informative (e.g., pointing out a word that has two ‹O›s in it), or didactic (e.g., encouraging the child to print the letter ‹O› by showing that it is like making a circle).

Parent teaching seems to have been neglected by researchers, perhaps because of a view of early childhood that favors child explorations about literacy as well as shared reading of children's books. To illustrate, a review of seventy articles on home literacy published from 1970 to 2013 revealed that only 34% (twenty-four) asked parents about their teaching practices at home in addition to shared reading activities. Even when teaching questions were included, these questions were not analyzed separately from the shared reading questions in the majority (54%) of the studies. Omnibus measures of home literacy or failure to consider parent teaching, however, might lead to false conclusions about the relations between specific activities and child early literacy. For instance, one could think that shared reading predicts early literacy when it is teaching activities that are robustly linked to early literacy.

Parent teaching about literacy seems more prevalent than once thought. Table 26.1 presents

Table 26.1 Mean Frequency of Teaching Literacy Reported by Parents as a Function of Home Language (12 Studies).

		Child Characteristics				Study Characteristics					
Author(s)	Year	Grade	SES	Country	Lang.	N	Parental Teaching	*M*	Range	Scale Anchor Labels	Separate Analysis[d]
Boudreau	2005	PS	Middle	US	E	37	Letters	3.3	1–5	1: Infrequently—5: Frequently	No
Foy and Mann	2003	PS	Middle	US	E	40	Reading	3.4	1–5	1: Never—5: Very Often	Yes
							Writing	3.5	1–5	1: Never—5: Very Often	No
Hood et al.	2008	PS	Low/ Middle	AUS	E	143	Letters	4.3[a]	2–5	1: Never—5: Very Often	Yes
							Write name	4.3[a]	2–5	1: Never—5: Very Often	No
							Reading	3.8[a]	1–5	1: Never—5: Very Often	No
Martini and Sénéchal	2012	K	High	CA	E	108	Letter names	4.4	2–5	1: Never—5: Very Often	Yes
							Letter sounds	3.7	1–5	1: Never—5: Very Often	Yes
							Letter printing	3.7	1–5	1: Never—5: Very Often	Yes
							Name printing	4.4	2–5	1: Never—5: Very Often	Yes
							Word printing	4.2	1–5	1: Never—5: Very Often	Yes
							Word reading	3.5	1–5	1: Never—5: Very Often	Yes
Phillips and Lonigan	2009	PS	Mixed	US	E	1044	Letters	4.3	0–6	0: Never—7: Daily	No
Sénéchal and LeFevre	2014	K	High	CA	E	110	Reading	3.5		1: Never—5: Very Often	Yes
Sénéchal et al.	1998	K	Middle/ High	CA	E	110	Writing	3.1		1: Never—5: Very Often	No
							Reading	3.1		1: Never—5: Very Often	No

Stephenson et al.	2008	K	Mixed	CA	E	61	Letters	3.3	0–5	0: Never—5: More than once/ day	Yes
							Letter sounds	2.8	0–5	0: Never—5: More than once/ day	No
							Reading	1.9	0–5	0: Never—5: More than once/ day	No
Sénéchal	2006	K	High	CA	F	90	Letters	4.0[b]	2–5	1: Never—5: Very Often	Yes
							Reading	3.0[b]	1–5	1: Never—5: Very Often	Yes
							Writing	4.0[b]	1–5	1: Never—5: Very Often	Yes
Manolitsis et al.	2011	K	–	GRC	G	70	Letters	2.4		0: Never—5: More than once/ day	Yes
							Letter sounds	2.5		0: Never—5: More than once/ day	No
							Reading	1.7		0: Never—5: More than once/ day	No
Silinskas, Leppänen, et al.	2010	K	Mixed	FIN	Fi	207	Letters (mother)	2.7[c]	1–4	1: Not at all—4: Many times/ week	Yes
							Letters (father)	2.4[c]	1–4	1: Not at all—4: Many times/ week	Yes
Silinskas, Parrila, et al.	2010	K	Mixed	FIN	Fi	1529	Letters	2.9	1–5	1: Not at all/rarely—5: Very often/daily	No
							Reading	2.2	1–5	1: Not at all/rarely—5: Very often/daily	No

Grade = Equivalence given age of children in the sample: PS = Preschool or 4-year-olds, K = kindergarten or 5-year-olds
Country: AUS = Australia; CA = Canada; GRE = Greece; FIN = Finland; USA = United States of America
Lang. (Language): E = English; F = French; Fi = Finnish; G = Greek
N = total sample size
[a] Means were calculated using the percentages given in the study
[b] Median
[c] Mean for parental teaching practices when children were in kindergarten
[d] Teaching questions analyzed separately from other home literacy activities

a synopsis of parent reports of teaching frequency in twelve correlational studies, representing five countries. In the eight studies conducted in English, the median frequency of teaching was 3.5 on 5-point scales, where 5 corresponds to very often or more than once a day. The results suggest that English-speaking parents tend to teach their young child about literacy. Interestingly, the frequency of teaching was generally more frequent in English and less frequent in the studies conducted with Finnish and Greek families. Because the written language of Finnish and Greek is more transparent than English, this raises the possibility that cultural differences in the frequency of parent teaching might be linked to the difficulty level of learning to read.

Examination of Table 26.1 shows that respondents used the entire range of choices and that there was some variability in responses across questions and across studies. This pattern, along with the lack of public campaigns encouraging parents to teach their young child about literacy, provides some evidence that social desirability might not unduly influence parents in their responses to these questionnaire items.

Also noteworthy in Table 26.1 is the variability of questions used by researchers, from teaching letters to teaching to read. Given this variability, it is worth examining whether each question has the same predictive value. Martini and Sénéchal (2012) found that most parents in their middle-class sample indicated they often or very often taught their 5-year-old child the names and sounds of alphabet letters as well as to print letters and words. In addition, the majority of parents reported teaching their child to read words, albeit to a lesser degree. According to participating parents, teaching early literacy skills is prevalent in middle-class homes. Interestingly, further analyses of the parent teaching variables revealed that teaching behaviors formed two distinct factors: teaching basic literacy skills such as letter names and sounds as well as printing one's name and teaching more advanced skills such as reading words. Although this distinction is obvious, it is important to consider when describing differences across families, socioeconomic contexts, and cultures. For instance, Sénéchal and LeFevre (2002) found that it was more advanced teaching behaviors that were a unique predictor of children's early literacy in their middle-class sample. With a more socioeconomically diverse French-Canadian sample, in contrast, Sénéchal (20006) found that the predictor was more basic literacy teaching.

Having established the distinction between two types of home literacy experiences, their relations to child outcomes can be examined. The central prediction of the home literacy model is that informal and formal literacy experiences hold different relations to children's oral language, early literacy, and phoneme awareness. Specifically, the exposure to varied and rich language afforded by informal literacy experiences should be robustly linked to young children's oral language, and the print-focused interactions afforded by formal literacy experiences should be reliably associated with children's procedural knowledge about reading and writing. Finally, formal and informal literacy experiences should be indirectly linked to children's phoneme awareness, via children's oral vocabulary and early literacy. Each of these predictions is examined next.

Home Literacy and Child Outcomes

Understanding the impact of home literacy on children's reading success requires a clear conceptualization of young children's emerging competencies. In this chapter, early literacy refers to children's procedural knowledge about reading and writing, such as their knowledge of letters and their initial attempts at decoding and writing. In contrast, oral language refers to children's vocabulary and listening comprehension. Moreover, children's phoneme awareness is a metalinguistic skill that is linked to both oral and written language. Each of these constructs is treated separately because Sénéchal, LeFevre, Smith-Chant, and Colton (2001) showed that there was a complex interplay among them. For instance, children's oral vocabulary as well as their alphabet knowledge at age four predicted phoneme awareness at age 5. Children's alphabet knowledge and phoneme awareness at age 5 were excellent predictors of word reading in grade one, whereas vocabulary at age 5 was an excellent predictor of reading comprehension in grade three. In the next sections, the associations between home literacy and each of these child outcomes are described.

ORAL LANGUAGE

Exposure to the complex language as well as the variety of narrative forms found in children's books could have a beneficial effect on a number of child language outcomes. In Sénéchal's research, parent reports of the frequency and variety of shared reading were a robust predictor of children's receptive and expressive vocabulary. This was the case for English-speaking children in kindergarten and grade one (Sénéchal & LeFevre, 2002, 2014; Sénéchal et al., 1996; Sénéchal, LeFevre, Thomas, & Daley, 1998; Sénéchal, Pagan, Lever, & Ouellette, 2008) as well as for French-speaking children (Sénéchal, 2000, 2006). In this research, multiple indices of shared reading were used, including parent reports (the frequency of reading, the number of books in the home) and checklist measures to assess parent knowledge of children's books (Sénéchal et al., 1996). A combination of indices of shared reading was necessary because Sénéchal et al. (1998) showed that across studies, parents tended to report similar frequencies, and these do not always predict child vocabulary (for more recent examples of mixed results see Roberts, Jurgens, & Burchinal, 2005; Weigel, Martin, & Bennett, 2006). It might be the case that parent reports of shared reading frequency are subject to social desirability biases, at least in North America where numerous public campaigns have promoted shared reading. In contrast, analyses using a combination of measures consistently showed that shared reading accounted for 8% to 10% unique variance in children's vocabulary after controlling for children's intelligence, parent print exposure, and maternal education (when correlated with outcomes) or parent literacy. Consistent with the predictions of the home literacy model, parent teaching is not reliably associated with child oral language.

In most of the research linking shared reading and child vocabulary the parent and child measures were assessed concurrently. This makes it impossible to determine the direction of relations. For instance, children with larger vocabularies might ask their parents to read to them more often than children with smaller vocabularies. In recent work, however, Sénéchal and LeFevre (2014) showed that parent reports of shared reading predicted growth in vocabulary from kindergarten to grade one. That is, the researchers controlled for kindergarten vocabulary in their examination of the longitudinal relation between shared reading in kindergarten and child vocabulary in grade one. Importantly, these correlational findings are consistent with those found in intervention studies as reported by the National Early Literacy Panel (2008; effect size .60).

Shared reading has also been associated with young children's comprehension of morphologically complex words (Sénéchal et al., 2008). This association remained significant after controlling for child intelligence, parent literacy, and education. In contrast, the relation between shared reading and the comprehension of syntactically complex sentences was entirely mediated by parent literacy. It might be the case that child syntax comprehension requires that parents model complex syntax in addition to shared reading. That is, parents who read more might also use more complex syntax during shared reading as well as in other interactions with their children. Sénéchal et al. also examined the relation between narrative knowledge and shared reading. They had hypothesized that frequent shared reading introduces children to characters, events, and situations across a variety of books, and that such exposure would help a child produce a cohesive narrative. Surprisingly, they did not find a statistically significant association between the frequency of shared reading and measures of narrative cohesion (e.g., vocabulary diversity, mean length utterance) and coherence (i.e., story grammars). This finding might suggest that exposure to stories in children's literature is not sufficient to promote changes in 4-year-old children's production of narratives. It might be the case that additional parental support is necessary such that the quality of parent–child interactions during shared reading might influence narrative production more than simple exposure (Reese & Cox, 1999). For instance, intervention studies that increased the quality of adult–child interactions during shared reading have shown positive effects in kindergarten children's narrative production (e.g., Lever and Sénéchal, 2011; Zevenbergen, Whitehurst, & Zevenbergen, 2003). It is also possible that frequency of occurrence interacts with the quality of adult reading—a possibility that is never tested in the correlational research on shared reading.

EARLY LITERACY

In Sénéchal's tests of the home literacy model three early literacy skills, namely, alphabet knowledge, early reading, and invented spelling, have been examined because of their predictive role in children's reading success in grade school (Ouellette & Sénéchal, 2008; Sénéchal et al., 2001). Early reading was measured by asking children to read familiar three-letter words. When children could not read, the experimenter provided support by asking children to sound out each letter and then blend the letters. Invented spelling refers to young children's ability to capture, in nonconventional ways, the phonological structure of the words with their limited alphabetic knowledge (e.g., spelling *rough* as <*ruf*>). Children were asked to spell familiar words as best they could, and their spellings were scored to reflect the degree to which they mapped onto the phonology and orthography of the words (see Deacon & Sparks, this volume). The highest score was awarded for an accurate spelling (e.g., <*rough*>), the second-highest score was awarded for an accurate phonological representation of the word (e.g., <*ruf*>), and so on. Positive correlations between the frequency of parent reports of teaching and the three early literacy measures are typically found. Most important, parent teaching reports accounted for 4% to 19% unique variance in children's early literacy after controlling for child vocabulary and phoneme awareness, child nonverbal intelligence, parent education and income, or parent literacy (Martini & Sénéchal, 2012; Sénéchal, 2006; Sénéchal et al., 1998). The findings hold for concurrent as well as longitudinal relations between parent reports of teaching frequency and child early literacy (Sénéchal & LeFevre, 2014). In contrast, shared reading is not a statistically significant predictor of early literacy according to these studies.

Phoneme Awareness

Phoneme awareness measured in kindergarten is one of the most important predictors of word reading in grade one. Although there are differing views, most researchers agree that it is awareness of the phonemic structure of spoken language that facilitates children's entry into reading (as opposed to rhyming ability or awareness of rimes, for example). The importance of considering phonemic awareness separately from other variables becomes evident when comparing the results of Sénéchal et al. (1998) with those of Sénéchal and LeFevre (2002). Sénéchal et al. (1998) included phonemic awareness, vocabulary, and listening comprehension in a single composite measure. They found that shared reading predicted this general oral language factor. The analyses in Sénéchal and LeFevre, however, revealed that when phonemic awareness was analyzed separately, it was not associated with shared reading or parent teaching about literacy. Subsequent research confirmed that any correlation between home literacy and phoneme awareness is mediated by child vocabulary and early literacy (Hood, Conlon, & Andrews, 2008; Sénéchal, 2006; Sénéchal & LeFevre, 2014). Whether other types of parent–child interactions promote phonemic awareness directly remains to be ascertained while controlling for differences in children's vocabulary and early literacy.

Longitudinal Links to Literacy in Grade School

Findings from longitudinal studies have shown that the home literacy environment prior to grade one was linked to children's eventual success in reading in both indirect and direct ways. For instance, shared reading was linked to child vocabulary in kindergarten, which, in turn, was a predictor of reading comprehension in grades three (Sénéchal & LeFevre, 2002) and four (de Jong & Leseman, 2001; Sénéchal, 2006). Parent reports of teaching early literacy were linked to child early literacy, which, in turn, predicted word reading in grade one. Word reading in grade one then predicted reading comprehension in grades three and four. Noteworthy is the Sénéchal (2006) finding that parent reports of teaching about literacy, but not storybook exposure, were directly linked to child reading fluency in grade four after controlling for grade one reading and grade four reading comprehension. In contrast, parent reports of shared reading when their child was 5 years old predicted the frequency with which children reported reading for pleasure in grade four, after controlling for parent education, child vocabulary, word reading, and reading comprehension. What is it about these early experiences that have such a pervasive impact? It could be that they are markers of different types of orientations toward reading acquisition (Evans, Fox, Cresmaso, & McKinnon, 2004; Lynch, Anderson, Anderson, & Shapiro, 2006). For example, parents who taught their 5-year-olds frequently may continue to provide more support for learning in the form of listening to their child read out loud, and these continued experiences may be the building blocks for reading fluency (see Sénéchal & LeFevre, 2014). In contrast, shared reading may nourish children's motivation to read.

The central prediction of the home literacy model (Sénéchal & LeFevre, 2002) that informal and formal literacy experiences have different relations to children's oral language, early literacy, and phoneme awareness has been replicated in English (Bingham, 2009; Hood et al., 2008) and extended to Canadian French (Sénéchal, 2006), Korean (Lee, Sung, & Chang, 2009), and Spanish (Farver & Lonigan, 2009). One or the other component of the model has also been replicated in other research. Two English-Canadian studies showed that parent teaching, not shared reading, predicted children's letter knowledge (Evans, Shaw, & Bell, 2000; Stephenson, Parrila, Georgiou, & Kirby, 2008). Another English-Canadian study showed that shared reading predicted vocabulary, not letter knowledge (Frijters, Barron, & Brunello, 2000). Finally, a study conducted in Chile found that parent teaching was correlated with children's letter knowledge (Strasser & Lissi, 2009).

Importantly, observations of the quality of parents' didactic interactions with their 5-year-old children predicted the children's reading and writing in grade one (Aram, Korat, & Hassunah-Arafat, 2013). This observational finding with Arabic families also provides support for the idea that it is interactions with and about print that promote early literacy. However, not all findings support the home literacy model. For instance, parent reports of teaching were not a significant predictor of child letter knowledge in a sample of Greek families (Manolitsis, Georgiou, & Parrila, 2011). Moreover, parent reports of teaching were not related to children's concepts about print in a sample of children in Head Start—a program designed to increase cognitive, social, and emotional development in preschoolers from low-income homes in the United States (Sparks & Reese, 2012). In the latter studies parents reported, on average, infrequent teaching. Another study, in which parent teaching information was not collected, showed that child storybook knowledge mediated the relation between shared reading and child vocabulary and letter knowledge in a Dutch sample (Davidse, de Jong, Bus, Huijbregts, & Swaab, 2011). Storybook knowledge referred to children's recall of characters and storylines from popular children's books (for a similar measure in English, see Sénéchal et al., 1996). This study raises the question as to whether child book knowledge is an index of the child's ability to learn from interactions with the parent.

In sum, the key contribution of the home literacy model is that it helps us to disentangle the complex pattern of relations among the home literacy environment and child outcomes. Failure to consider the entire pattern of relations may lead to the conclusion that storybook exposure or teaching about literacy have a wider range of associations to reading rather than a more focused impact (Burgess, Hecht, & Lonigan, 2002; de Jong & Leseman, 2001; Mol et al., 2008).

The results discussed in the section are correlational, and the parent teaching behaviors are based on parent responses to questionnaires. Even though stringent analyses that controlled for potential confounds were conducted, the results do not address the causal relations between variables. It is therefore important to examine the experimental and quasi-experimental research to assess whether parent–child activities have a positive effect on children's learning. Intervention research on parent–child shared reading has shown its efficacy for promoting vocabulary acquisition (for a meta-analysis, see National Early Literacy Panel, 2008). A synthesis of intervention research on child early literacy is presented in the next section.

Parent Teaching: A Meta-Analysis of Its Impact on Early Literacy

Causal statements about the role of parents in children's early literacy acquisition require intervention research. Sénéchal and Young (2008) conducted a meta-analysis on the impact of parent involvement in their child literacy acquisition. The initial search of the research literature was conducted between 2004 and 2006. To be included, interventions had to be conducted with parents of children from kindergarten up to grade three, the point at which most children have acquired sufficient fluency to use reading to learn about other academic domains (Indrisano & Chall, 1995). Presented in this section is an updated version of the meta-analysis for which the literature searches were conducted in the fall of 2013. Given the scope of this chapter, the findings presented are limited to children in preschool (five studies) and kindergarten (fourteen studies).

The meta-analysis focused on early literacy (alphabet knowledge, early reading, invented spelling) as well as reading (decoding, fluency, comprehension) and spelling. In some cases when researchers also measured metalinguistic skills known to predict reading (e.g., phonemic awareness), these measures were combined with the reading measures. The term "reading" is used subsequently to represent this set of variables. Importantly, the meta-analysis excluded measures of oral language or conceptual measures about print (e.g., concepts such as knowing that it is the print that is read in a book, not the pictures). These print concepts were excluded because there is no clear evidence that this conceptual knowledge robustly predicts reading in models that control for procedural literacy skills (Sénéchal et al., 2001). Standard meta-analytical procedures were used and described in detail in Sénéchal and Young (2008); a brief overview is presented next.

Overview of the Methodology

To find relevant research, searches of electronic databases were conducted with prespecified keywords as well as searches of the reference sections of review articles and all articles retained from the electronic searches. In addition to the age level and literacy measures criteria, studies selected had to meet the following criteria: (1) tested the hypothesis that parent involvement affects the acquisition of reading; (2) used an experimental or a quasi-experimental design (i.e., no random assignment of participants to conditions); (3) were published in a peer-reviewed journal; (4) included at least five participants; and (5) reported effect sizes or statistics permitting the calculation or estimation of effect sizes.

Updating the Sénéchal and Young (2008) meta-analysis resulted in fifteen articles on 4- and 5-year-olds. In four of these articles, the research design included four conditions of which two were of interest (Baker, Plotrkowski, & Brooks-Gunn, 1998; Chow, McBride-Chang, Cheung, & Chow, 2008; Harper, Platt, & Pelletier, 2011; Levin & Aram, 2012). The two conditions of interest were labeled treatment conditions, and they were randomly assigned one of the two remaining conditions, labeled controls. Each treatment-control pair was labeled Study 1 and Study 2, respectively. This approach added four independent samples to the analyses.

The primary statistic used to integrate and compare the nineteen studies was Cohen's *d,* a measure of effect size. Cohen's *d* is the standardized difference between the intervention group and the control group (or an estimate of the difference). Hence, an effect size of 1 represents a difference of one standard deviation between the intervention and the control groups. For example, if a study used a standardized test with a mean of 100 and a standard deviation of 15, then an effect size of 1 represents a 15-point advantage for the intervention group over the control group. Similarly, an effect size of .50 represents a 7.5-point difference between the intervention and the control groups, and an effect size of 0 represents no difference between the intervention and the control groups. Effect sizes can also have negative values that indicate that the mean for the control group is superior to that of the intervention group. In the present report, the description of effect sizes in terms of points gained on a test was used to gauge the magnitude of the effects.

For the 2013 update, two studies included two control groups where one received an alternative treatment and the other did not receive any treatments. In these cases, effect sizes using the alternative-treatment group were used as the control, thus controlling for possible halo effects. Moreover, there were two studies for which Cohen's *d* could not be calculated directly because the means and standard deviations were not reported for the treatment and control groups. For these studies, effect sizes were estimated from *F* statistics (St. Clair & Jackson, 2006) or Mann-Whitney *U* statistics (Drouin, 2009).

In studies that included multiple outcome measures, a single estimate of effect size per study was calculated to ensure effect size independence. Producing a single estimate of effect size for each study was done in five steps: (1) for studies using standardized and experimenter designed tests, only the standardized measures were used to optimize comparisons across studies; (2) for studies reporting composite scores as well as subtest scores for the same standardized test, only the composite scores were used; (3) for studies that included immediate and delayed posttests, only the measures for the immediate posttest were used to optimize the comparisons with the studies that included immediate posttests only; (4) in studies including multilingual families, only the measures in the language of the intervention were used; and (5) effect sizes were computed for each remaining measure and the median effect size was used as a single estimate for each study.

Combining effect sizes across studies was conducted using standard meta-analytic procedures (Cooper & Hedges, 1994; Hedges & Olkin, 1985). In all cases, mean effect sizes were weighted to acknowledge that studies with larger samples provide more reliable estimates of the population effect size (Hedges & Olkin, 1985). For each effect size, 95% confidence intervals are provided to assess whether it was statistically significantly different from zero. For the nineteen studies and each type of intervention, a heterogeneity statistic, *Q*, was computed to assess whether the variability in effect sizes across studies was greater than would be expected by chance (that is, $p \leq .05$). A significant *Q* statistic indicates that further analyses should be conducted to ascertain the locus of this variability. In cases when the variability cannot be explained, caution is warranted in interpreting the findings.

Results

Preliminary analyses revealed that there were no study results that were outliers (Hedges & Olkin, 1985). The 19 studies included in the meta-analysis are presented in Table 26.2. Combining the results of these intervention studies, representing 1342 families, showed that parent involvement had a positive impact on children's reading acquisition. The mean weighted effect size was moderately large at .36 and statistically significantly greater than zero, 95% CI [.24, .47]. This effect size corresponds to a 5.4-point gain on a standardized test (with a standard deviation of 15) for the intervention children compared to the control children. As Table 26.2 shows, however, there was considerable variability in the magnitude of effect sizes across studies, ranging from a low of -.04 to a high of 1.37. This variability across studies was greater than was expected by chance as indicated by a statistically significant test of heterogeneity, $Q = 48.00$, $p < .05$. As a consequence, the types of interventions were examined as a potential moderator that would explain this variability.

Sénéchal and Young (2008) made no a priori decisions about the types of parent–child activities to be included in their synthesis. Once collected and analyzed, three categories of activities emerged: Parents were asked to read to their child; parents were asked to listen to their child read books; and parents were trained to do literacy exercises with their children, serving as literacy tutors. Given our focus on children who are not yet receiving reading instruction at school (i.e., prior to the North American grade one), it is not surprising that we did not find any studies where parents were asked to listen to children read. Therefore, this category is not discussed further. In this update, however, we found five studies where parents were asked to read to their child as well as tutor them during other activities. These studies were kept separate in a new category labeled *Read books and tutor with specific activities*. Each relevant category is described next.

PARENTS READ TO CHILD

This category was the largest and included nine studies two from 2008 and seven from the 2013 update, in which parents were encouraged to read to their child. Parents in one study were instructed by researchers on appropriate reading practices such as how to read aloud effectively to their children, choose appropriate books, select a quiet environment and an optimum time of day, and ensure child interest in the books (Foster & Bitner, 1998). In three studies, parents were trained to use a dialogic reading technique with their children (Chow & McBride-Chang, 2003; Chow et al., 2008, Study 1; Chow, McBride-Chang, & Cheung, 2010). In dialogic reading, parents prompt their children to talk about the story, evaluate their children's responses, rephrase and add information to their children's responses, and prompt their children again in order to assess their children's learning. In another study, mothers read a storybook to their child and asked relevant questions promoting interactive discussions (Levin & Aram, 2012, Study 1). In two studies, parents were trained to reference print during shared reading by asking the children to show, for instance, a letter or the longest word on a page (Justice & Ezell, 2000); Justice, Skibbe, McGinty, Piasta, & Petrill, 2011). Lastly, the structured program Hope Instruction Program for Preschool Youngsters (HIPPY) was employed in two studies (Baker et al., 1998, Study 1 and Study 2). In this program, parents receive lessons that help them read and participate in activities at home with their child.

The mean weighted effect size for the *Read to child* category was small at .09, and it was not statistically significantly different from zero, 95% CI [-.08, .27]. As shown in Table 26.2, effect sizes across studies, representing 509 families, ranged from -.05 to .28, and the set was homogeneous, $Q = 1.33$, $p > .05$. As a whole, this set of nine studies shows that training parents to read to their child did not increase children's early literacy significantly.

PARENTS READ BOOKS AND TUTOR WITH SPECIFIC ACTIVITIES

This category included one study from 2008 and five studies from the 2013 update. In all these studies, parents received training on shared reading as well as teaching specific literacy skills. Parents in one study were encouraged to engage in scripted parent–child interactions and extended book-related discussions, as well as to complete book-based activities related to monthly themes with their child including eight language themes, three narrative themes, and one letter names and sounds theme (Jordan, Snow, & Porche, 2000). This specific study was the only shared reading study in the 2008 sample that added activities and, consequently, Sénéchal and Young (2008) included it in the *Parents read books* category. In the 2013 update, this study moved from the *Parents read books*

Table 26.2 Effect Sizes, Intervention, Child, and Study Characteristics as a Function of Intervention Type.

	Intervention				Child			Study						Cohen's *d*
	Training (h)	Support to Parents	Length (mo.)	Lang.	Grade	Dev. Level	SES	Random Assign.	*N*	Test time	Standard Test Given	Outcome Measure	Country	
Parents Read Books to Child (9 studies; 509 families; *d = .09*)														
Baker et al. Study 1(1998)	0.3	No	12	E	PS	2	1	No	69	I/D	Yes	3	US	0.28
Baker et al. Study 2 (1998)	0.3	No	12	E	PS	2	1	No	113	I/D	Yes	3	US	0.09
Chow and McBride-Chang (2003)	4	Yes	2	C	K	1	3	Yes	58	I	Yes	3	HK	0.18
Chow et al. (2010)	–	No	3	E[a]	K	1	2	No	34	I	Yes	1, 2, 3	CHI	–0.05
Chow et al. Study 1 (2008)	1	No	3	C	K	1	2	No	74	I	Yes	1, 2	CHI	0.03
Foster and Bitner (1998)	–	No	3	E	K	2	1	Yes	35	I	Yes	1	US	0.07
Justice and Ezell (2000)	0.25	No	1	E	PS	1	3	No	28	I	No	1	US	0.20
Justice et al. (2011)	–	Yes	3	E	PS	3	4	Yes	36	I	No	1	US	–0.04
Levin and Aram Study 1 (2012)	3	Yes	1.75	H	K	1	1	No	62	I/D	No	5	ISR	0.01
Parents Read Books and Tutor Specific Literacy Skills With Activities (6 studies; 551 families; *d = .33)**														
Fielding-Barnsley and Purdie (2003)	–	No	2	E	K	2	–	No	49	I/D	No	1	AUS	0.54*
Chow et al. Study 2 (2008)	–	No	3	C	K	1	2	No	74	I	Yes	1, 2	CHI	0.25
Drouin (2009)	–	No	1	E	PS	1	–	No	48	I/D	No	3, 4	UK	0.57*
Harper et al. Study 1 (2011)	13.5	Yes	9	E	K	1	–	No	55	I	Yes	1, 3, 4	CA	0.22
Harper et al. Study 2 (2011)	13.5	Yes	9	E[a]	K	2	–	No	77	I	Yes	1, 3, 4	CA	0.57
Jordan et al. (2000)	5	No	5	E	K	1	2	No	248	I	Yes	1	US	0.20

Parents Tutor Specific Literacy Skills with Activities (*N* = 4 studies; 282 families; *d* = .94*)														
Kraft et al. (2001)	2	Yes	4	E	K	1	2	No	43	I	Yes	3	US	0.41
Levin and Aram Study 2 (2012)	3	Yes	1.75	H	K	1	1	No	62	I/D	No	5	ISR	0.39
Niedermeyer (1970)	1.5	Yes	3	E	K	1	2	No	148	D	No	5	US	1.37*
St. Clair and Jackson (2006)	12.5	No	12	E[a]	K	2	1	No	29	I/D	Yes	5	US	1.18*

Note. Studies added in 2013 are in italics. A hyphen indicates missing information
Grade = Equivalence given age of children in the sample: PS = Preschool or 4-year-olds, K = kindergarten or 5-year-olds
Language: C = Cantonese, E = English, E[a] = English as a second language and testing in English, H = Hebrew
Dev. (Developmental) Level: 1 = normal, 2 = at risk, 3 = language impaired
SES: 1 = Low income, 2 = Middle/high income, 3 = mixed, 4 = low/middle
N = total sample size
Test time: I = Immediately after the intervention, D = After a delay, I/D = Immediate and delayed testing
Outcome measure: 1 = early literacy (alphabet knowledge, reading with help, invented spelling), 2 = phoneme awareness or morphological awareness, 3 = reading, 4 = spelling, 5 = composite measure
Country: AUS = Australia; CA = Canada; CHI = China; HK = China (Hong Kong); ISR = Israel; UK = England; US = United States of America; ISR = Israel
[a] Participants were matched
* $p \leq .05$

category to the *Parents read and tutor* category. In two studies, parents were trained to use dialogic reading as well as to teach alphabetic knowledge (Fielding-Barnsley & Purdie, 2003) or teach morphological awareness (an early predictor of reading in Chinese, Study 2, Chow et al., 2008). In Harper et al. (2011) Studies 1 and 2, parents were trained to give a structured program that included shared reading and teaching letter names and sounds. In another study, parents were trained in shared reading and supporting their children's learning of letter sounds, reading, writing their own name, and rhyming (Drouin, 2009).

The mean weighted effect size for the *Read and tutor* category was moderate at .33 and statistically significant, 95% CI [.09, .57]. As shown in Table 26.2, effect sizes across studies, representing 551 families, ranged from .20 to .57, and the set was homogeneous, $Q = 2.33$, $p > .05$. This effect size corresponds to a 5-point gain on a hypothetical standardized test (with a standard deviation of 15) for the intervention children as compared with the control children. Taken together, this small set of studies shows the efficacy of training parents in more print- or code-focused activities along with shared reading.

PARENTS TUTOR SPECIFIC LITERACY SKILLS WITH ACTIVITIES

The third category included four studies (two in 2008 and two in 2013) in which parents were trained to tutor their child on specific literacy skills. In four studies with kindergarten children, parents were trained to: (1) teach letter-sound correspondences and letter-sound blending (Kraft, Findlay, Major, Gilberts, & Hofmeister, 2001); (2) support their children's learning of letter names and sounds and sight words (St. Clair & Jackson, 2006); (3) help their children during writing activities (Levin & Aram, 2012 Study 2); and (4) implement practice exercises to learn to read one-syllable words (Niedermeyer, 1970).

The mean weighted effect size for the *Parents tutor* category was large at .94, 95% CI [.70, 1.19]. This effect size corresponds to a 14-point gain on a hypothetical standardized test (with a standard deviation of 15) for the intervention children as compared with the control children. As shown in Table 26.2, effect sizes across studies, representing 282 families, ranged from .39 to 1.37, and as a consequence of this wide range, the set was not homogeneous, $Q = 13.45$, $p < .05$. Therefore, one should use prudence in interpreting these positive findings.

COMPARING INTERVENTION TYPES

Training parents to tutor their child using specific reading activities produced the largest effect size at .94. This effect size for the *Parents tutor* category was significantly greater than the effects of the *Parents read and tutor* category (ES = .33), $z = 4.73$, 95% CI [.35, .84]. This result suggests that adding a shared reading component to interventions for which parents were also trained to tutor their child with specific activities did not increase the impact of the intervention on children's early literacy. In contrast, the set of studies asking parents to read to their child yielded a small effect size (.09) that was not statistically significantly different from zero. This comparison across types of early literacy interventions is important in light of Bus, van IJzendoorn, and Pellegrini's (1995) suggestion that shared reading was a source of early literacy learning based on their meta-analysis of correlational evidence. The intervention research in this meta-analysis did not support this claim.

Conclusion

In this chapter, converging correlational as well as experimental and quasi-experimental evidence was presented that supported a specific view of the home literacy environment. The evidence presented is in accord with a proposed home literacy model (Sénéchal 2006; Sénéchal & LeFevre, 2002). Informal literacy activities such as shared reading help children to learn about oral language whereas parent–child literacy interactions that focus on the form of print seem to be necessary for gains in early literacy.

Future Directions

SHARED READING

The first section of the chapter presented evidence that parents read books to their child for enjoyment and the quality time it affords. Also presented were reasons why shared reading is an occasion for learning oral language (and world knowledge, for that matter). In support of this, correlational evidence showed that the frequency of shared reading is linked to the breadth of children's vocabularies. Moreover, syntheses of experimental and quasi-experimental studies have shown that parents can be trained to increase the quality of parent–child interactions and that doing so increases children's expressive vocabulary. In contrast to these positive effects, the second section showed that shared reading is not a source of early literacy learning for children. Hence, it seems that parents of young children may be right in limiting the number of print-focused interactions during shared reading. This is food for thought for researchers and practitioners who might be tempted to transform shared reading into a source of early literacy learning.

Numerous questions still need to be addressed about shared reading. For example, rather than improving early literacy, it could be that shared reading increases children's motivation to read for pleasure. Hints of such a pattern were found in Sénéchal (2006), where the frequency of shared reading with 5-year-olds predicted the frequency of children reading for pleasure four years later.

PARENT TEACHING ABOUT LITERACY

The first section of the chapter showed that parents report teaching about the mechanics of reading and writing at home. The children of parents who report teaching more frequently tend to have stronger early literacy skills than children whose parents report teaching less. Moreover, parents seem to take advantage of naturally occurring occasions to impart knowledge about letters, reading, and writing. The second section showed that when parents are trained to do specific activities that focus on the mechanics of literacy, children do better than children whose parents were not trained or did alternative activities. Hence, we showed that parents do and can help their children learn early literacy skills. The issue to consider is whether parents should. The home literacy model was meant to describe what parents do at home and to show accurately the relation between what parents do and child outcomes. The model, however, was not intended to be prescriptive. Also, it is debatable whether the findings of the meta-analysis should be used to encourage parents to tutor their young children about literacy. Perhaps it is time to turn our attention to studying how to build strong partnerships between the home and the school in order to optimize early literacy in young children.

RESEARCH

In reviewing the research on home literacy environments, it is always striking how often the methodology and the reporting are weaker than they should be. This becomes evident when conducting a meta-analysis. Of the studies initially selected for the review, 60% did not provide an alternative treatment for the families in the control groups and 87% did not randomly assign families to conditions. Moreover, some studies had to be excluded because there was no control group or because the researchers did not report sufficient information to calculate effect sizes. Finally, 12% of the studies in Table 26.2 did not report any descriptive statistics. A challenge for researchers is to replicate study findings with the best methodological and reporting standards.

Acknowledgment

The author is indebted to Melissa Malette for her invaluable help in writing this chapter.

References

* Indicates articles included in the meta-analysis

Audet, D., Evans, M. A., Mitchell, K., & Reynolds, K. (2008). Shared book reading: Parental goals across the primary grades and goal-behavior relationships in junior kindergarten. *Early Education and Development*, *19*, 113–138.

Alexander, K. L., Entwisle, D. R., & Horsey, C. (1997). From first grade forward: Early foundation of high school dropout. *Sociology of Education*, *70*, 87–107.

Aram, D., Korat, O., & Hassunah-Arafat, S. (2013). The contribution of early home literacy activities to first grade reading and writing achievements in Arabic. *Reading and Writing: An Interdisciplinary Journal*, *26*, 1517–1536.

Aram, D., & Levin, I. (2004). The role of maternal mediation of writing to kindergartners in promoting literacy achievements in school: A longitudinal perspective. *Reading and Writing: An Interdisciplinary Journal*, *17*, 387–409.

*Baker, A., Plotrkowski, C. S., & Brooks-Gunn, J. (1998). The effects of the Home Instruction Program for Preschool Youngsters on children's school performance at the end of the program and one year later. *Early Childhood Research Quarterly*, *13*, 571–588.

Bingham, G. (2009, April). *Maternal literacy beliefs, joint book reading, home literacy environment, and children's early literacy development.* Poster presented at the biennial meeting of the Society for Research in Child Development, Denver, CO.

Boudreau, D. (2005). Use of a parent questionnaire in emergent and early literacy assessment of preschool children. *Language, Speech, and Hearing Services in Schools*, *36*, 33–47.

Burgess, S. R., Hecht, S. A., & Lonigan, C. J. (2002). Relations of the home literacy environment (HLE) to the development of reading-related abilities: A one-year longitudinal study. *Reading Research Quarterly*, *37*, 408–426.

Bus, A. G., van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, *65*, 1–21.

Butler, S. R., Marsh, H. W., Sheppard, M. J., & Sheppard, J. L. (1985). Seven-year longitudinal study of the early prediction of reading achievement. *Journal of Educational Psychology*, *77*, 349–361.

*Chow, B. W., & McBride-Chang, C. (2003). Promoting language and literacy development through parent-child reading in Hong Kong preschoolers. *Early Education and Development*, *14*, 233–248.

*Chow, B. W.-Y., McBride-Chang, C., & Cheung, H. (2010). Parent-child reading in English as a second language: Effects on language and literacy development of Chinese kindergarteners. *Journal of Research in Reading*, *33*, 284–301.

*Chow, B. W., McBride-Chang, C., Cheung, H. & Chow, C. S. (2008). Dialogic reading and morphology training in Chinese children: Effects on language and literacy. *Developmental Psychology*, *44*, 233–44.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis.* New York, NY: Russell Sage Foundation.

Crain-Thoreson, C., Dhalin, M. P., & Powell, T. A. (2001). Parent-child interaction in three conversational contexts: Variations in style and strategy. In J. Brooks-Gunn & P. Rebello (Eds.) *Sourcebook on emergent literacy* (pp. 23–37). San Francisco, CA: Jossey-Bass.

Davidse, N. J., de Jong, M. T., Bus, A. G., Huijbregts, S. C. J., & Swaab, H. (2011). Cognitive and environmental predictors of early literacy skills. *Reading and Writing: An Interdisciplinary Journal*, *24*, 395–412.

de Jong, P. F., & Leseman, P. P. M. (2001). Lasting effects of home literacy on reading achievement in school. *Journal of School Psychology*, *39*, 389–414.

Deckner, D. F., Adamnson, L. B., & Bakeman, R. (2006). Child and maternal contributions to shared reading: Effects on language and literacy development. *Applied Developmental Psychology*, *27*, 31–41.

*Drouin, M. (2009). Parent involvement in literacy intervention: A longitudinal study of effects on preschoolers' emergent literacy skills and home literacy environment. *Early Childhood Services*, *3*, 1–18.

Edwards, P. A. (1991). Fostering early literacy through parent coaching. In E. H. Hiebert (Ed.), *Literacy for a diverse society: Perspectives, practices, and policies* (pp. 199–213). New York, NY: Teachers College Press.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2005). First grade and educational attainment by age 22: A new story. *American Journal of Sociology*, *110*, 1458–1502.

Evans, M. A., Fox, M., Cresmaso, L., & McKinnon, L. (2004). Beginning reading: The views of parents and teachers of young children. *Journal of Educational Psychology*, *96*, 130–141.

Evans, M. A., Moretti, S., Shaw, D., & Fox, M. (2003). Parent scaffolding in children's oral reading. *Early Education and Development*, *14*, 363–388.

Evans, M. A. & Saint-Aubin, J. (2005). What children are looking at during shared storybook reading: Evidence from eye movements. *Psychological Science*, *16*, 913–920.

Evans, M. A., Shaw, D., & Bell, M. (2000). Home literacy activities and their influence on early literacy skills. *Canadian Journal of Experimental Psychology*, *54*, 65–75.

Farver, J. M., & Lonigan, C. J. (2009, April). *The development of ELL children's early reading skills in Spanish and English: What can immigrant families do?* Poster presented at the meeting of the Society for Research in Child Development, Denver, CO.

*Fielding-Barnsley, R., & Purdie, N. (2003). Early intervention in the home for children at risk of reading failure. *Support for Learning*, *18*, 77–82.

*Foster, S. M., & Bitner, T. R. (1998). A read-aloud project for at-risk kindergarten children and their parents. *Indiana Reading Journal*, *30*, 50–55.

Foy, J. G., & Mann, V. (2003). Home literacy environment and phonological awareness in preschool children: Differential effects for rhyme and phoneme awareness. *Applied Psycholinguistics*, *24*, 59–88.

Frijters, J. C., Barron, R. W., & Brunello, M. (2000). Direct and mediated influences of home literacy and literacy interest on prereaders' oral vocabulary and early written language skill. *Journal of Educational Psychology*, *92*, 466–477.

*Harper, S., Platt, A., & Pelletier, J. (2011). Unique effects of a family literacy program on the early reading development of English language learners. *Early Education and Development*, *22*, 989–1008.

Hart, S. A., & Petrill, S. A. (2009). The genetics and environments of reading: A behavioral genetic perspective. In Y. Kim (Ed.), *Handbook of behavior genetics Part II* (pp. 113–123). New York, NY: Springer.

Hayes, D. P., & Ahrens, M. (1988). Vocabulary simplification for children: A special case of motherese? *Journal of Child Language*, *15*, 395–410.

Hedges, L. V., & Olkin, O. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hindman, A. H., Connor, C. M., Jewkes, A. M., & Morrison, F. J. (2008). Untangling the effects of shared book reading: Multiple factors and their associations with preschool literacy outcomes. *Early Childhood Research Quarterly*, *23*, 330–350.

Hindman, A. H., Skibbe, L. E., & Foster, T. D. (2013). Exploring the variety of parental talk during shared book reading and its contributions to preschool language and literacy: Evidence from the Early Childhood Longitudinal Study-Birth Cohort. *Reading and Writing: An Interdisciplinary Journal*, *27*, 287–313.

Hood, M., Conlon, E., & Andrews, G. (2008). Preschool home literacy practices and children's literacy development: A longitudinal analysis. *Journal of Educational Psychology*, *100*, 252–271.

Indrisano, R., & Chall, J. (1995). Literacy development. *Journal of Education*, *177*, 63–83.

*Jordan, G. E., Snow, C. E., & Porche, M. V. (2000). Project EASE: The effect of a family literacy project on kindergarten students' early literacy skills. *Reading Research Quarterly*, *35*, 524–546.

*Justice, L. M., & Ezell, H. K. (2000). Enhancing children's print and word awareness through home-based parent intervention. American Journal of Speech-Language Pathology, *9*, 257–269.

Justice, L. M., Pullen, P. C., & Pence, K. (2008). Influence of verbal and nonverbal references to print on preschoolers' visual attention to print during storybook reading. *Developmental Psychology*, *44*, 855–866.

*Justice, L. M., Skibbe, L. E., McGinty, A. S., Piasta, S. B., & Petrill, S. A. (2011). Feasibility, efficacy, and social validity of home-based storybook reading intervention for children with language impairment. *Journal of Speech*, *Language*, *and Hearing Research*, *54*, 523–538.

*Kraft, B. L., Findlay, P., Major, J., Gilberts, G., & Hofmeister, A. (2001). The association between a home reading program and young children's early reading skill. *Journal of Direct Instruction*, *1*, 117–136.

Landry, S. H., Smith, K. E., Swank, P. R., & Guttentag, C. (2008). A responsive parenting intervention: The optimal timing across early childhood for impacting maternal behaviors and child outcomes. *Developmental Psychology*, *44*, 1335–1353.

Lee, K., Sung, M., & Chang, Y. E. (2009, April) *Relations of home literacy environments to preschoolers' vocabulary and reading.* Poster presented at the meeting of the Society for Research in Child Development, Denver, CO.

Lever, R., & Sénéchal, M. (2011). Discussing stories: On how a dialogic reading intervention improves kindergartners' oral narrative construction. *Journal of Experimental Psychology*, *108*, 1–24.

*Levin, I., & Aram, D. (2012). Mother–child joint writing and storybook reading and their effects on kindergartners' literacy: An intervention study. *Reading and Writing: An Interdisciplinary Journal*, *25*, 217–249.

Lynch, J., Anderson, J., Anderson, A., & Shapiro, J. (2006). Parents' beliefs about young children's literacy development and parents' literacy behaviors. *Reading Psychology*, *27*, 1–20.

Manolitsis, G., Georgiou, G. K., & Parrila, R. (2011). Revisiting the home literacy model of reading development in an orthographically consistent language. *Learning and Instruction*, *24*, 496–505.

Martini, F., & Sénéchal, M. (2012). Learning literacy skills at home: Parent teaching, expectations, and child interest. *Canadian Journal of Behavioural Science*, *44*, 210–221.

Mol, S. E., Bus, A. G., de Jong, M. T., & Smeets, D. J. H. (2008). Added value of dialogic parent-child book readings: A meta-analysis. *Early Education and Development*, *19*, 7–26.

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel.* Washington, DC: National Institute for Literacy.

*Niedermeyer, F. C. (1970). Parents teach kindergarten reading at home. *Elementary School Journal*, *70*, 438–445.

Ouellette, G., & Sénéchal, M. (2008). Pathways to literacy: A study of invented spelling and its role in learning to read. *Child Development*, *79*, 899–913.

Phillips, B. M., & Lonigan, C. J. (2009). Variations in the home literacy environment of preschool children: A cluster analytic approach. *Scientific Studies of Reading*, *13*, 146–174.

Reese, E., & Cox, A. (1999). Quality of adult book reading affects children's emergent literacy. *Developmental Psychology*, *35*, 20–28.

Roberts, J., Jurgens, J., & Burchinal, M. (2005). The role of home literacy practices in preschool children's language and emergent literacy skills. *Journal of Speech Language and Hearing Research*, *48*, 345–359.

Robins, S., & Treiman, R. (2009). Talking about writing: What we can learn from conversations between parents and their young children. *Applied Psycholinguistics*, *30*, 463–484.

Robins, S., Treiman, R., & Rosales, N. (2014). Letter knowledge in parent-child conversations. *Reading and Writing: An Interdisciplinary Journal, 27*, 407–429.

Robins, S., Treiman, R., Rosales, N., & Otake, S. (2012). Parent-child conversations about letters and pictures. *Reading and Writing: An Interdisciplinary Journal, 25*, 2039–2059.

Sénéchal, M. (1997). The differential effect of storybook reading on preschooler's expressive and receptive vocabulary acquisition. *Journal of Child Language*, *24*, 123–138.

Sénéchal, M. (2000). Morphological effects in children's spelling of French words. *Canadian Journal of Experimental Psychology*, *54*, 76–86.

Sénéchal, M. (2006). Testing the home literacy model: Parent involvement in kindergarten is differentially related to grade 4 reading comprehension, fluency, spelling, and reading for pleasure. *Scientific Study of Reading*, *10*, 59–87.

Sénéchal, M., & LeFevre, J. (2002). Parental involvement in the development of children's reading skill: A 5-year longitudinal study. *Child Development*, *73*, 445–460.

Sénéchal, M., & LeFevre, J. (2014). Continuity and change in the home literacy environment as predictors of growth in vocabulary and reading. *Child Development*, *85*, 1552–1559.

Sénéchal, M., LeFevre, J.-A, Hudson, E., & Lawson, P. (1996). Knowledge of picture-books as a predictor of young children's vocabulary development. *Journal of Educational Psychology*, *88*, 520–536.

Sénéchal, M., LeFevre, J., Smith-Chant, B. L., & Colton, K. (2001). On refining theoretical models of emergent literacy: The role of empirical evidence. *Journal of School Psychology, 39*, 439–460.

Sénéchal, M., LeFevre, J.-A., Thomas, E., & Daley, K. (1998). Differential effects of home literacy experiences on the development of oral and written language. *Reading Research Quarterly*, *32*, 96–116.

Sénéchal, M., Pagan, S., Lever, R., & Ouellette, G. (2008). Relations among the frequency of shared reading and 4-year-old children's vocabulary, morphological and syntax comprehension, and narrative Skills. *Early Education and Development*, *19*, 28–45.

Sénéchal, M., & Young, L. (2008). The effect of family literacy interventions on children's acquisition of reading from kindergarten to grade 3: A meta-analytic review. *Review of Educational Research*, *78*, 880–907.

Silinskas, G., Leppänen, U., Aunola, K., Parrila, R., & Nurmi, J.-E. (2010). Predictors of mothers' and fathers' teaching of reading and mathematics during kindergarten and Grade 1. *Learning and Instruction*, *20*, 61–71.

Silinskas, G., Parrila, R., Lekkanen, M.-K., Poikkeus, A.-M., Niemi, P., & Nurmi, J.-E. (2010). Mothers' reading-related activities at home and learning to read during kindergarten. *European Journal of Psychology of Education*, *25*, 243–264.

Sparks, A., & Reese, E. (2012). From reminiscing to reading: Home contributions to children's developing language and literacy in low-income families. *First Language*, *33*, 89–109.

Stadler, M. A., & McEvoy, M. A. (2003). The effect of text genre on parent use of joint book reading strategies to promote phonological awareness. *Early Childhood Research Quarterly*, *18*, 502–512.

*St. Clair, L., & Jackson, B. (2006). Effect of family involvement training on the language skills of young elementary children from migrant families. *School Community Journal*, *16*, 31–41.

Stephenson, K. A., Parrila, R. K., Georgiou, G. K., & Kirby, J. R. (2008). Effects of home literacy, parents' beliefs, and children's task-focused behavior on emergent literacy and word reading skills. *Scientific Studies of Reading*, *12*, 24–50.

Strasser, K., & Lissi, M. R. (2009). Home and instruction effects on emergent literacy in a sample of Chilean kindergarten children. *Scientific Studies of Reading*, *13*, 175–204.

Teale, W. H., & Sulzby, E. (1986). Emergent literacy as a perspective for examining how young children become writers and readers. In W. H. Teale & E. Sulzby (Eds.), *Emergent literacy: Writing and reading* (pp. vii–xxv). Norwood, NJ: Ablex.

Weigel, D. J., Martin, S. S., & Bennett, K. K. (2006). Mothers' literacy beliefs: Connections with the home literacy environment and preschool children's literacy development. *Journal of Early Childhood Literacy*, *6*, 191–211.

Zevenbergen, A., Whitehurst, G., & Zevenebergen, J. (2003). Effects of a shared-reading intervention on the inclusion of evaluative devices in narratives of children from low-income families. *Applied Development Psychology*, *24*, 1–15.

CHAPTER

27 Primary Grade Reading Instruction in the United States

Carol McDonald Connor *and* Stephanie Al Otaiba

Abstract

This chapter reviews recent US policy and research on literacy instruction in the primary grades, kindergarten through fifth grade. Four topics are discussed: policy and seminal reports, decoding and multitiered systems of instruction, comprehension, and individualized or differentiated instruction. The research review focuses on typically developing students, students who have or are at risk for reading disabilities, and English language learners. Research indicates that multitiered systems of instruction show promise although they have focused primarily on decoding and not on comprehension. For comprehension, multicomponent strategies may be more effective than single-strategy interventions. However, mixed results—with various studies on teaching students how to comprehend what they are reading showing positive, null, and negative effects—indicate that more research is needed. Individualizing the instruction provided to students based on assessment results appears to improve decoding and comprehension outcomes. Future directions for policy and research are also discussed.

Key Words: reading, literacy, comprehension, instruction, intervention, decoding, meaning-focused instruction, fundamental skill, differentiated instruction, individualized instruction

Teaching students how to read and write may be one of the most difficult but important endeavors facing a literate society—especially if that society considers literacy a right and not a privilege. Reading is a human invention that co-opts parts of the brain originally designed for other tasks, such as spoken language. In order to learn how to read, most young children require careful and highly technical instruction. If they are learning to read an alphabetic writing system, such as English, they must master the *alphabetic principle*—that letters stand for phonemes which in turn combine to form words (National Reading Panel, 2000). Alphabetic writing systems vary in the extent to which phonemes and letters correspond to each other. In English this correspondence is irregular for many words that children are expected to read, a significant barrier for some children. Children must also learn to attach meaning to what they have decoded (Kintsch, 1998; Rapp & van den Broek, 2005). This places demands on their oral language, world knowledge, and other cognitive systems such as working memory and attention. The aim of this chapter is to review research on primary grade reading instruction in the United States.

For this chapter, we define reading as the ability to decode words on a page and attach meaning to these words and the sentences they form in such a way as to form a coherent mental representation of the text. Building a coherent mental representation depends on students' ability to accurately and fluently decode words, to infer meaning using other information in the text and general world knowledge, to monitor their understanding of the text, and to use metacognitive and other strategies to build meaning. Comprehension also depends on the difficulty of the text and the students' perceptions of why they are reading and how important it is to understand the text (Snow, 2001). The chapter focuses on four topics on the current state of knowledge about effective primary reading instruction: (1) reading instruction in the United States and the policies and reports that shaped that instruction; (2) decoding interventions; (3) attaching meaning to text and comprehension instruction; and (4) individualizing (or differentiating) reading instruction to improve students' decoding and comprehension. We conclude with a summary and suggestions for future research, policy, and practice for a broader audience than just US schools.

Current State of Knowledge about Effective Primary Reading Instruction

Reading Instruction in the United States: Reports and Policies that Influence Classroom Reading Instruction

In the United States and many nations, most children are taught how to read during the early primary grades, kindergarten through second or third grade, or from 5 or 6 years to 8 or 9 years of age. There is evidence that children who are not reading proficiently by the end of second grade (about age 8 years) are much less likely to gain proficiency (Spira, Bracken, & Fischel, 2005), that consistently high-quality reading instruction from first through third grades provides stronger student reading outcomes (Connor et al., 2013), and that early effective instruction has a lasting impact on reading outcomes (Konstantopoulos & Chung, 2011). The motivation for much of the research described here is the persistent achievement gap in literacy between children from families of low and high socioeconomic status (SES); between children from the majority racial/ethnic group in the United States and minority peers; and between children for whom English is their first language and children for whom English is their second. The latter are often called English language learners (ELLs) (Chatterji, 2006; Duncan et al., 2007).

Two seminal reports published in 2000 and 2001, as well as state and federal policies, provide information about how reading is taught in schools throughout the United States. In 2001 2000, the National Reading Panel report (2000) was written with the purpose of reviewing the extant research and coming to consensus on the most effective ways to teach reading in the primary grades. The topics included in the report were those for which there was sufficient research. The report stated that effective approaches to teaching reading included: (1) explicit instruction in phonological awareness; (2) systematic phonics instruction; (3) methods to improve fluency; and (4) ways to improve comprehension, including instruction in vocabulary and strategies (www.nichd.nih.gov/research/supported/Pages/nrp.aspx#overview).

The second seminal report, the RAND report on *Reading for Understanding* (Snow, 2001), provided a heuristic for reading comprehension that is widely used today. It describes reading comprehension as the active extraction and construction of meaning from text, which is influenced by the reader's purpose and motivation for understanding the text as well as aspects of the text itself including difficulty, content, and genre. All of this is influenced by sociocultural contexts. Hence, effective instruction considers all aspects of the process of reading for understanding. Both of these reports—but particularly the National Reading Panel Report—influenced federal and state policies, such as Reading First (Gamse, Jacob, Horst, Boulay, & Unlu, 2008). This offered financial incentives to states to provide evidence-based reading instruction, assessment of student progress, and support for struggling readers.

From 2005, new policies that focused on preventing reading difficulties led to widespread adoption of multitiered systems of support or response to intervention (RTI) within the United States. Response to Intervention involves general education classroom teachers providing effective and high-quality reading instruction to all children as a foundation for learning. This foundational classroom instruction is known as Tier 1 instruction. If students fail to progress or respond to high-quality general Tier 1 instruction, the general education teacher (or other specialist) can provide more targeted and intensive intervention. A key aspect of multitier systems includes assessment. First, teachers use universal screening assessments, where all children are assessed in order to learn which children are and are not performing on grade level. These children then receive further diagnostic assessments, as well as ongoing assessment, to monitor their growth in reading and to determine whether they need additional supplemental interventions that increase in intensity (e.g., smaller group size, more frequent, more sessions, more individualized skills) (Connor, Alberto, Compton, & O'Connor, 2014). Supplemental interventions that increase in intensity are known as Tier 2 and Tier 3. In general, Tier 2 and Tier 3 interventions are provided to small groups of students with similar learning needs. It is expected that Tier 3 be more intensive than Tier 2. All fifty states in the United States now encourage RTI for prevention purposes, and a growing number of states allow it for identification of learning disabilities (Fuchs & Vaughn, 2012).

The United States does not have a national curriculum. Instead, each of the fifty states sets standards for what students are expected to learn, which vary widely from state to state. In an effort to make the English language arts (reading, writing, listening, and speaking) expectations for students more consistent throughout the United States, state departments of education, policymakers, and other stakeholders worked to develop the Common Core State Standards (www.corestandards.org), which were adopted and implemented by many states beginning in 2014. The standards set expectations for English language arts and literacy at each grade level from kindergarten through fifth grade and into middle and high school. For example, the reading standards focus on understanding key ideas and details, understanding text structure, and integrating knowledge and ideas when reading (see also Goldman & Snow, this volume).

In the rest of this chapter, we review recent research on the components of reading instruction defined by these reports and influenced by these policies. We start with code-focused instruction, then discuss comprehension instruction, including general instructional strategies, and then differentiated or individualized instruction in reading. As the reader will note, findings to date from RTI studies relate more to struggling students and code-focused skills, whereas comprehension instruction is relevant not only to struggling readers but also to the broader student population. Finally, we summarize the findings and implications.

Cracking the Code: Code-Focused Instruction and Multitiered Systems of Support

Although most of the recent research on code-focused interventions has targeted children who do not make adequate gains in reading skills even when they receive classroom instruction that is generally effective for their peers, research reveals that virtually all students require at least some explicit and systematic instruction in the alphabetic principle and phonics to learn to read and that some children require more than others (National Reading Panel, 2000).

There is a substantial body of evidence regarding the effectiveness of early reading interventions provided to students with reading difficulties in the primary grades (e.g., Benner, Nelson, Ralston, & Mooney, 2010; Cavanaugh, Kim, Wanzek, & Vaughn, 2004; Ehri, Nunes, Stahl, & Willows, 2001; National Reading Panel, 2000; Wanzek & Vaughn, 2007). Across these syntheses, small group explicit and systematic phonics and phonological awareness instruction was found to be highly effective in improving word reading. Standard treatment protocols, where teachers follow fairly scripted and consistent protocols for all students, were used in these studies. To date, a majority of these supplemental interventions have been implemented by researchers. However, interventions also appeared to be effective when administered by certified teachers or well-trained and supervised paraprofessionals or volunteers.

A recent synthesis (Connor et al., 2014) reviewed the research literature about students with reading disabilities or at risk of such disabilities because of such things as a family history of reading difficulties or weak language skills. The synthesis reviewed research in four areas: (1) assessment—universal screening, progress monitoring, assessment of ELL, and accommodations for students with disabilities; (2) contributions of basic cognitive processes to reading; (3) intervention—increasing intensity of instruction, improving fluency and preschool language, and promoting peer-assisted or collaborative learning; and (4) professional development—developing specialized knowledge and combining multiple strategies. The consensus was that the field has advanced in each of these areas. For example, research suggests that universal screening of all students followed by targeted diagnostic assessment for students who fail the screening is a reliable way to identify students who are likely to require more intense and targeted reading instruction. The reader is referred to the report for additional recommendations.

The *Institute for Education Sciences Practice Guide for RTI* (Gersten et al., 2008) reviewed the literature supporting multitier interventions. They made five recommendations based on evidence and expert knowledge that included (1) conducting universal screening, (2) providing a high-quality differentiated Tier 1 core reading program, (3) conducting frequent progress monitoring to assess growth of reading skills, (4) providing increasingly intensive tiers of intervention, and (5) ensuring that interventions are implemented with fidelity. These five core components have rapidly manifested in state laws or guidelines about RTI (Berkeley, Bender, Gregg Peaster, & Saunders, 2009; Zirkel & Thomas, 2010).

In one highly cited synthesis, Wanzek and Vaughn (2007) reviewed reading interventions that were extensive, which they defined as those occurring for over 100 sessions (Tier 2 & 3). These interventions were provided to students with or at risk for reading disabilities and led to higher reading achievement scores. In general, interventions with smaller group sizes were more effective than those with larger groupings (of three to eight). There was also evidence that stronger effects were found for kindergarten and first grade students than for second and third grade students.

Although most students benefit from supplemental intervention, 3% to 7% may need even more intensive intervention (e.g., Al Otaiba & Torgesen, 2007). Considerably less is known about what gains are possible for this small set of children who do not show adequate progress despite receiving well-implemented classroom instruction and supplemental interventions. To our knowledge, nine experimental or quasi-experimental studies have examined children who received the most intensive interventions (Al Otaiba et al., in press; Beach & O'Connor, 2013; Denton, Fletcher, Anthony, & Francis, 2006; Denton et al., 2013; Gilbert et al., 2013; Vaughn, Wanzek, Linan-Thompson, & Murray, 2007; Vellutino, Scanlon, Zhang, & Schatschneider, 2008). The findings suggest that there were large individual differences in the progress students made. It appeared to be easier to significantly impact word reading skills than fluency or comprehension. With regard to intervention components, only a small number of studies examined the effectiveness of classroom reading instruction (Hill, King, Lemons, & Partanen, 2012), and only three studies provided increasingly intensive intervention within a study year (e.g., Beach & O'Connor, 2013). A more common approach was to conduct a Tier 2 intervention for a year and then offer more intensive intervention the following year to those students who did not demonstrate adequate growth.

There is general concern that students with the weakest skills might be required to wait to receive the most intensive interventions, particularly in light of a recent literature review about the characteristics of children who continue to have difficulty learning to read with less intensive interventions (Lam & McMaster, 2014). Lam and McMaster extended prior syntheses describing responsiveness to multitier interventions (e.g., Tran, Sanchez, Arellano, & Swanson, 2011). They reported that students' initial word identification, understanding of letter–sound associations, phonemic awareness, and oral text reading fluency predicted their responsiveness to Tier 2 and Tier 3 interventions.

Our own concern about students with the weakest skills led us to explore two different RTI models—typical RTI and dynamic RTI (Al Otaiba et al., 2014). In typical RTI we conducted universal screening to assess initial reading skills of first-grade students and then students received Tier 1 intervention for eight weeks. If struggling students did not catch up to their peers, they received more intervention that increased in intensity across first grade. In contrast, dynamic RTI students received Tier 2 or Tier 3 interventions immediately according to their initial screening results. Thus students with the lowest skills could begin Tier 3 intervention at the beginning of the school year as soon as they were identified. A total of thirty-four first-grade classrooms and 522 first-grade students participated. The students attended ten socioeconomically and culturally diverse schools. The small group interventions were identical across conditions except for when intervention began. Reading assessments included letter-sound, word, and passage reading and teacher-reported severity of reading difficulties. Students in dynamic RTI showed an immediate achievement advantage compared with typical RTI, and effects accumulated across the year. Importantly, students in the dynamic condition who received Tier 2 and Tier 3 intervention ended the study with significantly higher reading scores than students in the typical RTI condition.

Two previous longitudinal follow-up studies of students who received RTI in first grade (Gilbert et al., 2013; Vellutino, Scanlon, Small, & Fanuele, 2006) had reported substantial increases in the proportion of students with reading skills falling below a standard score of 90 (typical mean = 100, *SD* = 15) from second to third grade. Specifically, Gilbert et al. reported that 46% of students who received Tier 3 in first grade had word reading and comprehension skills below a standard score of 90 by third grade. Vellutino et al. found that roughly a third of students who were difficult to remediate had basic reading skills below a standard score of 90 at this point. The findings of Al Otaiba et al. (2014) were somewhat more encouraging. They classified children as never at risk, less difficult to remediate, or requiring sustained intervention based on first-grade RTI. Even among those children who had required sustained intervention, by the end of second and third grade only 8.7% and 7.9%, respectively, scored below a standard score of 90 in word reading. By the end of third grade, none of the students who were less difficult to remediate in first grade had word reading scores below a standard score of 90.

In summary, emerging evidence suggests that RTI can improve students' decoding and word reading and reduce the percentage of students with weak reading skills. However, students who demonstrate inadequate response are likely to need ongoing intensive interventions. Although reading comprehension should be a key part of multitiered systems of instruction, with some exceptions (e.g., Al Otaiba et al., 2014), most of the research on Tier 2 and Tier 3 interventions has focused on code-focused skills (largely due to the needs of struggling students). New and unpublished research by the Reading for Understanding network researchers suggests that Tier 2 and Tier 3 interventions focused on oral language and comprehension are also effective in improving reading for understanding (http://ies.ed.gov/ncer/projects/program.asp?ProgID=62). Thus, incorporating reading comprehension interventions into multitiered systems of instruction may further improve students' reading proficiency.

Attaching Meaning to Text: Reading Comprehension Instruction

Teaching children how to attach meaning to what they have read has been more difficult than anticipated. At one point, it was assumed that once decoding issues were resolved, comprehension would improve as students became more fluent readers (e.g., Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001). However, as the well-supported simple view of reading holds (Hoover & Gough, 1990), proficient reading comprehension is the product of readers' ability to fluently decode text and their oral language skills (listening comprehension). If either is weak, then reading comprehension is likely to be weak as well. Many students do not have adequate language abilities, limiting their comprehension even when decoding skills are adequate. Several studies have identified students who have adequate decoding skills but difficulties with comprehension (Compton, Fuchs, Fuchs, Elleman, & Gilbert, 2008; Oakhill & Yuill, 1996). A review by García and Cain (2014) revealed that across 110 studies, decoding and reading comprehension had a correlation of about .74. This is lower than might have been expected and suggests that some children might have strong decoding skills but weaker comprehension skills. The association between decoding and comprehension was moderated by students' age (decoding was more highly correlated with comprehension for readers aged 10 years and younger) and, importantly, oral language skills (children with weaker listening comprehension generally had weaker reading comprehension).

There are recent reviews of research focusing on reading for understanding and reading comprehension (Block, Parris, Reed, Whiteley, & Cleveland, 2009; Shanahan et al., 2010; Swanson et al., 2011) as well as how to meet the needs of students for whom English is the second language (Melby-Lervåg & Lervåg, 2014; Slavin & Cheung, 2005) and students with or at risk of reading disabilities, including students living in poverty (Benner et al., 2010; Berkeley, Scruggs, & Mastropieri, 2010; Weiser & Mathes, 2011). We provide an overview of these reviews and syntheses as well as more recently published findings (see Oakhill & Berenhaus, this volume, for more discussion of comprehension).

The IES Practice Guide *Improving Reading Comprehension in Kindergarten Through 3rd Grade* (Shanahan et al., 2010) provides a practical review of the literature on reading comprehension and offers five recommendations. These recommendations have different levels of evidence. Strong evidence requires consistent findings across studies for a wide range of populations. Moderate evidence requires randomized controlled trials (RCTs), but findings may not generalize to wide populations. Minimal evidence is primarily correlational either because no RCT has been conducted or because it would not be practical or ethical to conduct an RCT. The authors based recommendations for second and third graders on studies of reading comprehension and recommendations for kindergarteners and first graders on studies of both listening and reading comprehension.

The first recommendation was to teach students how to use six comprehension strategies (strong evidence). These strategies included activating prior knowledge and predicting; developing and answering questions; visualizing the story action; monitoring and repairing comprehension; drawing inferences; and summarizing and retelling. These strategies could be taught individually or together. The panel also recommended that practitioners gradually reduce the support they offer to students for using the strategies.

The panel's second recommendation was to "teach students to identify and use the text's organizational structure to comprehend, learn, and remember content" (p. 17, moderate evidence). When children learn how narrative and expository texts are organized (e.g., narrative stories have a beginning, middle, and end), they are better able to comprehend what they read. Instruction about text structure can be begin as early as kindergarten and continue through the elementary grades and beyond (Swanson et al., 2011). For narrative text, instruction about characters, setting, goal, problem, plot, resolution, and themes appeared to aid students' comprehension.

Guiding students through focused, high-quality discussion on the meaning of text was the third recommendation (see also Goldman & Snow, this volume). The panel encouraged teachers to structure discussion to complement the text, the instructional purpose, and the students' reading ability and grade level; to develop discussion questions that ask students to think deeply about text; to ask follow-up questions to encourage and facilitate discussion; and to have students lead structured discussions in small groups. Although the evidence was deemed to be minimal at the writing of the practice guide, a more recent meta-analysis (Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009) provides evidence that discussion can improve students' comprehension. However, there are some caveats. One caveat is that most of the research was conducted with students in fourth grade or beyond. Another is that although both teacher and student talk increased and improvements in text comprehension were associated with discussion, there was little or no effect on students' literal or inferential comprehension and critical thinking and reasoning. Carlisle, Dwyer, and Learned (2013) also showed that discussion can be used to help students learn to reason, read, and write analytically. These are key skills that students are expected to master as part of the Common Core State Standards.

The practice guide panel (Shanahan et al., 2010) found minimal evidence for the fourth recommendation: select texts purposefully to support comprehension development. Nevertheless, they felt that it was an important recommendation in the context of the other recommendations and accumulating evidence on individual differences in learning to read, including the types of text that are accessible and interesting to different students (Snow, 2001). The panel encouraged teachers to use multiple genres of texts, texts of high quality, texts in line with students' reading ability, and texts that support the purpose of the instruction. Other work (Hiebert & Fisher, 2007; McNamara, 2013) has highlighted the importance of considering text complexity and accessibility to readers with different reading skills and academic knowledge.

The panel's fifth and final recommendation was to establish an engaging and motivating context in which to teach reading comprehension (moderate evidence). This recommendation was based on research showing that students who actively engage with a text are more likely to understand its meaning and that students are more likely to be actively engaged if they understand the purpose for reading and are motivated to make sense of the text (Guthrie, Anderson, Aloa, & Rinehart, 1999; McNamara, 2013). The panel suggested a number of teaching strategies, while noting that there is little strong evidence for the practices. These practices include conveying the purpose of the lesson, explaining how using comprehension strategies will help students learn, providing students with choices, and allowing classmates to collaborate with each other.

These five recommendations continue to be relevant in light of new research findings, which we have embedded in this description of the practice guide recommendations. Moreover, a meta-analysis (Berkeley et al., 2010) reveals that these strategies are also effective for students with reading disabilities, with a mean effect size (*d*) of .70. The meta-analysis found that, in particular, code-focused reading instruction and question/strategy instruction supported stronger reading comprehension skills for children with reading disabilities. Another meta-analysis of 24 studies revealed that effective reading instruction strategies are also effective for students with behavioral disorders (Benner et al., 2010). A common thread across effective interventions is that students were taught to attend to what they were reading more carefully and to think more systematically about text.

STUDENTS WHO ARE ENGLISH LANGUAGE LEARNERS

The number of students in the United States who speak a language other than English at home, or ELL students, continues to increase. In 1979, approximately 9% of US students were ELLs. This increased to 17% in 2005 (Slavin & Cheung, 2005). By 2008, 21% or almost 11 million students were ELLs (Melby-Lervåg & Lervåg, 2014). Students who are not proficient in English are more likely to drop out of school and many have weaker educational outcomes than their English-proficient and English-monolingual peers (Melby-Lervåg & Lervåg, 2014).

Slavin and Cheung (1999) reviewed research on reading instruction for ELL students. They compared English-only immersion programs, where ELLs are expected to learn English and their native language plays a minor (or no) role in instruction, and bilingual education, where students spend a substantial part of the school day using their native language, particularly in content areas such as social studies and science. Their most significant finding was that there were "far too few high-quality studies" (p. 273), and they called for longitudinal and randomized controlled studies. Across the seventeen studies they reviewed, bilingual education appeared to be more effective than English-only immersion programs in 70% of the studies (effect size $d = .33$). The remaining studies found no advantage for one type of program over the other. Notably, none of the studies found an advantage for English-only programs. The authors suggested that teaching children to read in their native language might offer a bridge to reading English. Of note, the bilingual programs studied in the 1991 meta-analysis were not those typically used in schools in the United States today. The programs described in the 1991 meta-analysis tended to be bilingual programs where English and the native language were taught at different times during the school day. In contrast, bilingual programs today tend to involve bilingual teachers or provide teacher aides who are proficient in the children's native language throughout the school day, or a separate program, rather than half of the day. Bilingual programs are far outnumbered today by English immersion programs, with some states essentially banning bilingual programs. Outside the United States this controversy extends to mother tongue instruction, particularly in developing countries where the mother tongue, in some instances, does not have a written form.

How do ELL students' literacy skills compare with those of their monolingual classmates? Melby-Lervåg and Lervåg (2014) carried out a meta-analysis of studies conducted in the United States and Canada that compared ELL students with monolingual peers on reading comprehension skills.[1] Across all eighty-two studies, ELL students had weaker reading comprehension skills than monolingual students ($d = .62$). However, this effect was moderated by students' language comprehension and decoding skills. Students' SES did not directly affect reading differences between ELL and monolingual peers. The researchers also observed that monolingual students had stronger decoding and phonological skills than did ELL students. This difference was smaller for comprehension than decoding and smaller for students in Canada than in the United States.

Taken together, there is clear evidence of an achievement gap in reading comprehension for ELL students compared with their monolingual peers. There has been and continues to be controversy in the United States regarding how to teach reading in English to students who may not be highly proficient in the language they are expected to read. Some states' policies do not align with emerging and established research (see Jared, this volume, for discussion of literacy development for students who are bilingual). For example, although paired bilingual programs, where both languages are taught but at different parts of the day, are generally more effective in closing the achievement gap than English-only programs, a number of US states now require English-only immersion programs by law.

RECENT STUDIES ON COMPREHENSION INSTRUCTION

Our search of the literature on reading comprehension revealed ten studies published since 2009 that used experiments or well-designed quasi-experimental studies to investigate effective methods for teaching reading comprehension that, in our opinion, provide insights beyond those provided in the other meta-analyses.

Reading Comprehension Interventions

Three of the studies focused on improving students' understanding of expository or informational text (Guthrie et al., 2009; Wijekumar, Meyer, & Lei, 2012; Williams, Stafford, Lauer, Hall, & Pollini, 2009). Guthrie and colleagues tested the efficacy of Concept-Oriented Reading Instruction (CORI) using a quasi-experiment with 156 fifth graders who were identified as low or high achieving. In CORI, reading comprehension is taught with an emphasis on inferencing and comprehension monitoring. Lessons include activities designed to improve and sustain motivation. Instruction was for 90 minutes per day for 12 weeks. The teacher taught lessons 3 days per week and a reading specialist taught 2 days per week. Compared with students who received traditional instruction, students participating in CORI generally exhibited higher scores on a reading comprehension test, a test of word reading, a test of fluency, and a test on the content area (science). There were no differences in effectiveness for high- versus low-achieving students. There was a marginal difference in performance on the test of inferencing, with higher-achieving students receiving CORI outperforming lower-achieving students, as might be expected. There was no treatment effect for motivation. Goldman and Snow (this volume) provide additional information about this intervention.

Williams et al. (2009) examined the efficacy of specific instruction on text structure of expository text on second graders' (215 students and 15 teachers) comprehension in an experiment. Using an animal encyclopedia, trade books, and researcher-developed texts, teachers taught students how to use clue words, graphic organizers (for example using a circle in the middle for the main idea and radiating lines to provide supporting details), summarizing, and compare-contrast strategies while focusing on vocabulary development and close analysis of text. There was also review at the end of each lesson. Students in the control condition received science instruction on the same content but no instruction on text structure and reading expository text. There were twelve 45-minute lessons taught in twenty-two sessions over the course of two months. There were significant effects of treatment on the researcher-developed assessments, suggesting that explicit comprehension instruction can be effective with second graders (much of the previous research was with older students) and that such instruction can be accommodated within the context of science instruction without jeopardizing students' learning of the content. A weakness of the study is that standardized assessments of reading comprehension were not administered after the intervention, so the effect of the intervention on standardized assessments is unavailable.

Wijekumar et al. (2012) assessed the efficacy of teaching text structure using an intelligent tutoring system (ITSS) to improve the reading comprehension of fourth graders from 131 classrooms. They stated that "structure strategy is designed to help readers use signals for text structures (i.e., clue words) in nonfiction (i.e., informational or expository text) to create strategically organized and efficient mental representations and use that knowledge to apply their memory of the text when needed" (p. 989). The signaling words then help students identify one of five different text structures (i.e., compare/contrast, problem/solution, cause/effect, sequence, description). The ITSS system was designed to only partially replace teacher-provided language arts instruction. It was based on the idea that technology can supplement and enhance teaching of the formal or core curriculum designated by the school or district. The authors note that the technology offers a number of features to support learning including consistency, practice, assessment, and feedback. The ITSS was implemented in a computer lab 30 to 45 minutes each week for 6 months. Students in the ITSS condition achieved significantly higher scores on a standardized reading comprehension test than students in the control condition, but the difference was modest (d = .10). Effect sizes were larger for researcher-developed assessments, with the largest effect (d = .49) on a task where students were asked to provide the main idea of a recently read passage. The size of the effect did not differ for students judged to be lower or medium/higher achieving. This result suggests that the technology supported students with a range of literacy skills.

These three studies together demonstrate the efficacy of multicomponent instructional interventions for improving students' content knowledge, including the ability to read and learn from informational text. They also underscore the importance of strategy instruction, including knowledge about text structure and graphic organizers, for supporting students' comprehension of expository and informational text. At the same time, all three instructional interventions were intensive and, at least to some extent, focused on meeting the learning needs of individual students either by allowing students to choose instructional materials (Guthrie et al., 2009) or through the use of technology (Wijekumar et al., 2012). Also, part of each of these interventions focused on building students' understanding of texts through discussion.

However, not all reading comprehension interventions that include strategy instruction are effective. James-Burdumy et al. (2012) conducted a multisite randomized controlled large-scale study of four different supplemental reading comprehension interventions with fourth and fifth graders. This massive 2-year study tested four interventions that were designed to supplement rather than replace regular instruction in reading, science, or social studies and for which there was some evidence of efficacy. These interventions were Project CRISS, ReadAbout, Read for Real, and Reading for Knowledge, which are all commercially available. All four interventions used what the authors called "explicit comprehension instruction" (p. 347), which included the explicit teaching of strategies. To varying degrees, teachers modeled the use of each strategy (e.g., summarizing was used in all four programs) and guided student practice. There were differences as well among the four interventions. For example, ReadAbout used computers and students received extensive and immediate feedback. Interventions were provided according to publisher guidelines for 2 years. Results after the first year revealed no significant positive effects of the interventions, and Reading for Knowledge had a significant negative effect. For the second year, Reading for Knowledge was dropped because schools did not want to use it. In the second year only ReadAbout had a positive treatment effect, but only on the social studies assessment. The authors note that ReadAbout, which was computer-based, was the only intervention that provided immediate feedback to students. These results suggest that reading strategy instruction alone is unlikely to make educationally important differences in students' reading comprehension. Another implication is the need for future research into technology use (e.g., e-readers, e-books that use text-to-speech technology) for students who do not read well enough to complete typical instruction in areas such as science and social studies.

One possible explanation for these findings is that teaching just reading comprehension strategies is not enough. Supporting this idea, McKeown, Beck, and Blake (2009) conducted a randomized controlled trials design study to examine the efficacy of three different approaches to building fifth graders' reading comprehension. The study was carried out with 116 students who were randomly assigned to classrooms. It focused on building content understanding, teaching comprehension strategies, and using core literacy curriculum materials, which the authors called the basal condition. The content approach used discussion at key points in the text. The strategy approach taught students specific strategies, including comprehension monitoring, summarizing, and predicting. These were prompted by the teacher at designated points in the text (similar to the interventions in the James-Burdumy et al., 2012, study described previously). The basal approach used the parts of the teachers' core curriculum that focused on comprehension, but did not use the parts that focused on word reading and other code-focused activities. The results showed improved comprehension in all three groups, with students in the content group making greater gains in comprehension than the students in other groups. The authors noted that "getting students to actively build meaning while reading does not necessitate knowledge of and focus on specific strategies but rather it may require attention to text content in ways that promote attending to important ideas and establishing connections between them" (p. 245).

We turn now to two additional studies that also used multiple strategies to improve students' reading comprehension. Block and colleagues (2009) assessed the efficacy of six different widely used strategies to teach comprehension to 660 second- through sixth-grade students from thirty classrooms in five schools. These included workbook practice on specific strategies, individualized schema-based learning (i.e., teacher-monitored silent reading followed by discussion), strategy instruction and practice during reading (called situated practice), conceptual learning, and transactional learning (discussion of material read individually and silently). These were provided to different classrooms in different random orders. These were contrasted with traditional instruction. The interventions were supplemental and designed to increase reading comprehension instruction by twenty minutes per day. All students, regardless of achievement level, received the instruction. There were treatment effects on standardized measures of reading comprehension and vocabulary. Regardless of which order of the six strategies was used, students had higher reading comprehension scores than students in the control group. Comparisons of the six instructional strategies suggested that transactional learning led to stronger summarizing skills, conceptual learning contributed to stronger grasp of the main idea, and schema-based learning led to stronger recalling of detail. Situated learning and workbook practice were not as effective as the other types of comprehension instruction.

Clarke, Snowling, Truelove, and Hulme (2010) compared three different reading comprehension interventions—a text comprehension intervention, an oral language training intervention, and a combination of the two—with a business-as-usual control. The study was conducted with eighty-four fourth graders who had reading comprehension difficulties. Interventions lasted 20 weeks for 30 minutes per session. The text comprehension intervention focused on teaching comprehension strategies (very similar to the other interventions previously discussed). The oral language intervention focused on improving students' listening comprehension and vocabulary. The combined intervention used both written and oral language strategies and lasted the same amount of time as the other two interventions. Students in the three intervention groups performed better on standardized assessments of reading comprehension and oral language than students in the business-as-usual control group. For the oral language and combined interventions, improvements in vocabulary mediated gains in reading comprehension. That is, comprehension gains were larger for students in these two groups because their oral language skills were stronger than those of students in the text comprehension intervention. The text comprehension intervention did not have an effect on vocabulary compared with the control. Long-term gains were greatest for students receiving the oral language intervention.

Across these studies, mixed results of effectiveness suggest that more nuanced research on comprehension instruction and intervention is needed before we fully understand the components of comprehension instruction that are effective and for whom they are effective. Further, it is important to learn more about how intervention effects are moderated by some of the variables we have described. In general, multicomponent interventions that included a focus on developing oral language (see also Goldman & Snow, this volume) and strategy instruction were more effective than single-strategy interventions. Corrective feedback—either explicit or implicit—supported students' gains in comprehension. Some interventions were designed for students with weak comprehension skills and so could easily be used in multitiered systems of instruction. Those studies that included both higher- and lower-achieving students generally demonstrated significant treatment effects (if there were treatment effects) regardless of achievement level. However, virtually all were small group interventions where teachers are more likely to be sensitive to students' individual learning needs.

Individualized Reading Instruction to Improve Students' Decoding and Reading Comprehension

Two different lines of research have focused on implementing reading instruction that accommodates individual student differences. The first (Reis, McCoach, Little, Muller, & Kaniskan, 2011) uses a school-wide enrichment model of reading (SEM-R) that is designed to increase interest and engagement in reading and espouses models from the gifted student literature and many constructivist principles (Dahl & Freppon, 1995). These principles, with their roots in the research of Piaget (1960), hold that children generate their own knowledge from meaningful experiences that interact with their ideas. The authors describe SEM-R as "an enrichment-based reading program designed to stimulate interest in and enjoyment of reading, leading to higher reading achievement, by enabling students to self-select and read high-interest books of personal choice that are slightly to moderately above current reading instructional levels independently with differentiated instruction provided in weekly teacher conferences" (p. 464). Teachers use assessment data to "respond to differences in student's readiness, interests, and learning profiles."

Reis and colleagues investigated the efficacy of SEM-R in an experiment with 1,192 second through fifth graders in five schools. Their results showed no overall effect of SEM-R across schools. Treatment effects across all schools ranged from −.11 to .27 across schools, and only one school showed a significant effect of treatment. Results were similar for oral reading fluency. Teachers reported that students in SEM-R generally enjoyed and were more engaged in reading and attributed these differences to SEM-R implementation. It is not clear why the program worked at one school and not others, and why the study failed to provide strong support for differentiating reading instruction using personal choice and weekly teacher conferences.

The second line of research on differentiated reading instruction provides strong evidence for the efficacy of individualizing literacy instruction (Connor et al., 2013; Connor, Morrison, Fishman, et al., 2011). This research follows bioecological and transactional theories of development (Bronfenbrenner & Morris, 2006; Morrison & Connor, 2009). These theories hold that there are multiple sources of influence on individual students, which act reciprocally. Sources of influence are hypothesized to move from child characteristics (e.g., genetics, temperament, aptitude), which are influenced by proximal (e.g., instruction, parents) and more distal sources of influence (state education policy). Generally, more distal influences operate through more proximal sources of influence. The intervention was also informed by cognitive development approaches to reading instruction (Morrison & Connor, 2002), including the simple view of reading (Hoover & Gough, 1990).

This instructional framework, called Individualized Student Instruction in Reading (ISI-R), relies on using teacher-led small group instruction, called flexible learning groups, and assessment to guide instruction. A key aspect of the intervention is Assessment-to-Instruction technology (A2i). The A2i algorithms compute recommended weekly amounts of four types of reading instruction—teacher/student-managed (teachers and students both participate in the learning opportunity) or student/peer-managed (students work independently or with peers), and code- or meaning-focused instruction following the simple view of reading (see Table 27.1 for examples of each type of activity). It used three types of valid and reliable assessments: word reading or decoding; comprehension, and vocabulary or word knowledge. The language arts materials already used by the classroom teachers are indexed to the four types of instruction. Hence, teachers use materials with which they are familiar but use them in different ways to meet their students' individual learning needs. Another component of A2i is teacher professional development, which supports implementation in the classroom.

The algorithms used in A2i to compute the recommended amounts and types of instruction were developed essentially by reverse-engineering the multilevel models that predicted students' end-of-year reading outcomes using the four types of instruction and the three fall assessment scores (Connor, Morrison, & Katch, 2004; Connor, Morrison, Schatschneider, et al., 2011 Connor, Morrison, &- Underwood, 2007). Based on students' spring reading scores, a target outcome is set, which is at least grade level reading by the end of the grade. This is used with the three test scores to compute the recommended amounts of each of the four types of instruction in minutes per day or week. The recommendations are provided by the A2i software along with progress monitoring charts, planning tools, and professional development resources.

In seven experiments with two different school districts, Connor and colleagues demonstrated the efficacy of ISI-R in kindergarten (Al Otaiba et al., 2011) through third grade for both decoding and reading comprehension (Connor et al., 2013). In the first study (Connor, Morrison, Fishman, et al., 2011), thirty-three teachers in eight schools and their 448 students were randomly assigned to ISI-R or a vocabulary intervention (Beck, McKeown, & Kucan, 2002), which was presented in the same way to all students. About half of the students were from low-SES families. Teachers learned to provide the A2i recommended amounts using small flexible learning groups. Professional development focused on classroom management, using assessment to guide instruction, and implementing research-based reading instructional activities effectively. Using classroom observation, the investigators demonstrated that teachers in the ISI-R condition were more likely to provide instruction that considered students' individual learning differences than were control teachers. Students in the ISI-R classrooms performed better on standardized tests of reading comprehension than students in the vocabulary control classrooms (d = .20). Another study with first graders revealed a significant treatment effect for word reading (Connor, Morrison, Schatschneider, et al., 2011). Finally, the closer the observed amounts of each type of instruction were to the A2i recommended amounts, the greater were students' comprehension gains. The investigators noted that the association between students' profile of language and literacy skills and recommended instruction was nonlinear and more complex than anticipated, which helps to explain why

Table 27.1 Examples of Teacher/Child-Managed and Child/Peer- Managed Code- and Meaning-Focused Instruction.

	Teacher/Child-Managed	Child/Peer-Managed
Code-focused	The teacher and a small group of students are working on phonological awareness activities. Students are changing words (e.g., "hat") to new words by changing one phoneme (e.g., /k/ for /h/ to produce "cat").	Students are at the computer center, each on a computer and using a phonics software program.
Meaning-focused	The teacher is reading the book *Stone Soup*[1] to the class. She stops and asks, "Why do you think the townspeople help make the stone soup?" The class then discusses the motives of the main character and the townspeople.	Students are reading books they have chosen quietly at their desks. Others are working together to write the class weekly newsletter.

[1] *Stone Soup* is a classic story where a wandering peddler (or soldier) starts to make stone soup using water and a stone. The villagers think about other ingredients that would make the stone soup taste better and so begin to bring meat and vegetables to add to the soup. Thus the peddler tricks the villagers into sharing their food with him.

teachers frequently have difficulty using assessment results to guide instruction (Roehrig, Duggar, Moats, Glover, & Mincey, 2008).

In another experiment, Connor et al. (2013) investigated whether the effects of ISI-R might accumulate from first through third grade by conducting a longitudinal study where first-grade teachers and their students were randomly assigned to ISI-R or a control math intervention. Students were followed into second grade and their second-grade teachers were randomly assigned to condition. Then, students were followed into third grade and their third grade teachers were randomly assigned to condition. In this study, which involved 95 teachers and 882 students, more than 45% of students were from low-SES homes. Results showed significant positive effects of ISI-R on a reading score comprising standardized word reading and reading comprehension assessments. Results also showed that ISI-R effects accumulated. Students who participated in ISI-R classrooms in all three grades made greater gains than students who were in ISI-R for fewer years or who were in control classrooms all three years (d = .73). Notably, participating in ISI-R in first grade appeared to be necessary but not sufficient for students to have higher reading scores than control students who participated in a math intervention all three years. For example, students who participated in ISI-R in first and second grades had significantly greater reading scores at the end of third grade than those in the control group. However, students who were in ISI-R classrooms in second and third grades fared no better than control students.

The differences in approaches for SEM-R and ISI-R, which are both designed to support teachers' efforts to provide differentiated reading instruction, speak to the contrasting theories that continue to inform reading instruction. These include constructivist principles, which are epitomized in the SEM-R intervention, and bioecological and transactional theories (Bronfenbrenner & Morris, 2006), which, along with cognitive theories of reading, informed the development of ISI-R. In SEM-R, instruction was differentiated through providing student choice, encouraging students to construct their own knowledge, and attending to students' individual learning styles. In ISI-R, the focus was on the content of the instruction (code- or meaning-focused), the dose of particular types of instruction including explicit instruction provided by the teacher and opportunities for independent and peer learning opportunities, and careful use of assessment to determine both dosage and challenge.

Until the writing of this chapter, neither line of research cited the other (mea culpa—although we did cite McCoach, O'Connell, Reis, & Levitt, 2006, in one of our papers). Indeed, Reis et al. claimed that there was little experimental evidence for differentiating instruction, although Connor and colleagues published their first experiment in 2007 (Connor, Morrison, Fishman, Schatschneider, & Underwood, 2007). We hope to see the junction of these two different lines of research become an example of consilience (Wilson, 1998), where competing theories and separate lines of research converge to improve theories and models of instruction that improve student learning.

Conclusions and Future Directions for Research and Policy

To summarize this chapter, we make several observations about the state of primary reading instruction in the United States, which include salient findings about policy and from research. We also discuss future directions for policy, research, and practice.

The literacy achievement gaps between children from higher and lower SES homes and for ELLs versus native speakers have closed somewhat since 1998 (NAEP, 2013), but not to acceptable levels. They remain a complex and perplexing problem. These gaps are also evident globally. Policy has put increasing focus on what happens in the classroom and holds teachers accountable for their students' achievement. Already, the US Common Core State Standards (CCSS) have profoundly impacted reading instruction in the primary grades and will continue to do so in the future. While acknowledging that code-focused skills are critical, particularly during the early years of schooling, the CCSS target greater use of expository and informational text across the content areas, increasing from kindergarten on, involving deeper reading that requires analysis of text and higher-order reading comprehension. What will be the impact of the CCSS on student achievement in the United States, particularly for students who are most vulnerable—those with reading disabilities, those living in poverty or those for whom English is a second language? How can we improve the CCSS as we learn more about their impact? There is an unfortunate tendency in US education of following the latest fad rather than relying on research to effect gradual improvement in practices that lead to stronger student outcomes.

Research over the past few decades shows that virtually all students learning to read an alphabetic writing system, such as English, require at least some explicit and systematic instruction in the alphabetic principle and phonics to learn to read and that some children require more than others. Although there is less research on encoding, the research that is available suggests that encoding and decoding may develop synergistically, each supporting the other. Greater focus on encoding and writing to improve reading outcomes is likely to be a fruitful line of research (see also Kessler & Treiman, Deacon & Sparks, and Caravolas & Samara, this volume).

Despite converging evidence that RTI can reduce the percentage of students with weak reading skills, schools and practitioners need additional guidance from researchers. In particular, how do we establish the best practices for moving students up and down tiers? Whereas some data-guided methods exist, there is no uniform definition for what constitutes adequate versus inadequate response to instruction and intervention. It seems clear that students who still have reading difficulties even when provided intensive and tailored Tier 3 instruction will need ongoing help to maintain their word reading skills and to improve their fluency and comprehension. Thus, as the gap between their current instructional skill level and general education instruction at Tier 1 grows in response to higher standards set by the CCSS, finding methods to support the achievement of students with reading difficulties will be vital (see also Goldman & Snow, this volume). It will also be challenging to ensure that students with reading disabilities receive evidence-based accommodations and differentiated instruction that allow them to progress toward college and career readiness, particularly with regard to content area literacy (i.e., science, social studies). The inclusion of Tier 2 and Tier 3 reading comprehension interventions within multitiered systems of instruction might provide even stronger student outcomes, although this remains to be tested. Relatedly, it is vital to ensure alignment across core instruction and intervention for all students, but particularly for students who also receive instruction in their native language.

With regard to reading comprehension instruction, a common thread across effective interventions is that students are taught to attend to what they are reading more carefully and to think more systematically about the meaning of text. Accumulating research points to the centrality of more sophisticated language skills and reasoning in reading for understanding and suggests reasons why comprehension may break down even when decoding skills are adequate. Supporting children's development of the kinds of language they are expected to use to talk about their understanding of science, social studies, and narrative text is challenging for teachers.

Accumulating evidence demonstrates the efficacy of multicomponent instructional interventions for improving students' grasp of content knowledge, including the ability to read and learn from informational text. Research reveals the efficacy of strategy instruction, particularly text structure, for supporting students' reading comprehension of expository and informational texts. Effective interventions frequently include a focus on building students' understanding of texts through discussion (see also Goldman & Snow, this volume). Nevertheless, mixed results across studies of reading comprehension interventions suggest that we still have much to learn before we fully understand the multiple components and active ingredients of comprehension interventions that are effective—and this is likely to vary depending on student characteristics. In general, multicomponent interventions that include a focus on developing oral language skills, as well as strategy instruction and discussions about the meaning of text, are likely to be more effective than single-strategy interventions. Immediate corrective feedback—either explicit or implicit—appears to support students' gains in comprehension. However, this remains to be tested.

Although there are competing theories regarding the design, implementation, and evaluation of effective differentiated or individualized reading instruction, we hope to see disparate lines of research converge to provide insights on how to meet the needs of students with different skills and aptitudes. Evidence suggests that the effects of effective reading instruction accumulate from kindergarten through third grade. Efficacious instruction maximizes the amount of time spent in meaningful instruction and includes both code-focused and meaning-focused instruction. In general, effective individualized reading instruction uses small group assessment-informed and interactive teacher-led instruction to address students' individual learning needs provides opportunities for students to work independently and with peers, nurtures students' motivation to read, uses multiple strategies, and empowers teachers to make key instructional decisions. These instructional regimes should be effective in higher poverty schools that serve many students at risk of academic underachievement. Although differentiating reading instruction is difficult for teachers, they can succeed with support and appropriate professional development. How do we bring these practices to schools?

As of this writing, the Reading for Understanding Network of researchers (http://ies.ed.gov/whatsnew/newsletters/july10.asp?index=roundncer) is conducting important and systematic research to develop and evaluate effective ways of improving reading comprehension for students from prekindergarten through high school, with much of their research cited in this handbook and more forthcoming. Improving reading for understanding will be the challenge of this decade. In the United States and worldwide, research investments in meeting educational challenges promise a faster pace for finding ways to improve primary reading instruction and thereby ensure that all students achieve the highest levels of reading proficiency possible for them. In this chapter we have highlighted some of the most pressing research needs. These include making sure all children reach their highest reading potential; bringing policy in line with what we know about effective reading instruction; developing more effective RTI multitiered models and protocols that include both decoding and comprehension intervention; improving students' oral language, including content-area knowledge; improving comprehension skills; and figuring out how to meet the diverse needs of students in general education classrooms. We have useful research findings, but unfortunately they do not always make it into the classrooms where they are needed. Finding better ways of bringing research into the classroom, understanding how to effectively partner with schools and districts, and making research accessible to teachers, educational leaders, and policymakers will be important for rigorous research, practice, and policy.

Acknowledgments

The writing of this chapter was supported in part by the National Institute of Child Health and Human Development, *Florida Center for Reading Research, Learning Disabilities Research Center* (P50 HD052120) and by the US Department of Education, Institute of Education Sciences, *Reading for Understanding* (R305F100027).

Note

1 Note that the authors used the terms "first" and "second language learners."

References

Al Otaiba, S., Connor, C. M., Folsom, J. S., Greulich, L., Meadows, J., & Li, Z. (2011). Assessment data-informed guidance to individualize kindergarten reading instruction: Findings from a cluster-randomized control field trial. *Elementary School Journal*, *111*, 535–560.

Al Otaiba, S., Connor, C. M., Folsom, J. S., Greulich, L., Wanzek, J., Schatschneider, C., & Wagner, R. K. (2014). To wait in Tier 1 or intervene immediately: A randomized experiment examining first grade response to intervention (RTI) in reading. *Exceptional Children*, *8*, 11–27.

Al Otaiba, S., & Torgesen, J. K. (2007). Effects from intensive standardized kindergarten and first grade interventions for the prevention of reading difficulties. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 212–222). New York, NY: Springer.

Beach, K. D., & O'Connor, R. E. (2013). Early response-to-intervention measures and criteria as predictors of reading disability in the beginning of third grade. *Journal of Learning Disabilities*, *44*, 67–183.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York, NY: Guilford.

Benner, G. J., Nelson, J. R., Ralston, N. C., & Mooney, P. (2010). A meta-analysis of the effects of reading instruction on the reading skills of students with or at risk of behavioral disorders. *Behavioral Disorders*, *35*, 86–102.

Berkeley, S., Bender, W. N., Peaster, L. G., & Saunders, L. (2009). Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities*, *42*, 85–95.

Berkeley, S., Scruggs, T. E., & Mastropieri, M. A. (2010). Reading comprehension instruction for students with learning disabilities, 1995–2006: A meta-analysis. *Remedial and Special Education*, *31*, 423–436.

Block, C. C., Parris, S. R., Reed, K. I., Whiteley, C. S., & Cleveland, M. D. (2009). Instructional approaches that significantly increase reading comprehension. *Journal of Educational Psychology*, *101*, 262–281.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology: Theoretical models of human development* (6th ed., Vol. 1, pp. 793–828). Hoboken, NJ: Wiley.

Carlisle, J. F., Dwyer, J., & Learned, J. (2013). Discussion as a means of learning to reason, read, and write analytically. In P. B. Miller, P. McCardle, & R. Long (Eds.), *Teaching reading and writing: Improving instruction and student achievement* (pp. 83–92). Baltimore, MD: Brookes.

Cavanaugh, C. L., Kim, A.-H., C. L., Wanzek, J., & Vaughn, S. (2004). Kindergarten reading interventions for at-risk students: Twenty years of research. *Learning Disabilities: A Contemporary Journal*, *2*, 9–21.

Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the Early Childhood Longitudinal Study (ECLS) kindergarten to first grade sample. *Journal of Educational Psychology*, *98*, 489–507.

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science*, *21*, 1106–1116.

Compton, D. L., Fuchs, D., Fuchs, L. S., Elleman, A. M., & Gilbert, J. K. (2008). Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. *Learning and Individual Differences*, *18*, 329–337.

Connor, C. M., Alberto, P. A., Compton, D. L., & O'Connor, R. E. (2014). Improving reading outcomes for students with or at risk for reading disabilities: A synthesis of the contributions from the Institute of Education Sciences Research Centers. http://ies.ed.gov/ncser/pubs/20143000/

Connor, C. M., Morrison, F. J., Fishman, B., Crowe, E. C., Al Otaiba, S., & Schatschneider, C. (2013). A longitudinal cluster-randomized controlled study on the accumulating effects of individualized literacy instruction on students' reading from first through third grade. *Psychological Science*, *24*, 1408–1419.

Connor, C. M., Morrison, F. J., Fishman, B., Giuliani, S., Luck, M., Underwood, P. S., . . . Schatschneider, C. (2011). Testing the impact of child characteristics × instruction interactions on third graders' reading comprehension by differentiating literacy instruction. *Reading Research Quarterly*, *46*, 189–221.

Connor, C. M., Morrison, F. J., Fishman, B. J., Schatschneider, C., & Underwood, P. (2007). The early years: Algorithm-guided individualized reading instruction. *Science*, *315*, 464–465.

Connor, C. M., Morrison, F. J., & Katch, E. L. (2004). Beyond the reading wars: The effect of classroom instruction by child interactions on early reading. *Scientific Studies of Reading*, *8*, 305–336.

Connor, C. M., Morrison, F. J., Schatschneider, C., Toste, J., Lundblom, E. G., Crowe, E., & Fishman, B. (2011). Effective classroom instruction: Implications of child characteristic by instruction interactions on first graders' word reading achievement. *Journal for Research on Educational Effectiveness*, *4*, 173–207.

Connor, C. M., Morrison, F. J., & Underwood, P. (2007). A second chance in second grade? The independent and cumulative Impact of first and second grade reading instruction and students' letter-word reading skill growth. *Scientific Studies of Reading*, *11*, 199–233.

Dahl, K. L., & Freppon, P. A. (1995). A comparison of inner city children's interpretations of reading and writing instruction in the early grades in skills-based and whole language classrooms. *Reading Research Quarterly*, *30*, 50–74.

Denton, C. A., Fletcher, J. M., Anthony, J. L., & Francis, D. J. (2006). An evaluation of intensive intervention for students with persistent reading difficulties. *Journal of Learning Disabilities*, *39*, 447–466.

Denton, C. A., Tolar, T. D., Fletcher, J. M., Barth, A. E., Vaughn, S., & Francis, D. J. (2013). Effects of tier 3 intervention for students with persistent reading difficulties and characteristics of inadequate responders. *Journal of Educational Psychology*, *105*, 633–648.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., & Klebanov, P. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428–1446.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*, 393–447.

Fuchs, L. S., & Vaughn, S. (2012). Responsiveness-to-intervention: A decade later. *Journal of Learning Disabilities*, *45*, 195–203.

Gamse, B. C., Jacob, R. T., Horst, M., Boulay, B., & Unlu, F. (2008). *Reading First Impact Study final report (NCEE 2009-4038)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education.

García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, *84*, 74–111.

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2008). *Assisting students struggling with reading: Response to intervention and multi-tier intervention for reading in the primary grades. A practice guide. NCEE 2009-4045*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. URL: http://ies.ed.gov/ncee/wwc/PracticeGuide.aspx?sid=3

Gilbert, J. K., Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Barquero, L. A., & Cho, E. (2013). Efficacy of a first-grade responsiveness-to-intervention prevention model for struggling readers. *Reading Research Quarterly*, *48*, 135–154.

Guthrie, J. T., Anderson, E., Aloa, S., & Rinehart, J. (1999). Influences of concept-oriented reading instruction on strategy use and conceptual learning from text. *Elementary School Journal*, *99*, 343–366.

Guthrie, J. T., McRae, A., Coddington, C. S., Klauda, S. L., Wigfield, A., & Barbosa, P. (2009). Impacts of comprehensive reading instruction on diverse outcomes of low- and high-achieving readers. *Journal of Learning Disabilities*, *42*, 195–214.

Hiebert, E., & Fisher, C. W. (2007). Critical word factor in texts for beginning readers. *Journal of Educational Research*, *101*, 3–11.

Hill, D. R., King, S. A., Lemons, C. J., & Partanen, J. N. (2012). Fidelity of implementation and instructional alignment in response to intervention research. *Learning Disabilities Research and Practice*, *27*, 116–124.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, *2*, 127–160.

James-Burdumy, S., Deke, J., Gersten, R., Lugo-Gil, J., Newman-Gonchar, R., Dimino, J., . . . Liu, A. Y.-H. (2012). Effectiveness of four supplemental reading comprehension interventions. *Journal of Research on Educational Effectiveness*, *5*, 345–383.

Kintsch, W. (1998). *Comprehension: A paradigm for cognitions*. New York, NY: Cambridge University Press.

Konstantopoulos, S., & Chung, N. (2011). The persistence of teacher effects in elementary grades. *American Educational Research Journal*, *48*, 361–386.

Lam, E. A., & McMaster, K. L. (2014). Predictors of responsiveness to early literacy intervention: A 10-year update. *Learning Disabilities Quarterly*, *37*, 134–147.

McCoach, B., O'Connell, A. A., Reis, S. M., & Levitt, H. A. (2006). Growing readers: A hierarchical linear model of children's reading growth during the first 2 years of school. *Journal of Educational Psychology*, *98*, 14–28.

McKeown, M. G., Beck, I. L., & Blake, R. G. K. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading Research Quarterly*, *44*, 218–253.

McNamara, D. S. (2013). The epistemic stance between the author and reader: A driving force in the cohesion of text and writing. *Discourse Studies*, *15*, 1–17.

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first- and second-language learners. *Psychological Bulletin*, *140*, 409–433.

Morrison, F. J., & Connor, C. M. (2002). Understanding schooling effects on early literacy. *Journal of School Psychology*, *40*, 493–500.

Morrison, F. J., & Connor, C. M. (2009). The transition to school: Child-instruction transactions in learning to read. In A. Sameroff (Ed.), *The transactional model of development: How children and contexts shape each other* (pp. 183–201). Washington, DC: American Psychological Association.

Murphy, P. K., Wilkinson, I. A. G., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology*, *101*, 740–764.

NAEP. (2013). *National Assessment of Educational Progress: The nation's report card*. http://nces.ed.gov/nationsreportcard/.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institutes of Health.

Oakhill, J., & Yuill, N. (1996). Higher order factors in comprehension disability: Processes and remediation. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and interventions* (pp. 69–92). Hillsdale, NJ: Erlbaum.

Piaget, J. (1960). *The psychology of intelligence*. Paterson, NJ: Littlefield, Adams.

Rapp, D. N., & van den Broek, P. (2005). Dynamic text comprehension: An integrative view of reading. *Current Directions in Psychological Science*, *14*, 276–279.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, *2*, 31–74.

Reis, S. M., McCoach, D. B., Little, C. A., Muller, L. M., & Kaniskan, R. B. (2011). The effects of differentiated instruction and enrichment pedagogy on reading achievement in five elementary schools. *American Educational Research Journal*, *48*, 462–501.

Roehrig, A. D., Duggar, S. W., Moats, L. C., Glover, M., & Mincey, B. (2008). When teachers work to use progress monitoring data to inform literacy instruction: Identifying potential supports and challenges. *Remedial and Special Education*, *29*, 364–382.

Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. K. (2010). *Practice guide: Improving reading comprehension in kindergarten through 3rd grade*. (NCEE 2010-4038). Washington, DC: Department of Education.

Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English Language Learners. *Review of Educational Research*, *75*, 247–284.

Snow, C. E. (2001). *Reading for understanding*. Santa Monica, CA: RAND Education and the Science and Technology Policy Institute.

Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: The effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology*, *41*, 225–234.

Swanson, E., Vaughn, S., Wanzek, J., Petscher, Y., Heckert, J., Cavanaugh, C., . . . Tackett, K. (2011). A synthesis of read-aloud interventions on early reading outcomes among preschool through third graders at risk for reading difficulties. *Journal of Learning Disabilities*, *44*, 258–275.

Tran, L., Sanchez, T., Arellano, B., & Swanson, L. H. (2011). A meta-analysis of the RTI literature for children at risk for reading disabilities. *Journal of Learning Disabilities*, *44*, 283–295.

Vaughn, S., Wanzek, J., Linan-Thompson, S., & Murray, C. (2007). Monitoring response to supplemental services for students at risk for reading difficulties: High and low responders. In S. R. Jimerson, M. K. Burns & A. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 234–243). New York, NY: Springer.

Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities*, *39*, 157–169.

Vellutino, F., Scanlon, D., Zhang, H., & Schatschneider, C. (2008). Using response to kindergarten and first grade intervention to identify children at-risk for long-term reading difficulties. *Reading and Writing: An Interdisciplinary Journal*, *21*, 437–480.

Wanzek, J., & Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *School Psychology Review*, *36*, 541–561.

Weiser, B., & Mathes, P. (2011). Using encoding instruction to improve the reading and spelling performances of elementary students at risk for literacy difficulties: A best-evidence synthesis. *Review of Educational Research*, *81*, 170–200.

Wijekumar, K. K., Meyer, B. J. F., & Lei, P. (2012). Large-scale randomized controlled trial with 4th graders using intelligent tutoring of the structure strategy to improve nonfiction reading comprehension. *Educational Technology Research and Development*, *60*, 987–1013.

Williams, J. P., Stafford, K. B., Lauer, K. D., Hall, K. M., & Pollini, S. (2009). Embedding reading comprehension training in content-area instruction. *Journal of Educational Psychology*, *101*, 1–20.

Wilson, E. O. (1998). *Consilience: The unity of knowledge*. New York, NY: Alfred A. Knopf.

Zirkel, P. A., & Thomas, L. B. (2010). State laws and guidelines for implementing RTI. *Teaching Exceptional Children*, *43*, 60–73.

CHAPTER

28 African American English and Its Link to Reading Achievement

Holly K. Craig

Abstract

African American English (AAE) is a major American dialect. Recent research has focused on student patterns of AAE feature usage and found important relationships between AAE and reading achievement. This chapter provides background information on the nature of dialects and then focuses on AAE, identifying the major features that characterize child discourse. Intrinsic student factors and extrinsic influences on feature production are discussed as well. An important influence on AAE feature production is style shifting: the changes a speaker makes to his or her speaking patterns in response to differences in the communication context. The chapter discusses recent research that shows an inverse relationship between AAE feature production and reading achievement, and the mounting evidence that a student's ability to style shift from AAE to Standard American English in literacy tasks is positively related to reading achievement. A final section of the chapter identifies needed directions for future research.

Key Words: African American English, American dialect, reading achievement, style shifting, literacy, feature production, elementary students, bidialectalism

All languages have dialects. American English is the sum of many dialects, and Americans all speak some dialect of American English. The dialects of American English are regionally and socioculturally determined, and scholars propose that they are distinguished by thousands of differences (Dictionary of American Regional English [DARE], 2014). This chapter will focus specifically on African American English (AAE) and will describe the significant impacts AAE has on students' reading achievement. African American English is a major sociocultural dialect and is one of the most widely spoken in the United States. Child AAE is increasingly well understood, benefiting from a surge in recent research. This interest is due at least in part to national testing and accountability movements that reveal substantial differences between the test scores of African American students and other racial groups.

Of practical and theoretical importance, recent investigations have found an inverse relationship between the amounts of AAE feature production characterizing a student's discourse and his or her reading scores. This new information should be applicable to the study of dialects in other language communities and suggests directions for exploring dialect-reading acquisition links in other countries. The goals of this chapter are to provide the reader with an overview of AAE, particularly child use of the dialect; to discuss the evidence for relationships between AAE and reading achievement; and to conclude with a discussion of needed directions for future research.

American Dialects

Most dialects trace their distinctive patterns to the histories of the settlers of their region or sociocultural group. However, dialects are not static but dynamic, with features that are constantly changing, so today's usages are markedly different from those of past generations and from decades and centuries ago. Shared speaking patterns within a community are important to the speakers' perceptions of self-identity as well as in defining membership within the community (Wolfram, 2004). Further, knowledge of more than one community's speaking patterns can have tangible payoffs. For example, for AAE-speaking students, bidialectalism relates to higher reading achievement scores. For AAE-speaking adults, bidialectalism translates into better wages (Grogger, 2011).

When describing American dialects, scholars typically use Standard American English (SAE) as the frame of reference. They identify the specific features associated with regional and sociocultural dialects in terms of their differences from the way the same meanings would be expressed in the more general forms of SAE. African American English is as rule governed and complex as SAE, showing distinctive and systematic differences from SAE in sound patterns, grammar, word choices, and the vocal characteristics of pitch and cadence.

Dialectal variations are governed by rules; they are not random variations from a reference dialect. One dialect is distinguished from another primarily on the basis of systematic sound and word differences. For example, the Mid-Atlantic Dialect (Labov, 2007) spoken in and around the state of Maryland, adds an /ɹ/ after the vowel /a/ so that "Washington" sounds more like "Wa<u>r</u>shington" to speakers who do not share the regional dialect. (Underlining is used throughout the chapter to highlight the form of interest.) Dialects show differences based on smaller regions as well, and can be distinct to specific communities. For example, dropping of the /ɹ/ sound is characteristic of residents of the city of Boston. Dialectal differences can be quite complex. Consider the sentence "Pa<u>r</u>k the ca<u>r</u>." A Bostonian would probably omit the /ɹ/, saying: "Pa_k the ca_," but not so if followed by a word beginning with a vowel: "Pa_k the ca<u>r</u> over there." Regional synonyms exist as well, which are familiar to local speakers. In my home state of Michigan, the word "long john" refers to a type of underwear, as it does in other states. But Michiganders also recognize that when offered a long john, this probably refers to an iced, oblong-shaped, jelly-filled pastry (DARE, 2014). Regional and sociocultural dialects are not mutually exclusive, but show intersecting influences. For example, AAE, a sociocultural dialect, differs if the speaker is from Detroit as opposed to other regions. In Detroit, the first syllable of a number of two-syllable words is typically stressed, whereas it is the second syllable which receives the primary stress by speakers of other dialects. African American English–speaking Detroiters would say "<u>po</u>lice" and "<u>De</u>troit," whereas AAE speakers in other parts of the Midwest would probably follow the more general pattern of pronunciation: "po<u>lice</u>" and "De<u>troit</u>," emphasizing the second syllable. With increasingly easy access to information on a national and global scale, some dialectal forms, once exclusively associated with a particular speech community, have become popularized and widely adopted. You may hear the "<u>De</u>troit" pronunciation now across a variety of discourse venues.

The SAE dialect is the set of linguistic behaviors associated with educated and professional American discourse. It is used in more formal contexts, and it aligns more closely with written English than with other dialects. An opposition between informal language dialects associated with homes, communities, and informal vernaculars and the more formal language dialect of a country's politically and socially recognized or sanctioned dialect characterizes many countries and many languages. For example, like SAE, Classical Arabic (or Modern Standard Arabic) is the dialect of the Arabic language that is valued for academic and public discourse.

Dialects reflect their communities and thus have varying social currency from high to low prestige. In the United States, SAE is a highly valued dialect. Speakers of lower prestige dialects may find themselves the victims of linguistic prejudice. Often, judgments are made by listeners about a speaker's education level and intelligence based on his or her dialect use. Speakers whose dialects show many substantial differences from SAE may be devalued or stigmatized based on the way they talk.

The sociopolitical history of African Americans in the United States and the breadth of feature differences between AAE and SAE have resulted in substantial linguistic prejudice, which lingers even today. Both black and white listeners have rated speakers who use AAE features as lower in social status, socioeconomic well-being, intelligence, and even personal attractiveness (Koch, Gross, & Kolts, 2001; Rodriguez, Cargile, & Rich, 2004). Housing discrimination has been linked to linguistic prejudice (Massey & Lundy, 2001; Purnell, Idsardi, & Baugh, 1999). Grogger (2011) calculated that African Americans who "sound black" paid a substantial financial penalty, earning approximately 10% less than African American peers whose discourse was perceived as less racially distinctive. After controlling for measures of skill and family background, African American workers with fewer AAE speech characteristics were on par with the wages of whites who had similar work skills.

Linguistic prejudice may be experienced by African Americans at very early ages. For example, teachers' academic expectations for their African American students tend to be lower than for other students, and this includes estimations of their reading skills (Cecil, 1988). Teachers tend to correct more miscues in reading tasks by their African American students (Cunningham, 1976/1977; Markham, 1984), even when these variations are dialectal in nature. Teachers' perceptions of AAE as a low-status dialect can place AAE-speaking students at risk for reading difficulties (Goodman & Buck, 1973). Reading skills are foundational to classroom learning. Thus, negative teacher interactions can impede the acquisition of reading skills with detrimental consequences undermining achievement in the full range of content areas.

Child African American English

Children begin their formal education speaking the language of their communities. This means that many African American students speak AAE at the time of school entry. Until recently, most of what was known about AAE was derived from the study of older students and adults. However, the last twenty years or so have yielded a wealth of new information about dialect use by young African American students. Although dialectal differences reflect systematic variations across the full linguistic range of sound production, grammar, vocabulary, and prosody, the study of child AAE has focused primarily on its morphosyntactic and phonological characteristics.

Table 28.1 provides examples of commonly produced morphosyntactic and phonological features of child AAE, based on the reports of Oetting and McDonald (2002) in Louisiana; Renn and Terry (2009) in North Carolina; and Craig, Thompson, Washington, and Potter (2003) in Michigan. In contrast to these common features, some features were fairly rare in the discourse of children. For example, the remote past "been" is used as part of the verb to express something that took place in the distant past: "he <u>been</u> reading story books." The zero copula and subject-verb agreement features (see Table 28.1) were frequently observed features in a longitudinal data set of African American students residing in North Carolina; however, the remote past "been" was one of the more "obscure" features in those student samples (van Hofwegen & Wolfram, 2010). These Midwestern and Southern data sets indicate that specific patterns of common and rare usage should be considered central to the dialect and not simply regional in nature.

Phonological knowledge is fundamental to learning how to read (see Ehri, this volume). Students use their knowledge about the sounds of words as a base for building print knowledge of those words. AAE-speaking students have extensive knowledge of the phonological representation of SAE words. It does not appear that having alternative ways to produce words in AAE and SAE interferes with or confuses AAE speakers during reading acquisition. Specifically, high dialect feature producers have been found to accept SAE productions of words and to produce SAE versions of words as frequently as low dialect producers in structured tasks (Terry, 2014; Terry & Scarborough, 2011). Differences in metalinguistic awareness skills predicted reading achievement rather than differences in the phonological representations of dialect-sensitive words. Terry (2006) found that differences between dialect speakers and SAE speakers in the spelling of inflections were mediated partially by the students' morphosyntactic awareness. In other work, Terry and Scarborough (2011) found that phonological awareness fully mediated the contribution of dialect differences to reading. Craig, Kolenic, and Hensel (2014) also found that phonological awareness, morphosyntactic awareness, and pragmatic awareness predicted reading scores. Together, these recent studies indicate that it is not more or less knowledge of SAE that explains reading differences among dialect speakers. It is AAE speakers' awareness of linguistic forms, their sensitivity to context differences, and their ability to adapt their linguistic forms to the context requirements that influence reading acquisition.

Variability in Feature Production

African American English feature production varies widely among speakers. Not all African Americans speak AAE, and those who do may not use AAE features at all in some contexts or may vary in the degree to which they produce AAE features across different discourse contexts. For example, in a sample of fifty adults residing in the Chicago area, Craig and Grogger (2012) found that two participants produced no morphosyntactic AAE features at all during a brief interview, whereas others produced up to thirty-five exemplars during 4-minute samples of their discourse.

Variability in the frequencies of feature production used by child speakers of AAE can be traced

Table 28.1 Common Morphosyntactic and Phonological Features of Child AAE.

Morphosyntactic Features	Examples
Ain't, a negative auxiliary for *is+not, are+not, do+not, have+not*	"nope it ain't down there"
Invariant "be," used to express habitual states	"I be watchin' a lot of reality series"
Multiple Negation, where more than one negative form is used to express a negative meaning	"she didn't have nowhere to go"
Subject-verb Agreement Variations, where "-s" is not added to present tense verb forms	"the sign say danger"
Zero Copula, copula and auxiliary forms of the verb "to be" are variably included and excluded	"I don't know what she _ doing"
Zero Modal Auxiliary, "can," "will," "do," and "have" are variably included and excluded	"his grades and stuff _ been dropping"
Zero Past Tense "-ed," where the simple past tense form "-ed" is variably included and excluded	"the frog jump_ out and it was gone"
Zero Possessive "-s," the "-s" possessive form is variably included and excluded	"he yelled out the window and called the frog_ name"
Phonological Features	**Examples**
Consonant Cluster Reduction, the deletion of phonemes from consonant clusters	"the little /lɪl/ girl dropped her papers"
"g" Dropping/ Nasal fronting, the substitution of /n/ for /ŋ/ in final word position	"and he blowing /bloʊɪn/ a whistle to stop the cars"
Monophthongization of Diphthongs, the neutralization of diphthongs	"and kids walking on the sidewalk /sɑdwɔk/"
Postvocalic Consonant Reduction, the deletion of consonant singles following vowels	"and they had /hæ/ fell in the puddle"
Substitution for /θ/ and /ð/, where /t/ and /d/ substitute for /θ/ and /ð/ in prevocalic positions and /f, t/ and /v/ substitute for /θ/ and /ð/ in intervocalic and postvocalic positions	"they /deɪ/ making fire" "both /boʊf/ of these boys had done it"

to four major sources of influence. Linguistic context is a major influence on whether the AAE or SAE form is more likely to be produced. For example, the simple past tense form "-ed" is more likely to be excluded after consonant sounds ("he check_ under his boots") than after vowels ("we hurried"), or when the addition of "-ed" adds a separate syllable to the bare verb form ("he started walking"). Similarly, the plural "-s" is more likely to be excluded after consonant sounds /t/, /d/, /b/, /p/, /k/, and /g/ ("two wonderful cake_") and more likely to be included after vowels ("my shoes"). The plural "-s" is also more likely to be excluded after a number word or the demonstratives "them" and "those" ("those two cupcake_").

Another major type of influence on feature production derives from extrinsic social characteristics of the speakers, in particular their community, socioeconomic status (SES), gender, and education. Charity (2007) examined regional differences in AAE feature production among 5- to 8-year-old students in New Orleans, Louisiana; Washington, DC; and Cleveland, Ohio. They found that both morphosyntactic and phonological feature production rates were higher for New Orleans. Charity suggested that geographic, social, and historical factors may intermingle and contribute to the higher AAE feature rates in New Orleans. In particular, students may perceive greater acceptance of AAE features according to local speech norms.

Within relatively small geographic regions, substantial community-by-community differences can be observed in AAE. The reasons for these more local dialect differences within regions warrant investigation, especially the extent to which segregation patterns contribute to the differences. Students who are exposed to SAE compared with those exposed only to AAE are likely to differ in the extent to which their discourse includes AAE features. Bountress (1983) found lower levels of a small subset of AAE features used by students attending integrated schools compared with students living in the same region who attended schools where African American students constituted 99% of the student body. Mean differences in feature production rates within our southeastern Lower Michigan communities are suggestive of segregation-related differences. Although the number of communities was small, our data indicate that students living in communities with low segregation levels used fewer AAE features in their discourse than African American students living in moderately or highly segregated communities, as reported in Table 28.2. Differences are reported in the table as *dialect density measures* (DDMs), which are the ratios of number of instances of AAE feature production to the number of words in the sample.

African Americans from low-income homes reportedly produce higher frequencies of AAE features than those from middle-income homes (Horton-Ikard & Miller, 2004; Washington & Craig, 1998). It is noteworthy that African Americans are three times more likely than Caucasians to live in low-income homes (Brooks-Gunn, Klebanov, & Duncan, 1996). In gender studies, the discourse of African American boys showed higher frequencies of AAE feature production than that of African American girls (Charity, 2007; Washington & Craig, 1998). However, both SES and gender effects on frequencies of AAE feature production diminished with increasing time in school and more exposure to SAE (Craig & Washington, 2004).

Academic achievement level is associated with rates of AAE feature production. Craig and Grogger (2012) observed that adults with less than a high school or graduate equivalent degree produced AAE morphosyntactic features approximately five times more often than college graduates. The reasons for this relationship remain unclear, and may not be causal or direct. African Americans are more likely to reside in low SES homes than Caucasians, and SES has a number of covariables such as inter generational illiteracy, fewer high-quality school and community resources, poor health, and limited healthcare. However, higher education provides greater exposure to SAE, probably making speakers more capable of and more interested in reducing their production of AAE features and adopting the SAE dialect, which is associated with educated and professional discourse (Mufwene, 2001). Some of these extrinsic social influences covary, complicating the overall interpretation of dialect patterns.

A student's language developmental status is important to characterizing the likelihood of AAE or SAE dialect usage. Grade is a significant influence on feature production rates during the elementary school years. When evaluated in school settings using tasks that prompted for SAE, some but not all AAE-speaking students decreased the amount they produced AAE features over the course of the elementary grades (Bountress, 1983; Craig & Washington, 2004; Isaacs, 1996; Ivy & Masterson, 2011). The feature production rates of preschoolers and kindergartners were observed to be more than

Table 28.2 Kindergartener Mean (Standard Deviation) Dialect Density Measures (DDMs) From Four Communities Within a Single Geographic Region, Southeastern Lower Michigan.

City	Census Characterization	Segregation Level[a]	Mean DDM	Feature per Words
Ann Arbor (n = 56)	Midsize city	33.4 low-moderate	.053 (.043)	1 feature/19 words
Flint (n = 79)	Midsize city	74.6 very high	.073 (.043)	1 feature/14 words
Jackson (n = 17)	Small city	46.6 moderate	.075 (.046)	1 feature/13 words
Oak Park (n = 80)	Large suburb	45.6 moderate	.095 (.058)	1 feature/11 words

[a] Index of Dissimilarity (Lewis Mumford Center for Comparative Urban and Regional Research, 2013): Range = 0–100, where ≤ 30 is low, 40–50 is moderate, ≥ 60 is very high

twice the average rates of first through fifth graders in a study of 400 African American students during an oral narrative task (Craig & Washington, 2004). First grade was the time when substantial changes appeared to take place, corresponding for these students to their enrollment in full-day public school programs.

Craig and Washington (2004) reported a significant downward shift in AAE feature usage at first grade, from approximately one morphosyntactic feature per ten words at preschool and kindergarten to only one per twenty-six words at first grade through fifth grade. These averaged values reflected sharp declines in AAE feature production for many of the individual students. Sixty-eight percent of the first through fifth graders clustered in a low AAE feature production group. However, approximately one-third of students in first through fifth grades continued to produce moderate to high levels of AAE features. Similarly, van Hofwegen and Wolfram (2010) found a significant decrease in feature production from age 48 months to grade one.

To illustrate these grade-related differences, Table 28.3 presents a segment of language by an African American boy "reading" the wordless story book *Frog, Where Are You?* (Mayer, 1969). Wordless story books represent an emergent reading context for young students, and they are appropriate for students in the early elementary grades from culturally diverse backgrounds (Muñoz, Gillam, Peña, & Gulley-Faehnle, 2003; Schachter & Craig, 2013). The student and the examiner looked through a story book depicting a boy, a dog, and a frog in which the frog runs away and the boy and dog search for it. The segments in Table 28.3 were produced by a student when he was a kindergartener, first grader, and second grader. They are excerpted from the end of his stories and show his sentence knowledge and knowledge of narrative structure growing over time, as well as changes in his dialect usage. In this emergent reading task, the transcripts show a dialect shift away from the production of AAE at kindergarten toward SAE forms at first and again at second grade.

Across the age–grade span of 48 months to tenth grade, van Hofwegen and Wolfram (2010) characterized AAE feature use as a roller coaster trajectory, a term representing the nonlinear increases and decreases in feature production observable across development. Prior to third grade, AAE feature production rates have been shown to decrease (Craig, Thompson, Washington, & Potter, 2004). However, van Hofwegen and Wolfram observed a significant increase in AAE feature production from grade six to grade eight, and then a significant decrease from grade eight to grade ten. Across studies, the trajectories depended in part on whether total AAE feature production was examined or only a subset of the most common features. Total rates tended to show a decline with age and grade; however, production of the most common types of features stayed level.

The linguistic systems of early elementary grade students are still undergoing substantial development. Major phonological and grammatical rules of the child's language system are acquired prior to school entry; however, vocabulary, prosody, and

Table 28.3 Closing Segments of the Reading of a Wordless Storybook by a Male African American Student.

Kindergarten	First Grade	Second Grade
and they fell<u>ed</u> off into the water	and he almost fell	and then they fell into a pond
and they fell in water with their heads down	he heard the frog	then he heard the frogs
and the dog _ on his head	the dog was being loud	the boy told the dog to be quiet
he _ telling the dog to be quiet	and he was telling him to be quiet	they saw two frogs
and they got on this tree	and they looked over the bridge	then they saw little frogs
_ dog found some frogs	and they found him	the boy took one of the frogs home with him
he found a lot of frogs	he said bye and took the little frog with him	that's the end
it's done	the end	

advanced grammatical forms continue to develop through the elementary grades. To illustrate, my research in Michigan communities shows three interesting patterns in AAE morphosyntactic feature development that contribute to variations in feature production among students, and these are discussed in what follows.

First, features involving more complex verb forms with multiple verbal elements, especially completive "done," double auxiliaries and modals, and remote past "been," were not used by students until later in development. These multiple verbal elements conveyed multiple meanings, (e.g., past tense plus completed action: "then the dog done run away"). In contrast, the features subject-verb agreement variations, "was/were," zero copula, zero past tense "-ed," and zero "-ing" occurred earlier in development. These features were conceptually simpler in that the verbal elements tended to express single meanings: for example, "and the bees _ tryin' to get him."

Often early feature production reflected an overly broad application of the feature to a variety of linguistic contexts; this range of contexts becomes more narrowly defined in adult usage. One example involved the remote past "been" feature. Early elementary grade students used the form "been" to communicate something that occurred in the past ("The dog been running because the bee coming"), not just the distant past, as is more typical of adult usage, as in "she been know that since kindergarten." Another example involved the preterite "had" feature. Unlike the use of "had" as a simple past-tense main verb ("and the dog had his nose in a jar looking at it"), the preterite "had" in AAE occurs prior to simple past-tense regular and irregular verb forms in narratives and functions to introduce a complicating action in the plot (Ross, Oetting, & Stapleton, 2004; Schachter & Craig, 2013) ("and the boy had saved him before he got to hurt hisself"). Green (2011) observed that preverbal "had" combines with present-tense verbs early in the child's acquisition of AAE forms. In support of Green's proposal, we have found that early elementary grade students did not restrict "had" to simple past-tense contexts but also produced the form in present-tense contexts where the function may not be specific to the introduction of a complicating action ("and then he had say what you doin' in the tub with your dog"). In our data the main verb was most typically an irregular, further emphasizing the child's immature linguistic status as she or he sorts out regular and irregular verb forms. The double modal feature of AAE used by adults results in combinations of "would," "might," and "could," as in "we might could go there." However, our early elementary grade students tended to combine copula forms, such as "I'm is making footprints." The principle of multiple marking was the same for the young students as for the adults, but the verb element was consistent with the child's less mature auxiliary system.

Stylistic factors are another major influence on the production of AAE features, such that speakers adjust their patterns of feature production based on characteristics of the discourse context (Preston, 1991). These changes, identified as style shifting, dialect shifting, or code switching, involve grammatical features more frequently than phonological features (Wolfram, 2004). As will be discussed in the next section of this chapter, style changes in response to context differences are especially important for understanding reading achievement by AAE-speaking students.

Common grammatical features are more likely than rare features to show changes during style shifting (Bell, 1984; Craig & Grogger, 2012). For adults, the topic of the conversation influences whether AAE forms are more or less likely to be used. "Casual," "intimate," or more "ethnic" informal topics have been associated with higher frequencies of AAE features than "message-oriented," more "formal," or more "mainstream" topics (Baugh, 1983; Craig & Grogger, 2012; Labov, 1972; Rickford & McNair-Knox, 1994). Craig and Grogger (2012) found that adults shifted their patterns of AAE feature production when responding to questions that were expected to elicit differences in feature production, specifically situational compared with metaphorical (Blom & Gumperz, 1972), formal versus informal (Baugh, 1983; Labov, 1972), and mainstream and message-oriented about work, meetings, and the conveyance of specific information compared with personal topics (Linnes, 1998; Rickford & McNair-Knox, 1994). The speakers significantly decreased their rates of AAE feature production when talking about what they would say in a job interview or during a medical appointment (formal context, message-oriented, metaphorical), compared with talking about their leisure activities (informal, personal, situational). The AAE features most commonly used across speakers were the features that decreased the most between informal and formal topics.

Although the study of style shifting by child speakers of AAE remains limited, children do show differences in their feature production rates based on the type of discourse genre. For example, third graders showed significant differences in feature production rates when asked to read aloud and prepare written narratives, compared with their use of features in oral narratives (Thompson, Craig, & Washington, 2004). Style shifting has been observed as early as preschool (Connor & Craig, 2006).

In summary, speakers of AAE show many differences from speakers of SAE. However, SAE is the dominant and valued dialect of academic discourse. Students vary widely in their rates of AAE feature production, spanning high to low feature users. A number of factors underlie the rules for AAE feature production. As discussed in the next section of the chapter, one of these sources is the style shifting of morphosyntactic forms from AAE to SAE as the context requires, which advantages students for the acquisition of reading skills.

Relationships Among Child AAE and Literacy Achievement

Now that the characteristics of child AAE have been discussed, its implications for reading are considered.

The Black–White Achievement Gap

Far too many students in the United States fail to achieve competency in literacy in general, and reading in particular. Low achievement levels are especially pronounced for racial and ethnic minority students and for those from socioeconomically disadvantaged homes (National Center for Education Statistics [NCES], 2011, 2012). African American students are a large minority population in the nation's schools, and a long-standing black–white test score gap characterizes the persistent differences between school-age African American students and their non-Hispanic white peers (Jencks & Phillips, 1998). The National Assessment of Educational Progress (NAEP) is used in the United States to evaluate and track academic achievement from first through twelfth grades. For example, at fourth grade, 79% of white students were reading at or above basic levels on the 2013 NAEP, defined as partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at that grade. The same was true for only 50% of black students. Only 46% of white students were reading at or above proficient levels, defined as solid academic performance and competency over challenging subject matter; however, the same was true for only 18% of black students (US Department of Education, 2013). The test score gap was observed as early as school entry, and continued across all of the content areas through twelfth grade (US Department of Education, 2009). Unfortunately, these disparities affected most achievement indicators, such that African Americans were more likely to be held back a grade, be suspended or expelled, and drop out of high school; college enrollment and graduation rates were lower, as were median earnings and unemployment levels (Hoffman & Llagas, 2003).

Many factors have been identified as potential contributors to the gap, including unequal educational opportunities due to racial segregation, low teacher expectations, high levels of poverty in this population, low cognitive skills, low levels of home literacy, and dialect interference in the learning of literacy skills (see Washington & Craig, 2001, and Washington, 2001, for more discussion). An often-cited explanatory variable for academic underachievement is the high rates of poverty endured by many African American families. African American students are three times as likely to live in poverty as their non-Hispanic white peers (Aud, Fox, & KewalRamani, 2010). When families have limited access to food, clothing, shelter, and medical care, a child's health and cognition can be compromised (Bradley & Corwyn, 2002). Further, when caregivers cannot provide adequately for their children, homes can have high levels of stress, leading to socioemotional problems (McLoyd, 1990). All of these factors can have an impact on learning in negative ways. High-quality school experiences can mitigate some of the effects of poverty for African American students (Craig, Connor, & Washington, 2003). Unfortunately, children residing in poor families are also likely to reside in low-SES communities, resulting in attendance at low-quality schools. Poverty relates to reading achievement both directly and indirectly (Craig et al., 2003; Nievar & Luster, 2006). Poverty shows the strongest effects in the earliest grades (Lee & Burkham, 2002). After first grade, oral language, particularly rates of AAE feature production, showed greater influences on standardized reading scores than SES (Craig, Zhang, Hensel, & Quinn, 2009).

Poverty is not something that schools can resolve, but perhaps there are other more malleable factors that can be addressed. With improved understanding of the importance of strong literacy skills for success across all academic content, research questions have been reframed from the influences of racial and poverty gaps to emphasize the role of language differences, especially dialect, in reading.

Role of African American English

Whereas early oral language skills help to create the foundation for reading acquisition (Snow, Burns, & Griffin, 1998; Storch & Whitehurst, 2002), dialect differences in the oral language of African American and non-Hispanic white peers have been hypothesized to contribute to the achievement gap. Probing for a potential relationship between dialect and reading achievement initially used a research strategy that selected a small set of specific morphosyntactic or phonological features and looked for statistical associations with reading scores. These attempts showed no significant relationships (Goodman & Buck, 1973; Harber, 1977; Seymour & Ralabate, 1985). Later research designs included more holistic characterization of dialect usage, adopting rate-based DDMs of phonological features and morphosyntactic features.

Today's rate-based approaches have revealed inverse relationships between the amounts of AAE feature production and performances on a broad range of literacy measures. The more dense the AAE feature production the lower the standardized reading scores (Craig et al., 2009), while controlling for SES, general oral language, and general writing skills. Amounts of AAE feature production were also inversely related to letter identification, word and nonword reading (Charity, Scarborough, & Griffin, 2004; Connor & Craig, 2006), accuracy and rate of oral reading (Craig et al., 2004), passage comprehension (Charity et al., 2004), spoken accuracy and spelling of inflections and nonwords (Kohleret al. 2007), receptive vocabulary (Craig & Washington, 2004), and writing (Ivy & Masterson, 2011).

African American students in moderate to highly segregated communities may enter school producing forty or more morphosyntactic and phonological features of AAE (Craig & Washington, 2004, 2006). It has been hypothesized that these many differences may result in mismatches between their spoken word productions and the spelling of printed words (Labov, 1995; LeMoine, 2001), resulting in confusion and increased error rates during reading. Similarly in other language communities, Ibrahim (1983) and Alrabaa (1986) proposed that poor acquisition of literacy skills in Arabic can result from mismatches between colloquial Arabic dialects and Classical Arabic. Saiegh-Haddad (2003) observed that the learning of Classical Arabic needed for reading and writing was more difficult for students in northern Israel who spoke their local dialect of Northern Palestinian. Overall, it appears that when the dialect of the home community and the language of the classroom do not align well, learning to read may be negatively impacted.

Dialect Shifting–Reading Achievement Hypothesis

Students developing language competence learn that people speak differently in different contexts. African American English feature production has been found to relate to context. It has been hypothesized that the problem for reading acquisition may not be a lack of knowledge of SAE or interference between phonological representations of words that are consistent with AAE pronunciation but not with their corresponding SAE spellings. Rather, the critical variable for dialect speakers learning to read has been observed to be the extent to which a student learns that she or he must style shift in literacy in tasks with expectations for SAE (Craig et al., 2014; Terry, 2014). Craig et al. (2009) examined rates of AAE feature production in two contexts: the writing of a short narrative and the telling of a short oral narrative, tasks that are more and less likely to elicit SAE, respectively. They found that the rates of AAE feature production during the oral narrative did not predict scores on standardized reading, whereas lower rates of AAE feature production by these same students in the writing context predicted better reading scores.

Students begin to show genre sensitivity as early as preschool. Connor and Craig (2006) found that preschoolers decreased their production of AAE features during imitation of SAE sentences—a task with high expectations for SAE—compared with their feature production during storytelling using a wordless story book, a task with more moderate expectations for SAE. The language samples presented in Table 28.4 illustrate feature production patterns by genre, where one student shifts away from AAE forms during a writing task compared with his peer, who does not demonstrate linguistic adaptation to the writing context. Analyses focus on the students' morphosyntactic features because morphosyntactic features are the component of the dialect system that speakers alter when style shifting (Wolfram, 2004).

The hypothesized relationship between shifting away from AAE in literacy tasks and higher achievement may not be limited to reading but may impact writing skills as well. Ivy and Masterson (2011) examined the use of six AAE features in speaking and writing contexts and

Table 28.4 Oral and Written Narrative Samples by Two Male Third Graders.

The first student shows evidence of style shifting in writing, and the second student shows no evidence of style shifting. The samples are presented as written by the students.

Student 1:

Oral Picture Description Narrative of Ice Skating	Written Description of A Boy Playing Basketball, Showing Evidence of Style Shifting
there's a danger sign that says danger-thin-ice and umm there's cuts in the ice and there's two boys tryin' to grab each other and one's falling in the water and somebody _ fallin' and it's a fire and it's a snowman and it's a shovel	Ounce upon a time there was a boy named Jack. He liked to play basketball every day. One day when he came from school he road his bike to the store to by his mom some stuff to cook. The next day he was at school playing basketball with his friends from a derfrent class. Then he took his fawlshot and made for one point

Student 2:

Oral Picture Description Narrative of An Accident	Written Description of Visiting A Cousin, Showing No Evidence of Style Shifting
it's he, he um _ wrapped up in some covers and a police _ tellin' cars to stop and it's a car right there that's already in a accident and this one, and these people they umm _ feelin' sad for this little boy gettin' hit by the car and they _ takin' him in a trunk, in a truck	On Sunday I went to my cousin_ house and played his game. We went to the store and we had walked. I had a little cousin to he's is sometimes bad. We was playing basketball down staris. I was playing in my little cousin_ power wheel it go_ relly fast

found that third graders produced AAE features with similar frequencies during speaking and writing tasks. However, eighth graders significantly decreased their AAE feature production during writing.

Recently, a small number of longitudinal studies have been undertaken to confirm and probe the relationships that were observed between reading and AAE in prior cross-sectional research designs. The challenges of longitudinal research with this population of students are substantial due to the high levels of attrition among participants. Many school districts with large minority enrollments are in communities with serious economic problems. For example, Flint, Michigan, has partnered with us in many research projects for a number of years. The major employer in Flint is the auto industry, and the public school systems have been negatively affected by losses to this industry. The primary reasons for attrition in our research tend to be that the student moved away when a caregiver became unemployed or the school was closed and contact was lost due to cuts in administrative staff (Craig et al., 2014).

Although limited in number, longitudinal studies of dialect and literacy achievement are beginning to be reported. They are especially important at this time to confirm findings based on averaged cross-sectional data but at the level of the individual student when so much intra- and inter subject variation in feature production characterizes the discourse samples under study. Terry and colleagues (Terry & Connor, 2012; Terry, Connor, Petscher, & Conlin, 2012) have examined the relationships between dialect and literacy achievement in students with typical and atypical language development, most of whom were African American. These studies confirmed that students with greater levels of SAE showed greater growth in reading skills. In our Michigan program of research, we followed 102 typically developing, primarily low-SES, AAE-speaking students from kindergarten through second grade, assessing AAE feature production, general oral language and cognitive skills, and reading three times a year (Craig et al., 2014). In this study, a *style shifting coefficient* was created that modeled a student's feature production changes over time and between contexts. We found that style shifting was not related to grade; further, evidence of style shifting at any grade predicted better reading scores at second grade.

The Current Educational Context

Widespread adoption across the United States of the Common Core State Standards (CCSS, National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010) heralds major changes in expectations for both students and teachers. Currently, by the end of the elementary grades, students must demonstrate command of standard English grammar when writing, speaking, reading, or listening. Students must also understand how language functions in different contexts and be able to make effective choices for meaning and style specific to a variety of contexts.

The expectations for formal and conventional forms of English can be quite challenging to teachers for a number of reasons. One reason is that many teachers do not share the culture of their African American students and may have little personal knowledge of AAE. In 2011, 84% of US teachers were Caucasian (National Center for Education Information) and 30.6% of students were African American (United States Census Bureau). Many teachers with little personal knowledge or experience of AAE are being asked to teach African American students to change their dialect usage in classroom contexts. These teachers do not have the skills or the language knowledge necessary to help AAE-speaking students learn to be bidialectal and, specifically, how to style shift to SAE for academic tasks (Fogel & Ehri, 2000).

Two types of teacher management of AAE prevail and can be observed widely across the United States in schools with high levels of minority language enrollments. Teachers who do not understand AAE often classify its forms as errors. This can result in negative interactions between students and teachers (Cunningham, 1976/1977; Washington & Miller-Jones, 1989), which can hinder student learning. Alternatively, teachers sometimes take an eradicationist approach to the dialect, viewing AAE use as inferior to SAE use and therefore attempting to remove it completely from students' speech (Lippi-Green, 1997; Smitherman, 1974, 2000). Similarly, in Arabic classrooms, bidialectalism is discouraged. Students must abandon their home language forms and learn the forms of Classical Arabic (Maamouri, 1998). These eradicationist practices are highly problematic. Regardless of whether teacher correction is rooted in ignorance or in intentional discrimination, these teacher management strategies devalue the students' cultural-linguistic backgrounds (Green, 2002; Rickford, 1998). Further, frequent negative classroom interactions foster high levels of student disengagement (Fordham, 1999; Wolfram, Adger, & Christian, 1999) and academic underachievement (Erikson, 1987). It is not that AAE is wrong, as eradicationist approaches suggest, but that there is a standard mainstream alternative (SAE for English and MSA for Arabic) that will be more advantageous in learning to read. Effective new methods that are respectful and constructive are greatly needed.

Future Research

Some African American students decrease the extent to which they use AAE features in classroom contexts naturally as a by-product of schooling and without formal instruction (Adler, 1992; Battle, 1996; Craig & Washington, 2004; Isaacs, 1996). A critical research question is whether style shifting to SAE can be taught. With evidence accumulating that the ability to become bidialectal and use SAE for academic purposes relates to better reading and writing achievement, and with increasing demands on teachers to teach the conventional forms of SAE language to their students, the malleability of dialect patterns has become an important question. Theoretically, becoming bidialectal reflects learning new language behaviors. There is fifty or more years of research in the fields of communication disorders and second language learning in the United States showing that structural and pragmatic language behaviors can be taught. However, teachers in Cyprus have found that repeated correction of Greek Cypriot dialect has yielded discouraging results. Students find it difficult to express themselves in standard Modern Greek, and their standardized test scores remain low when tests presented in Modern Greek are used to measure educational achievement (Pavlou & Papapavlou, 2004; Papapavlou & Pavlou, 2007).

It is noteworthy that although small in number, research reports of attempts to teach SAE have been successful with African American students. These studies have included third and fourth graders (Fogel & Ehri, 2000); fourth, fifth, and sixth graders (Sweetland, 2007); and kindergarteners and first graders (Craig, 2014). Some programs have used storybooks written in dialect to help students begin to read in SAE (Labov, 1995), and some have emphasized metalinguistic awareness in their methods (LeMoine, 2001; Wheeler & Swords, 2010; Wolfram, 1999). Others combined teaching dialect awareness with the production of common AAE features (Craig, 2014; Sweetland, 2007). These teaching programs have identified a core set of principles that were critical to their effectiveness and that can be applied to the creation and refinement of future interventions. These principles are the following.

The programs need to teach dialect recognition skills as a first and foundational starting point, prior to focusing on the production of standard alternatives to dialectal forms. Students need to learn to recognize AAE and SAE forms in a positive and constructive set of classroom interactions and to distinguish in which contexts each dialect should be used. When a positive focus on dialect awareness is not part of the emerging understanding of language differences, student self-efficacy can decrease (Fogel & Ehri, 2000). However, self-confidence can increase when dialect awareness is incorporated into the intervention (Sweetland, 2007).

The teaching of older students and adults typically has framed the concept of dialect differences in the context of racial differences in which AAE is presented as a dialect used by many African Americans (Sweetland, 2007; Wheeler & Swords, 2006). However, teaching dialect awareness and recognition to elementary grade students is also successful when using a formality metaphor (Craig, 2014; Wheeler & Swords, 2006). The terms "formal" and "informal" easily lend themselves to concrete real-life experiences, and AAE can be linked to informal everyday language and SAE to formal school discourse. The formality metaphor avoids the inevitable introduction of concepts like prejudice and bigotry when the students are so young. In our program of research (Craig, 2014) we developed methods, activities, and materials that efficiently moved students through lessons associating informality and formality to clothing (e.g., play clothes/uniform), places (e.g., home/school), and types of talk (casual/very polite; AAE/SAE).

Contrastive analysis (CA) methods compare and contrast students' home language with SAE and teach students to translate highly frequent features from AAE to SAE. Two building blocks constitute key instructional contexts for CA: instructional activities, which are the lesson types and formats in which CA is presented, and instructional discourse types, which are the modeling and feedback of teachers.

A major challenge for program development when teaching style shifting between AAE and SAE is to modify and adapt familiar lesson activities to the learning of bidialectalism. The English language arts time within most classrooms lends itself well to the types of language activities involved in style shifting teaching programs. Using existing class time in a modified way can still meet established goals and practices, such as vocabulary building, but include a focus on code differences between AAE and SAE within the same lessons and without adding time to the teacher's established schedule.

Instructional discourse in CA is characterized by constructive teacher comments and is respectful. The teacher's communications consistently provide models and feedback that do not devalue the students' home community language forms but make formal SAE alternatives highly salient. Teacher corrections consistently emphasize that dialects are adapted to the discourse context in which the student is engaged, and not because they are inherently good or bad. Teacher feedback is not critical in tone, but is positive and affirming (Craig, 2014; Wheeler & Swords, 2006).

Gaining teacher support and compliance have been serious problems in prior attempts to develop contrastive analysis programs (e.g., Ai, 2002). However, when the dialect teaching curriculum can be integrated into regular classroom pedagogy, avoiding the creation of additional work for the teacher and minimizing the amount of instruction in linguistics prior to implementing the program, the curriculum has been well accepted (Craig, 2014). The English language arts curriculum, comprising activities such as vocabulary building, sentence construction, and storybook reading, is an ideal context for the application of CA.

It is not yet clear how much linguistic competence teachers need in order to teach students to style shift to SAE. Wheeler and Swords (2004) provide teachers with an assessment strategy to learn about the dialect forms used by their students, and grammar lessons for teaching written language forms of SAE for students in grades three to six. These become powerful mini-lessons for the teachers about AAE features and their SAE alternatives. Alternatively, teachers can apply highly structured and scripted lesson sequences without a deep knowledge of specific dialects or the rules governing AAE (Craig, 2014).

Although this body of published literature remains small, it demonstrates that dialect change can be taught to students in all grades. Linguistic competence in SAE can be taught to later elementary, middle school, and older students, resulting in improved reading achievement (Sweetland, 2007; Wheeler & Swords, 2006). Until recently, these methods were only suitable for older students who have the cognitive maturity to think abstractly about their language forms, making explicit comparisons between the grammars of SAE and AAE. Methods are now available to teach AAE-speaking kindergarten and first-grade students to become bidialectal by scaling down contrastive analysis methods to be appropriate for students as they begin formal instruction in reading and writing (Craig, 2014). Rather than waiting until students are older and falling behind academically, dialect teaching can be incorporated into routine classroom activities. Early teaching programs can be grounded in the processes involved in socializing beginning students to the community of the classroom and academic learning.

Research shows that increased production of SAE corresponds to better reading achievement. Unfortunately, this body of research falls short of linking dialect changes causally to gains in reading and writing achievement. Randomized controlled trials will be required to establish the efficacy of dialect teaching programs, and pre- and postmeasures developed to evaluate changes in AAE feature production following dialect teaching.

Summary and Conclusions

Many African American students who live in large, highly segregated urban centers speak the dialect of AAE when they begin formal public schooling. Students who fail to achieve competence in SAE across the elementary grades show greater difficulty in reading and in other school subjects than their peers who become bidialectal. The greater the amounts of AAE feature production characterizing a student's discourse, the lower the reading scores. A number of variables influence the likelihood of an AAE feature or SAE alternative being produced at any time. Recent research has shown that students who are skilled at style shifting from AAE to SAE outperform their peers who are not able to make this language adaptation. Interest is increasing in the development of dialect teaching programs to help teachers help their students learn to style shift. These new research directions have considerable promise, and positive outcomes will offer an important new direction in the search for ways to resolve the black–white achievement gap.

References

Adler, S. (1992). *Multicultural communication skills in the classroom*. Boston, MA: Allyn & Bacon.

Ai, X. (2002). *Academic English Mastery Program 2000–2001 evaluation report* (Planning, Assessment, and Research Division Publication No. 111). Los Angeles, CA: Los Angeles Unified School District, Program Evaluation and Research Branch.

Alrabaa, S. (1986). Diglossia in the classroom: The Arabic case. *Anthropological Linguistics*, *28*, 73–79.

Aud, S., Fox, M. A., & KewalRamani, A. (2010). *Status and trends in the education of racial and ethnic groups* (NCES 2010–015). Washington, DC: US Department of Education, National Center for Education Statistics.

Battle, D. E. (1996). Language learning and use by African American children. *Topics in Language Disorders*, *16*, 22–37.

Baugh, J. (1983). *Black Street English: Its history, structure, and survival*. Austin: University of Texas Press.

Bell, A. (1984). Language style as audience design. *Language in Society*, *13*, 145–204.

Blom, J., & Gumperz, J. J. (1972). Social meaning in linguistic structures: Code-switching in Norway. In J. Gumperz & D. Hymes (Eds.), *Directions in sociolinguistics: The ethnography of communication* (pp. 407–434). New York, NY: Holt, Rinehart, and Winston.

Bountress, N. G. (1983). Effect of segregated and integrated educational settings upon selected dialectal features. *Perceptual and Motor Skills*, *57*, 71–78.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*, 371–399.

Brooks-Gunn, J., Klebanov, P., & Duncan, G. (1996). Ethnic differences in children's intelligence test scores: Role of economic deprivation, home environment, and maternal characteristics. *Child Development 67*, 396–406.

Cecil, N. L. (1988). Black dialect and academic success: A study of teacher expectations. *Reading Improvement*, *25*, 34–38.

Charity, A. H. (2007). Regional differences in low SES African-American children's speech in the school setting. *Language Variation and Change*, *19*, 281–293.

Charity, A. H., Scarborough, H. S., & Griffin, D. M. (2004). Familiarity with School English in African American children and its relation to early reading achievement. *Child Development*, *75*, 1340–1356.

Connor, C. M., & Craig, H. K. (2006). African American preschoolers' language, emergent literacy skills, and use of African American English: A complex relation. *Journal of Speech, Language, and Hearing Research*, *49*, 771–792.

Craig, H. K. (2014). *ToggleTalk™:Bi-dialectal fluency for young children*. Sun Prairie, WI: Ventris Learning.

Craig, H. K., Connor, C. M., & Washington, J. A. (2003). Early positive predictors of later reading comprehension for African American students: A preliminary investigation. *Language, Speech, and Hearing Services in Schools*, *34*, 31–43.

Craig, H. K., & Grogger, J. T. (2012). Social and style constraints on adult usage of African American English variants. *Journal of Speech, Language, and Hearing Research*, *55*,1274–1288.

Craig, H. K., Kolenic, G. E., & Hensel, S.L. (2014). African American English speaking students: A longitudinal examination of style shifting from kindergarten through second grade. *Journal of Speech, Language, and Hearing Research*, *57*, 143–157.

Craig, H. K., Thompson, C. A., Washington, J. A., & Potter, S. L. (2003). Phonological features of child African American English. *Journal of Speech, Language, and Hearing Research*, *46*, 623–635.

Craig, H. K., Thompson, C. A., Washington, J. A., & Potter, S. L. (2004). Performance of elementary grade African American students on the Gray Oral Reading Tests. *Language, Speech, and Hearing Services in Schools*, *35*, 141–154.

Craig, H. K., & Washington, J. A. (2004). Grade-related changes in the production of African American English. *Journal of Speech, Language, and Hearing Research*, *47*, 450–463.

Craig, H. K., & Washington, J. A. (2006). *Malik goes to school: Examining the language skills of African American students from preschool–5th grade*. Mahwah, NJ: Erlbaum.

Craig, H. K., Zhang, L., Hensel, S. L., & Quinn, E. J. (2009). African American English speaking students: An examination of the relationship between dialect shifting and reading outcomes. *Journal of Speech, Language, and Hearing Research*, *52*, 839–855.

Cunningham, P. M. (1976/1977). Teachers' correction responses to black-dialect miscues which are non-meaning-changing. *Reading Research Quarterly*, *12*, 637–653.

Dictionary of American Regional English online (2014). http://dare.wisc.edu/

Erikson, F. (1987). Transformation and school success: The politics and culture of educational achievement. *Anthropology and Education Quarterly*, *18*, 335–356.

Fogel, H., & Ehri, L. (2000). Teaching elementary students who speak Black English Vernacular to write in standard English: Effects of dialect transformation practice. *Contemporary Educational Psychology*, *25*, 212–235.

Fordham, S. (1999). Dissin' "the Standard": Ebonics and guerilla warfare at Capitol High. *Anthropology and Education Quarterly, 30*, 272–293.

Goodman, K. S., & Buck, C. (1973). Dialect barriers to reading comprehension revisited. *The Reading Teacher*, *27*, 6–12.

Green, L. J. (2002). *African American English: A linguistic introduction*. Cambridge, England: Cambridge University Press.

Green, L.J. (2011). *Language and the African American Child*. Cambridge: Cambridge University Press.

Grogger, J. (2011). Speech patterns and racial wage inequality. *Journal of Human Resources*, *46*, 1–25.

Harber, J. R. (1977). Influence of presentation dialect and orthographic form on reading performance of Black, inner-city children. *Educational Research Quarterly*, *2*, 9–16.

Hoffman, K., & Llagas, C. (2003). *Status and trends in the education of Blacks* (NCES 2003-034). Washington, DC: US Department of Education, National Center for Education Statistics.

Horton-Ikard, R. & Miller, J. F. (2004). It is not just the poor kids: The use of AAE forms by African-American school-aged children from middle SES communities. *Journal of Communication Disorders*, *37*, 467–487.

Ibrahim, M. H. (1983). Linguistic distance and literacy in Arabic. *Journal of Pragmatics*, *7*, 507–515.

Isaacs, G. J. (1996). Persistence of non-standard dialect in school-age children. *Journal of Speech and Hearing Research*, *39*, 434–441.

Ivy, L. J., & Masterson, J. J. (2011). A comparison of oral and written English styles in African American students at different stages of writing development. *Language, Speech, and Hearing Services in Schools*, *42*, 31–40.

Jencks, C., & Phillips, M. (Eds.). (1998). *The black–white test score gap*. Washington, DC: Brookings Institution Press.

Koch, L. Gross, A., & Kolts, R. (2001). Attitudes toward Black English and code switching. *Journal of Black Psychology*, *27*, 29–42.

Kohler, C. T., Bahr, R. H., Silliman, E. R., Bryant, J. B., Apel, K., & Wilkinson, L. C. (2007). African American English dialect and performance on nonword spelling and phonemic awareness tasks. *American Journal of Speech-Language Pathology*, *16*, 157–168.

Labov, W. (1972). *Language in the inner city*. Philadelphia: University of Pennsylvania Press.

Labov, W. (1995). Can reading failure be reversed? A linguistic approach to the question. *Literacy Among African-American Youth*, 39–68.

Labov, W. (2007). Transmission and diffusion. *Language*, *83,* 344–387.

Lee, V. E., & Burkham, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school.* Washington, DC: Economic Policy Institute.

LeMoine, N. (2001). Language variation and literacy acquisition in African American students. In J. L. Harris, A. G. Kamhi, & K. E. Pollock (Eds.), *Literacy in African American communities* (pp. 169–194). Mahwah, NJ: Erlbaum.

Lewis Mumford Center for Comparative Urban and Regional Research. (2002). Segregation of whole population, Census 2000 data. http://mumford.albany.edu/census

Linnes, K. (1998). Middle-class AAVE versus middle-class bilingualism: Contrasting speech communities. *American Speech*, *73*, 339–367.

Lippi-Green, R. (1997). *English with an accent.* New York, NY: Routledge.

Long john. (n.d.). In *Dictionary of American Regional English online*. http://dare.news.wisc.edu/state-by-state/Michigan/

Maamouri, M. (1998). *Language education and human development: Arabic diglossia and its impact on the quality of education in the Arab region*. Discussion paper prepared for the World Bank. Philadelphia: International Literacy Institute, University of Pennsylvania.

Markham, L. R. (1984). "De dog and de cat": Assisting speakers of Black English as they begin to write. *Young Children*, *39,* 15–24.

Massey, D. & Lundy, G. (2001). Use of Black English and racial discrimination in urban housing markets: New methods and findings. *Urban Affairs Review*, *36*, 452–469.

Mayer, M. (1969). *Frog Where Are You??* New York, NY: Dial.

McLoyd, V. C. (1990). The impact of economic hardship on black families and children: Psychological distress, parenting, and socioemotional development. *Child Development*, *61*, 311–346.

Mufwene, S. (2001). *The ecology of language evolution*. Cambridge, England: Cambridge University Press.

Muñoz, M. L., Gillam, R. B., Peña, E. D., & Gulley-Faehnle, A. (2003). Measures of language development in fictional narratives of Latino children. *Language, Speech, and Hearing Services in Schools*, *34*, 332–342.

National Center for Education Information. (2011). Profile of teachers in the US. www.ncei.com/Profile_Teachers_US_2011.pdf.

National Center for Education Statistics. (2011). *The nation's report card: Reading 2011* (NCES 2012-457). Institute of Education Sciences, US Department of Education, Washington, DC.

National Center for Education Statistics. (2012). *The nation's report card: Writing 2011* (NCES 2012–470). Institute of Education Sciences, US Department of Education, Washington, DC.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common core state standards English Language Arts standards*. National Governors Association Center for Best Practices, Council of Chief State School Officers, Washington, DC.

Nievar, M. A., & Luster, T. (2006). Developmental processes in African American families: An application of McLoyd's theoretical model. *Journal of Marriage and Family*, *68*, 320–331.

Oetting, J. B., & McDonald, J. L. (2002). Methods for characterizing participants' nonmainstream dialect use within studies of child language. *Journal of Speech*, *Language*, *and Hearing Research*, *45*, 505–518.

Papapavlou, A., & Pavlou, P. (2007). The interplay of bidialectalism, literacy and educational policy. In A. Papapavlou & P. Pavlou (Eds.), *Sociolinguistic and pedagogical dimensions of dialects in education* (pp. 101–121). Newcastle, England: Cambridge Scholars.

Pavlou, P., & Papapavlou, A. (2004). Issues of dialect use in education from the Greek Cypriot perspective. *International Journal of Applied Linguistics*, *14*, 243–258.

Preston, D. (1991). Sorting out the variables in sociolinguistic theory. *American Speech*, *66*, 33–56.

Purnell, T., Idsardi, W., & Baugh, J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology*, *18*, 10–30.

Renn, J., & Terry, J. M. (2009). Operationalizing style: Quantifying the use of style shift in the speech of African American adolescents. *American Speech*, *84*, 367–390.

Rickford, J. R. (1998). The Creole origins of African-American vernacular English: Evidence from copula absence. In S. S. Mufwene, J. R. Rickford, G. Bailey, & J. Baugh (Eds.), *African-American English: Structure, history and use* (pp. 154–200). London, England: Routledge.

Rickford, J. R., & McNair-Knox, F. (1994). Addressee- and topic-influenced style shift: A quantitative sociolinguistic study. In D. Biber & E. Finegan (Eds.), *Sociolinguistic perspectives on register* (pp. 235–276). New York, NY: Oxford University Press.

Rodriguez, J., Cargile, A., & Rich, M. (2004). Reactions to African-American vernacular English: Do more phonological features matter? *Western Journal of Black Studies*, *28*, 407–414.

Ross, S. H., Oetting, J. B., & Stapleton, B. (2004). Preterite *had* + V-*ed*: A developmental narrative structure of African American English. *American Speech*, *79*, 167–193.

Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics*, *24*, 431–451.

Schachter, R., & Craig, H. K. (2013). Students' production of narrative and AAE features during an emergent literacy task. *Language*, *Speech*, *and Hearing Services in Schools*, *44*, 227–238.

Seymour, H. N., & Ralabate, P. K. (1985). The acquisition of a phonologic feature of Black English. *Journal of Communication Disorders*, *18*, 139–148.

Smitherman, G. (1974). Soul 'n style. *English Journal*, *63*, 14–15.

Smitherman, G. (2000). *Blacktalk: Words and phrases from the hood to the amen corner* (2nd ed.). Boston, MA: Houghton Mifflin.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, *38*, 934–947.

Sweetland, J. (2007). Teaching writing in the African American classroom: A sociolinguistic approach. *Dissertations Abstracts International*, *67* (09). UMI No. 3235360.

Terry, N. P. (2006). Relations between dialect variation, grammar, and early spelling skills. *Reading and Writing: An Interdisciplinary Journal*, *19*, 907–931.

Terry, N. P. (2014). Dialect variation and phonological knowledge: Phonological representations and metalinguistic awareness among beginning readers who speak nonmainstream American English. *Applied Psycholinguistics*, *35*, 155–176.

Terry, N. P., & Connor, C. M. (2012). Changing nonmainstream American English use and early reading achievement from kindergarten to first grade. *American Journal of Speech-Language Pathology*, *21*, 78–86.

Terry, N. P., Connor, C. M., Petscher, Y., & Conlin, C. R. (2012). Dialect variation and reading: Is change in nonmainstream American English use related to reading achievement in first and second grades? *Journal of Speech*, *Language*, *and Hearing Research*, *55,* 55–69.

Terry, N. P., & Scarborough, H. S. (2011). The phonological hypothesis as a valuable framework for studying the relation of dialect variation to early reading skills. In S. A. Brady, D. Braze, & C. A. Fowler (Eds.), *Explaining individual differences in reading: Theory and evidence* (pp. 97–120). New York, NY: Taylor & Francis.

Thompson, C. A., Craig, H. K., & Washington, J. A. (2004). Variable production of African American English across oracy and literacy contexts. *Language*, *Speech*, *and Hearing Services in Schools*, *35*, 269–282.

United States Census Bureau. (2011). *School enrollment* [data file]. www.census.gov/hhes/school/data/cps/2011/tables.html

US Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessments.

US Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2013 Reading Assessments.

Van Hofwegen, J., & Wolfram, W. (2010). Coming of age in African American English: A longitudinal study. *Journal of Sociolinguistics*, *14*, 427–455.

Washington, J. A. (2001). Early literacy skills in African-American children: Research considerations. *Learning Disabilities Research and Practice*, *16*, 213–221.

Washington, J. A., & Craig, H. K. (1998). Socioeconomic status and gender influences on children's dialectal variations. *Journal of Speech*, *Language*, *and Hearing Research*, *41*, 618–626.

Washington, J. A., & Craig, H. K. (2001). Reading performance and dialectal variation. In J. Harris, A. Kamhi, & K. Pollock (Eds.), *Literacy in African American Communities* (pp. 147–168). Mahwah, NJ: Erlbaum.

Washington, V. M., & Miller-Jones, D. (1989). Teacher interactions with nonstandard English speakers during reading instruction. *Contemporary Educational Psychology*, *14*, 280–312.

Wheeler, R. S., & Swords, R. (2004). Codeswitching: Tools of language and culture transform the dialectally diverse classroom. *Language Arts*, *81*, 470–480.

Wheeler, R. S., & Swords, R. (2006). *Code-switching: Teaching Standard English in urban classrooms.* Urbana, IL: National Council of Teachers of English.

Wheeler, R., & Swords, R. (2010). *Code-switching lessons : Grammar strategies for linguistially diverse writers.* Portsmouth, NH: Firsthand.

Wolfram, W. (2004). Social varieties of American English. In E. Finegan & J. R. Rickford (Eds.), *Language in the USA: Themes for the twenty-first century* (pp. 58–75). Cambridge, England: Cambridge University Press.

Wolfram, W. (1999). Repercussions from the Oakland Ebonics controversy—The critical role of dialect awareness programs. In C. Adger, D. Christian, & O. Taylor (Eds.), *Making the connection: Language and academic achievement among African American students* (pp. 61–80). Mahwah, NJ: Erlbaum.

Wolfram, W., Adger, C. T., & Christian, D. (1999). *Dialects in schools and communities*. Mahwah, NJ: Erlbaum.

CHAPTER

29 Teachers' Knowledge About Beginning Reading Development and Instruction

Anne E. Cunningham *and* Colleen Ryan O'Donnell

Abstract

This chapter focuses on the body of disciplinary and pedagogical knowledge required to provide high-quality beginning reading instruction to young children. The chapter examines quality literacy instruction from a historical perspective, reviews scientific evidence on the successful teaching of reading, explores why teachers are not consistently teaching beginning reading in ways that are aligned with best practices, and provides recommendations for how the field can support teachers in developing the knowledge needed to improve student reading outcomes. The goal is to provide research-based suggestions for strengthening both the content and delivery of teacher professional development in the area of literacy and to demonstrate that these suggestions have the power to affect child outcomes.

Key Words: teacher knowledge, teacher education, professional development, reading, word recognition

> Those who can, do; those who understand, teach.
>
> —Shulman, 1986

In most academic subjects, it is obvious to the layperson why teachers need disciplinary competence in order to be effective. For example, most people would agree that it would be difficult to provide high-quality instruction about the principles and concepts of physics without deep knowledge and understanding of those principles and concepts. A physics teacher must be able to guide students in the creation of detailed conceptual frameworks, respond to student inquiries, and provide nuanced clarifications—tasks that would be impossible without deep understanding of the field. However, the need for discipline-specific knowledge or competence can be less obvious when considering a teacher's ability to provide high-quality instruction in more fundamental academic tasks such as reading. It is easy to assume that being a skilled reader creates a sufficient knowledge base for providing reading instruction. Although the connection may be less obvious, content-specific knowledge may be particularly important in the teaching of fundamental academic skills such as reading and associated skills including spelling and writing (Brady & Moats, 1997; Wong-Fillmore & Snow, 2003). A convincing body of empirical research provides strong evidence that successful reading instructors need to have highly specialized skills and knowledge—skills akin to those required of a physics teacher, but specific to literacy (Connor, Son, Hindman, & Morrison, 2005).

High-quality reading instruction is partially defined by the knowledge that teachers of reading must possess to provide effective instruction for their students (Snow, Burns, & Griffin, 1998; National Early Literacy Panel, 2008; National Reading Panel, 2000). The research suggests that, like teachers of physics, teachers of reading require domain-specific knowledge and expertise—expertise, for example, in the language of instruction, knowledge about reading development and its component skills, and the ability to use that knowledge in educational encounters with children (Connor et al., 2005; Cunningham, Etter, Platas, Wheeler, & Campbell, 2015; Cunningham & Zibulsky, 2009; Foorman & Moats, 2004).

The content knowledge required for effective instruction and intervention in the United States includes knowledge of the American English spelling system. English is a morphophonemic or deep alphabetic orthography (Venezky, 1999), which means that its spelling is bound by meaning (as in *magician*) as well as sound (as in *magic*). Although its spellings map onto speech sounds quite predictably, especially for words encountered during the earliest years of reading instruction, the correspondences can be complex and variable. In order to provide explicit and complete explanations of both predictable and less predictable relationships (only some of which are caused by meaning overriding predictable sound–symbol correspondences), teachers must be knowledgeable about the complex English spelling system (Moats, 1994; Wong-Fillmore & Snow, 2003). Because decoding problems underlie the difficulties of most primary grade students with reading problems (Catts, Hogan, & Adlof, 2005), explicit and accurate word recognition instruction is necessary. Instruction about sound–symbol correspondences is particularly important, as is instruction about less predictable words that are high in frequency (such as *was* or *from*). Knowledge of the spelling system, along with facility in methods known to be effective in teaching it, is fundamental background knowledge for teachers. Thus, a prerequisite knowledge base for the delivery of high-quality beginning reading instruction in the United States must include understanding reading development, linguistic concepts, and features of the English language and its spelling. This type of specialized disciplinary knowledge, referred to as *pedagogical content knowledge* by Shulman (1987), captures the amalgam of disciplinary knowledge and pedagogy needed to effectively teach a skill such as literacy.

Equipped with the understanding that there is a prerequisite knowledge base for the delivery of quality literacy instruction, a complementary question is whether the average teacher recognizes the need for these competencies, possesses the required knowledge, and values the pursuit of this knowledge. The goal of this chapter is to review the literature regarding the disciplinary and pedagogical knowledge that is necessary to teach beginning reading in English to young children (see Goldman & Snow, this volume, for discussion of teaching for adolescents), what teachers know, what information they lack, and what further opportunities they need in order to acquire this critical knowledge. We begin by examining quality reading instruction from a historical perspective, review the scientific evidence on the successful teaching of reading, and then define and contextualize aspects of learning to read and discuss teachers' need for knowledge of the American English spelling system. We then outline the prerequisite knowledge base for the delivery of quality beginning literacy instruction, explore why teachers are not consistently teaching reading in ways that are aligned with best practices, and provide recommendations for how to support teachers in developing knowledge needed to improve students' reading outcomes.

Defining Quality Reading Instruction

High-quality reading instruction makes a difference in the literacy development and performance of students, and in this section we discuss the nature of such instruction.

Historical Perspectives on Quality Reading Instruction

To provide context for our discussion of the current understanding of quality reading instruction, a brief historical perspective on the factors that have traditionally driven reading instruction in the United States may be helpful. For the better part of the twentieth century, prominent figures in education debated vehemently about the most effective way to teach children to read (Chall, 1967, 1992; Stanovich & Stanovich, 1995). Theorists and educators generally adopted one of two perspectives with respect to their thinking about how children learn to read and what unit of language (i.e., the sentence, word, or phoneme) should be the focus of instruction. In one camp were those who advocated a *whole- language* approach, arguing that learning to read is analogous to learning to speak and that the most effective means of teaching children to read is to immerse them in print, eschewing more analytic approaches (Goodman, 1986; Smith, 1971). In the other camp were those who subscribed to the skill-based and more analytic *phonics* approach, which involves the direct teaching of letter–sound correspondences and combinations of letters and their corresponding sounds. Those in this camp emphasized the importance of providing children with direct instruction in the *alphabetic principle*. According to the alphabetic principle, letters and combinations of letters are the symbols used to represent the speech sounds of a language based on systematic and predictable relationships between written letters, symbols, and spoken words (Adams, 1990; Bond & Dijkstra, 1967; Chall, 1967; Ehri, this volume).

Toward a Scientific Definition of Quality Reading Instruction

Because the field of education in the 1970s and 1980s was yet to be influenced by the idea that instruction should be guided by scientific inquiry and converging empirical evidence, educators in the United States were easily persuaded by movements that were driven predominantly by deeply rooted philosophical perspectives, observation, and personal experience (Stanovich, 2000). Across much of his writing, Stanovich (e.g., 1993) maintained that a reliance on a political/ideological rather than a scientific model for making instructional decisions has created many problems for reading education. He argued that the extreme pendulum swings that have characterized reading education might be avoided by equipping teachers with a scientific model of decision-making. Concomitantly, the end of the twentieth century brought nationwide concern in the United States regarding academic achievement, especially among disadvantaged students (Lyon, 1999a; National Assessment of Educational Progress, 1995). As a result, there was a push to end the reading wars and identify, from a scientific perspective, the most effective approaches to reading instruction.

Research Regarding the Critical Features of a Quality Reading Curriculum

One of the most notable large-scale investigations in the United States was conducted by the National Research Council (NRC) in 1998 (Snow, Burns, & Griffin, 1998). Noting the increasing demand for literacy in a technologically advanced society and the repercussions for those who have low levels of literacy, the US Department of Education and the US Department of Health and Human Services asked the National Academy of Sciences to establish a committee focused on determining, from an empirical perspective, how to best support the reading development of children and prevent reading difficulties. The committee reviewed the research on reading development and instruction, the factors associated with reading failure, and the interventions and instructional approaches known to prevent reading difficulties and promote optimal reading outcomes. In summarizing the results of their research, the committee argued that more focus must be placed on improving the quality of reading instruction for both struggling readers and beginning readers. They noted that, although the needs of struggling readers vary depending on their skills and abilities, effective teachers use evidence-based materials and strategies to craft an appropriate mix of learning opportunities for every student. The NRC report argued that the ability to craft an ideal combination of instructional techniques requires, at a minimum, deep knowledge and understanding of reading development as well as familiarity and facility with the pedagogical strategies known to be most effective in supporting reading development (or remediating delay).

Multiple skills have been shown to be essential for successful reading acquisition. These include *phonological awareness* (the ability to detect and manipulate the sounds or phonemes in language), *print knowledge* (the combination of elements of alphabet knowledge, concepts about print, and early decoding), *fluency* (the ability to quickly and efficiently process text), *vocabulary, background knowledge*, and *comprehension* (the ability to derive meaning from written text). For beginning reading acquisition, the NRC report (Snow et al., 1998) highlighted the importance of accurate word identification and the role of explicit instruction to help children develop an appreciation for the sound structure of language to facilitate decoding. This included knowledge of specific letter–sound correspondences, common spelling patterns, and high-frequency irregular words. Additionally, according to the NRC report, repeated opportunities to practice both silent and oral reading of high-quality engaging texts promote reading fluency. Moreover, when children receive explicit instruction in comprehension strategies, their understanding of texts is facilitated. Finally, the facilitative effects of reading exposure across a wide variety of topics provide distributed practice that promotes fluency and reading comprehension. Such exposure across a variety of texts further promotes children's vocabulary and conceptual knowledge.

Research Regarding the Pedagogical Methods Known to be Effective in Supporting Student Learning

The report of the NRC (Snow et al., 1998) provides more than a theoretical approach to reading instruction. It draws attention to the fact that teachers must have deep understanding of the process of reading development and provides a synthesis of the research supporting the claim that teachers must be able to provide quality instruction in five areas: phonological awareness, phonics, fluency, comprehension, and oral language development including vocabulary.

Although the report of the NRC (Snow et al., 1998) provides an analysis of the skills, environments, and experiences that are critical to the acquisition of reading, the committee did not address the specific instructional approaches that are most efficacious in bringing about positive outcomes. Thus, Congress asked the National Institute of Child Health and Human Development (NICHD), in collaboration with the Secretary of Education, to convene a panel of experts to review the research on the effectiveness of common approaches to teaching children to read. The National Reading Panel (NRP, 2000) engaged in a comprehensive review of the major variables found to contribute to skilled reading. Based on the consensus synthesis of the NRC, the NRP (2000) focused on research pertaining to word recognition (i.e., phonological awareness, learning the alphabetic principle through phonics instruction, vocabulary, fluency, and comprehension). This work differed from the NRC report in that it attempted to synthesize experimental and quasi-experimental work in reading instruction through a meta-analysis of the research on reading instruction. Meta-analyses are effective tools for summarizing the research in a specific area in that they provide a statistical analysis of the results of multiple individual studies and integrate findings more rigorously than traditional narrative or descriptive review methods (Glass, 1976).

The NRP (2000) concluded that there was an impressive body of converging evidence in the area of the word recognition to guide the field. The meta-analysis suggested that specific skills must be mastered in the course of reading development and that not all strategies or forms of instruction are effective for all students at all levels of development. For example, the NRP found that teaching children to manipulate phonemes in words was highly effective and that teaching phonemic awareness improves reading significantly more than instruction that does not include instruction in segmenting and blending phonemes. Moreover, this finding extends to a variety of learners across a range of grade and age levels. Likewise, after reviewing thirty-eight independent studies on the teaching of phonics, the NRP (2000) found that systematic phonics instruction benefits students in kindergarten through sixth grade who are having difficulty learning to read. Yet phonics instruction had the greatest impact for students in kindergarten through second grade. The NRP argued, based on converging evidence from a variety of studies, that explicit, systematic phonics instruction that includes phonological awareness is an essential part of a beginning reading curriculum.

In addition to word recognition, the NRP (2000) reviewed data on reading fluency and comprehension. One instructional procedure they found to be highly effective was *guided repeated oral reading*. This approach encourages students to read passages orally with systematic and explicit guidance and feedback from teachers. The NRP reviewed sixteen studies of this approach and found that guided repeated oral reading procedures had significant benefits (weighted effect size average of 0.41) for the development of word recognition, fluency, and comprehension across a range of grade levels. Furthermore, these results apply to all students, including those having difficulty reading. The NRP also reviewed 205 independent studies of reading comprehension and found that text comprehension improved when readers actively related the ideas in texts to their own knowledge. There was substantial evidence to suggest the need for direct instruction in text comprehension strategies such as questioning, summarizing with words and pictures, drawing maps of stories, cooperative work, and monitoring one's own comprehension. It was determined that a combination of comprehension strategies was most effective. Supporting reading comprehension through vocabulary development—both explicit teaching of vocabulary and incidental exposure to vocabulary—was also found to be critical.

Across a number of domains, the National Reading Panel (2000) provided information about which instructional strategies were most effective at which level of development, for which specific reading skills, and with which types of students (e.g., typically developing children, children at risk for reading failure, and second language learners). These detailed findings highlight the importance of ensuring that teachers have a deep knowledge base that can be skillfully woven into acts of teaching—from explicit instructional opportunities to the ability to supply fruitful explanations, analogies, examples, and materials to each student at the right time. That is, if teachers are to be effective, they must be equipped to evaluate, understand, and respond to each student's instructional needs. When it comes to having demonstrated competence in reading instruction, it is insufficient for a teacher to be able to identify the essential components of an effective reading curriculum or define literacy-related terms. Instead, competence in teaching reading entails a wider variety of factors. In the area of word recognition, the primary focus of this chapter, teachers of beginning reading in the United States must possess a good knowledge about the American English spelling system to be able to provide this level of instruction (Moats, 1994; Brady & Moats, 1997).

The Need for a Deep Knowledge of the American English Spelling System

Expert teaching of reading requires knowledge of language structure, and in this section we discuss the nature of American English spelling and teachers' knowledge about it.

The Complex American English Spelling System

English is considered a deep orthography because it has a lower degree of letter–sound correspondence than many other alphabetic writing systems (Besner & Smith, 1992). This letter-sound irregularity leads to a complex spelling system, which is the main hurdle for beginning readers. The complex letter–sound system of English must be taught because it is not necessarily intuitive to beginning readers. Although describing each rule in the spelling system is beyond the scope of this chapter, two main concepts of the American English spelling system are discussed in what follows.

Cummings (1988) suggested that spelling rules are of at least two types, tactical and procedural. *Tactical rules* cover the rules for letter–sound correspondence and contextual constraints of spelling. For example, there are a number of ways to spell the /k/ sound (i.e., ‹c›, ‹k›, ‹ck›, ‹ch›, ‹q›, and ‹cq›), which, interestingly, depend on the context in which the phoneme occurs. *Procedural* rules govern the way prefixes, bases, and suffixes combine to form written words. In other words, procedural rules underlie the morphological structure of spelling. For example, the spelling of "running" is *run+n+ing* rather than *run+ing*, the doubling of the *n* serving to reinforce the pronunciation of the first syllable as closed (*run*) not open (*ru*). Cummings notes that "the important products of these tactical and procedural rules are *correspondences,* the conventionalized relationships that exist between sounds and their spellings" (p. 10). These relationships assist beginning readers learning new words, and instruction about these relationships should be based on a solid understanding of how the spelling system works (see Kessler & Treiman, this volume, and Kessler, Treiman, & Evans, 2007, for further discussion of the spelling system).

Knowledge of the Spelling System Is Not Intuitive for Teachers

The knowledge and skills required to implement an effective early literacy curriculum are not necessarily intuitive to skilled readers. Once a reader becomes fluent, attention moves away from code translation toward comprehension (Oakhill, Berenhaus, & Cain, this volume). The vast majority of teachers became skilled readers far too long ago to rely on their intuitive knowledge of phonology and orthography as guides for instruction. Knowledge of conventional spelling can obscure the ability to attend to language at the sound level, thus making a teacher believe, for instance, that /s/ is the third sound in the word *music* rather than /z/. Consequently, awareness of subtleties of word structure needed to guide students is, ironically, often obscured by the teacher's personal reading competencies.

In her seminal study, Moats (1994) found that teachers' knowledge of phonology, orthography, and morphology was "surprisingly poor." She was one of the first to suggest that many teachers "understand too little about spoken and written language structure to be able to provide sufficient instruction in these areas" (p. 81). This does not mean that teachers lack reading and spelling ability. Instead, her initial results suggested that teachers' own literacy does not guarantee them detailed insights into structural aspects of phonology, orthography, and morphology. For example, when administered a series of multiple choice questions, Moats (1994) found that only 27% of the teachers could successfully count the number of morphemes in a word (e.g., *salamander* = 1, *pies* = 2, *unbelievable* = 3). Also, only 10% could identify a consonant cluster (i.e., two or three consonants that blend to make a distinct consonant sound such as /skr/ in *scratch* and /st/ in *first*), and none consistently identified consonant digraphs (i.e., two consonant letters that together make a single sound such as /θ/ in *think*). There was likewise substantial evidence that many experienced teachers have misconceptions about the principles of grapheme–phoneme correspondence, such as the number of ways to spell /k/, the reason for doubling the ‹m› in words such as *comment* and *commitment*, and the way in which the following vowel signals whether the letter ‹g› is pronounced as /g/, /d͡ʒ/, or /ʒ/, as in *god, gem*, and *rouge*.

Subsequent studies provided converging evidence that teachers generally lack sufficient knowledge of many of the linguistic concepts needed to successfully teach beginning readers (e.g., Bos, Mather, Dickson, Podhajski, & Chard, 2001; Cunningham, Zibulsky, & Callahan 2009; Cunningham, Perry, Stanovich, Stanovich, & Chappell, 2001; Piasta, Connor, Fishman, & Morrison, 2009). These studies have similarly shown that teachers have difficulty counting phonemes and morphemes in words, recognizing phonetically irregular words, classifying words by syllable type (open as in *hi*, closed as in *him*, r-controlled as in *bird*, silent -e as in *mate*, vowel team as in *bread*, consonant-le, as in *little*), and understanding how syllable-division patterns affect pronunciation (for example, a vowel-consonant-vowel (VCV) sequence can be divided in two ways to produce different syllable types—VC/V, as in *wom/an* [the first syllable is closed] and V/CV, as in *hu/man* [the first syllable is open]). The implications are that teachers are limited in their ability to interpret and respond to students' errors, pick appropriate examples for teaching decoding and spelling, effectively organize and sequence instruction, use morphology to demystify various aspects of spelling, and integrate the components of literacy instruction (Cunningham, Zibulsky, Stanovich, & Stanovich, 2009; Moats, 1999).

Why Teachers Need Deep Knowledge of the American English Spelling System

To illustrate how limited knowledge of these concepts may result in weak instruction, consider the following examples. A teacher who cannot consistently count phonemes in a word might mistakenly believe that a child who spelled *exit* as ‹eksit› was not successfully representing each of the sounds in the spoken word in the spelling. A teacher who mistakenly believes that /θ/ is a blend of /t/ and /h/ might not provide corrective feedback to a student who attempts to sound out the written word *thin* by saying, /t/ /h/ /ɪ/ /n/. A teacher who does not know that ‹c› is pronounced like /s/ when followed by the letters ‹e›, ‹i›, and ‹y› is less equipped to support a student who haphazardly uses the sounds /k/ and /s/ when decoding. Or a teacher who is unable to consistently recognize an irregular word as irregular may cause confusion by including the word *give* as an example of a typical word with a silent ‹e› or encouraging a student to sound out the word *was*. These practices are problematic given that, as mentioned earlier, the ability to decode words is the source of reading difficulty among many beginning readers (Catts et al., 2005; Cunningham & Stanovich, 1997). Without training, many teachers cannot explain the underlying system of the English spelling system to students.

What Teachers Know About Literacy Instruction

As compared with disciplines such as mathematics and social studies, studies of teachers' declarative knowledge in the domain of literacy are not well developed. With only a few exceptions, the majority of research in this area has occurred within the last two decades (Bos et al., 2001; Brady et al., 2009; Cunningham et al., 2001; Mather, Bos, & Babur, 2001; McCutchen et al., 2002; Moats, 1994). There has recently been a substantial increase in the amount of research going beyond the documentation of knowledge levels to investigating the factors related to variations in teachers' knowledge as they relate to classroom practice (e.g., Cheesman, McGuire, Shankweiler, & Coyne, 2009; Cunningham, Perry, Stanovich, & Stanovich, 2004; Mather et al., 2001; Lopes, Spear-Swerling, Gabriela Velasquez, & Zibulsky, 2014; Spear-Swerling, Brucker, & Alfano, 2005).

For example, Mather et al. (2001) examined the potential impact of years of experience on teachers' attitudes and knowledge of effective classroom practices for the teaching of reading. They explored teachers' awareness of the importance of direct, explicit, code-based literacy instruction and their knowledge of phonics terminology, such as consonant blend, diphthong, digraph, and schwa. Specifically, 293 teachers attending a university to receive their teaching credential (preservice teachers) and 131 teachers in the field (in-service teachers) were asked to complete a rating scale inquiring about their beliefs about various practices in literacy instruction along with an assessment of their knowledge of the structure of language. Results of the study suggested that more experienced teachers generally had a more positive view of the role that explicit, code-based instruction plays in supporting the reading development of children. Mather and her colleagues hypothesized that this finding might indicate that experiences with beginning and struggling readers increased teachers' appreciation for the importance of code-based instruction. Less encouraging was that neither group of teachers had a clear understanding of the importance of letter–sound correspondences as a foundation for accurate word recognition. Instead, the large majority of teachers believed that the use of context was the most beneficial strategy for identifying an unknown word. Moreover, consistent with the findings of Moats (1994), neither preservice nor in-service teachers had sufficient knowledge of the phonological, orthographic, and morphological structures of the English language to effectively teach reading at a basic, code-based level. In other words, although experience may have supported teachers in the development of an appreciation for the importance of code-based instruction, it had less impact on the development of the knowledge required to successfully provide that instruction.

Do Teachers Know What They Don't Know?

It is counterintuitive to think that teachers might recognize the importance of a particular instructional approach yet fail to develop their own skills in a manner that would enable them to provide that instruction. However, this recognition may be an important precursor to knowledge gain, and it is therefore crucial to evaluate whether teachers are aware of what they do not know (Cunningham et al., 2004). Cunningham et al. argued that it is only when individuals recognize gaps in their knowledge that they are inclined to seek out and attend to the information they do not possess. Well-calibrated thinking about one's own knowledge has significant consequences in terms of how likely one is to improve the quality of instruction by targeting the areas of weakness through professional development.

To investigate this aspect of teachers' knowledge, Cunningham and her colleagues (2004) assessed the actual and perceived reading-related subject matter knowledge of 722 teachers of kindergarten through third-grade pupils. The researchers evaluated teachers' actual knowledge through the use of direct measures of knowledge of phonemic awareness (e.g., the number of speech sounds heard in the words *exit* and *sun*), and phonics (multiple choice questions related to concepts such as syllables and speech sounds). Perceived knowledge was evaluated by asking teachers to respond to the following questions: How would you describe your current skill level, based on past success, in your knowledge of children's literature, ability to provide instruction in phonemic awareness, and ability to provide instruction in phonics? Teachers were asked to make one of four choices: no experience, minimal skills, proficient, or expert. Based on their responses, two subgroups of teachers were identified for each category of knowledge. Consistent with previous research, teachers had limited knowledge in all these domains. However, this study made an additional contribution to the literature by showing the majority of teachers overestimated their levels of knowledge in word recognition but not children's literature. Teachers were particularly poorly calibrated in the essential domains of phonemic awareness and phonics, with the majority of kindergarten to third-grade teachers failing to recognize the limits of their knowledge of skills known to be critical to quality literacy instruction. Spear-Swerling et al. (2005) replicated this study and obtained similar findings for general and special educators (teachers who educate students with disabilities or special needs).

The limits of teachers' knowledge in this area are not unique to elementary school teachers. Given their involvement with children at a critical time of language development, teachers of children aged 3 to 5 in preschool represent an important bridge to literacy acquisition. They are in a position to begin helping children develop an awareness of the linguistic elements of language. As the demands of schooling and literacy increase, preschool teachers are increasingly called on to provide explicit and systematic instruction that helps students develop the phonological skills necessary for later efficient word recognition. Cunningham, Zibulsky, and Callahan (2009) examined the knowledge of early childhood educators to determine whether they possess the necessary competencies to guide the literacy development of their students. Similar to the results of studies investigating the knowledge of elementary school teachers, Cunningham, Zibulsky, and Callahan found that preschool teachers lack the disciplinary knowledge required to promote early literacy and also overestimate what they know. The researchers made the case that overconfidence in one's ability to teach young children essential language and literacy skills creates a potential obstacle for seeking additional information or professional training.

Attitudes Toward the Teaching of Phonics

The elusiveness of foundational concepts of language may affect teachers' attitudes about their instructional responsibilities. Cunningham, Zibulsky, Stanovich, et al. (2009) investigated first-grade teachers' priorities and preferences in beginning reading instruction, and showed that this group's preferred time allocation for instruction typically did not conform to models of reading instruction substantiated by the National Reading Panel report (2000). Cunningham, Zibulsky, Stanovich, et al. found that teachers preferred to spend their reading instruction time on literature-based activities and independent reading and writing. Although teachers with more knowledge of letter–sound correspondences were somewhat more inclined to spend time teaching phonics, the majority of teachers did not allocate their time in ways consistent with research recommendations. Surprisingly, even special education teachers' overall content knowledge was quite low and they did not favor intensive code-based instruction for struggling readers. These findings have been replicated by Lopes et al. (2014).

Impact of Teachers' Knowledge on Practices and Students' Outcomes

In response to evidence that teachers lack the knowledge to provide quality instruction in early literacy, research efforts turned toward investigations of the relationship between teachers' knowledge, classroom practices and students' outcomes. McCutchen et al. (2002) proposed that if knowledge of phonology is essential for children as they acquire literacy (Adams, 1990; Cunningham & Stanovich, 1997; National Reading Panel, 2000), then knowledge in this area must likewise be important for teachers. They examined the content knowledge of teachers of children ages 6 to 9 in phonology, evaluating the extent to which teachers' knowledge varied based on grade level or classroom placement (i.e., regular education vs. classrooms of children with special needs) and investigated the impact of teachers' knowledge on students' learning.

Results of this study provided further evidence that teachers at all grade levels generally lack sufficient knowledge of phonology (McCutchen et al., 2002). Furthermore, significant relationships were observed between teacher content knowledge and instructional practices. Across all three grades, teachers' reading-related content knowledge was related to their observed instructional practices. For example, knowledge of phonology was related to the instructional practices used in focusing children's attention on sounds and letter–sound relationships. However, in exploring the relationships between teachers' content knowledge and students' outcomes, significant relationships were only observed between kindergarten teachers' phonological knowledge and their students' reading achievement (i.e., the more teachers knew about phonology, the better the performance of their students). This link was not observed among first- and second-grade teachers. The researchers argued this finding may be due to methodological reasons. They also noted that the link between kindergarten teachers' phonological knowledge and their students' reading performance was troubling because overall, teachers' phonological knowledge was quite low.

Spear-Swerling and Brucker (2004) conducted a study involving 147 novice teachers of special education students. They examined the impact of teacher training and direct instruction emphasizing basic linguistic concepts on teachers' knowledge and the progress of their students in the development of basic reading and spelling skills. Teacher instruction was focused on the importance of systematic and explicit teaching of word decoding to beginning readers and children with reading difficulties. Teachers were also taught about the characteristics of language that are reflected in the writing system such as phonemes, graphemes, and morphemes. Additional central concepts included phonemic awareness, the role of orthographic and morphemic units in reading and spelling, common syllable types in English, multi syllable words, and common phonetically irregular words. The results of the study suggested that teachers who received direct instruction about the English spelling system had greater knowledge of how writing reflects language than teachers who had not received such instruction. A similar pattern of results were observed when McCutchen, Green, Abbott, and Sanders (2009) trained a sample of teachers of older students (ages 10–12). They found that teachers' linguistic knowledge uniquely predicted lower-performing students' end-of-year performance in reading, spelling, writing, and vocabulary.

Although some of the data suggest a link between teachers' knowledge and students' outcomes, investigations of this nature have also produced null results. For example, two experimental studies by Carlisle and her colleagues (Carlisle, Correnti, Phelps, & Zeng, 2009; Carlisle, Kelcey, Rowan, & Phelps, 2011) demonstrated only small and nonsignificant relationships between teachers' knowledge of early literacy and students' performance on tests of decoding, word recognition and reading comprehension. In interpreting these results, the researchers point to methodological weaknesses in the field's approach to studying teachers' knowledge and to the complexity of the factors that influence teachers' acquisition of knowledge. For example, Carlisle and colleagues highlight the potential for limited alignment between measures of teachers' knowledge (e.g., knowledge of linguistics) and their knowledge of how to effectively embed this information into their reading instruction (e.g., understanding of how to effectively use knowledge of linguistics in instructing students, or what Shulman, 1987, described as pedagogical content knowledge).

Piasta et al. (2009) sought to examine these links between teachers' knowledge, classroom practice, and students' growth. They examined first-grade teachers' knowledge about early literacy concepts such as phonological awareness and the alphabetic principle, the amount and type of decoding instruction teachers provide, and their students' outcomes. Piasta et al. hypothesized that teachers' knowledge impacts students' outcomes through the type of instruction teachers provide. Their results suggested that students' gains were predicted by the interaction between teachers' knowledge and the amount of explicit decoding instruction that the students received. Students with more knowledgeable teachers demonstrated stronger gains with increased time in explicit instruction. Conversely, even with an explicit code-focused curriculum, teachers possessing low levels of knowledge produced weaker gains in skill with increased time in explicit phonics instruction. Piasta et al. maintained that their data demonstrated that explicit code-focused curricula cannot replace the expert teaching of highly knowledgeable teachers.

Explaining Gaps in Teachers' Knowledge

As the science of reading development has matured over the past thirty years, a convincing body of evidence has emerged regarding how children learn to read and the best practices in the teaching of reading. This research base is now coupled with a growing literature suggesting that many teachers do not possess the knowledge and skills needed to provide quality early reading instruction. In many ways, the field of teacher education has not kept up with the literature on reading development and instruction. Necessarily, we turn to the role that teacher preparation at the university level may play in the development of disciplinary content knowledge.

For over a decade, researchers and policymakers have been investigating the types of educational opportunities that are afforded to and required of trainee teachers. Lyon (1999b) cautioned that most teachers receive insufficient formal instruction in reading development and reading disabilities or disorders during their undergraduate preparation. Specifically, the average teacher completes just one or two reading courses prior to receiving a degree. Although noted to be insufficient, the statistics presented by Lyon also raise the question of whether the courses offered to or taken by preservice teachers align with our knowledge of instructional best practices. Additionally, it is unclear whether the content of these courses provides the necessary instruction in reading development, the effective components of literacy instruction, and sufficient knowledge of the English language (Walsh, Glaser, & Wilcox, 2006).

In 2001, a study of teacher preparation programs was ordered by the National Center for Education Evaluation and Regional Assistance at the US Department of Education's Institute of Education Sciences (No Child Left Behind Act, 2002). In addition to a direct evaluation of preservice teachers' knowledge about the essential components of early reading instruction, the study included a survey of preservice teachers' perceptions of the extent to which their training focused on the essential components of literacy instruction. Specifically, 2,237 preservice teachers from 99 institutions responded to a survey inquiring about the degree to which their teacher education programs emphasized the five essential components of effective reading instruction of phonemic awareness, phonics, fluency, vocabulary, and comprehension (National Reading Panel, 2000), and provided field-based practica. On average, preservice teachers rated their training programs as placing only little or moderate emphasis on the essential components of reading instruction. Interestingly, preservice teachers were twice as likely to report that there was a stronger focus on the essential components of reading in their field experiences, whereas they reported the opposite emphasis when answering about their coursework.

The National Council on Teacher Quality (NCTQ) (2006) also examined what preservice teachers learn about reading instruction in their undergraduate teacher preparation programs. In this study, analyses were not limited to preservice teacher's perceptions of their coursework. Rather, NCTQ (2006) examined the content of coursework by reviewing the syllabi and texts for the reading-related courses required of students training to be teachers of children ages 6 to 12 at 72 of the 1,271 elementary education programs in the United States. Courses were analyzed to assess the extent to which they provided instruction in the five essential components of effective reading instruction. The results suggested that education schools are not consistently teaching the principles and practices that recent evidence has demonstrated to be effective. Of the seventy-two randomly selected schools, 85%

earned what the study called a failing grade for their instruction of students in the scientific evidence related to reading development and instruction. This finding is noteworthy given that schools could earn a passing grade even if less than 20% of the lectures in a reading-related course were devoted to the science of reading. Additionally, passing grades merely required that course materials reference each of the five essential components of reading instruction and did not require any demonstration that the information was presented accurately or sufficiently. What is perhaps of greater concern is that much of the instruction discussed approaches to literacy instruction that were not scientifically based as if they were as effective as approaches which research does support. The NCQT (2006) report made the additional point that not only are the majority of teacher candidates not receiving sufficient exposure to scientifically based methods of reading instruction but that teacher candidates are often advised to develop their own unique approach to teaching reading.

One hypothesized explanation for why teacher preparation programs are not adequately preparing their students is that the professors who teach the students lack sufficient awareness or knowledge of these critical elements of reading development and instruction. In a 2009 investigation, Joshi et al. examined whether instructors in teacher training programs have sufficient knowledge of reading development, linguistic concepts, and features of the English language and spelling system. Joshi et al. found that, although teachers in these training programs could provide the definition of and manipulate syllables, many had insufficient knowledge of phonemes and concepts related to morphemes. In addition, many instructors of preservice teachers were unable to define phonological awareness (erroneously identifying it as letter–sound correspondence) and failed to recognize phonics as a desirable method for beginning reading instruction. Joshi et al. concluded that professional development opportunities for professors and instructors in teacher preparation programs may be necessary for ensuring proper training of elementary school teachers (see also Binks-Cantrell, Washburn, Joshi, & Hougen, 2012).

Teacher preparation programs are not solely to blame for the fact that teachers lack sufficient knowledge to provide quality literacy instruction. In order to be considered highly qualified under the laws established in the United States, such as the Elementary and Secondary Education Act: No Child Left Behind (2002), teachers must not only hold a degree from a 4-year institution but must also be fully certified or licensed by the state in which they work and demonstrate competency in the core academic subjects in which they teach. Thus, teacher preparation programs do not operate in isolation. Licensing criteria and licensing exams must also do their part to verify that teachers have sufficient knowledge. Although in the United States there are state-by-state differences in licensing exams and the requirements for the number of reading courses that must be taken, research by Stotsky (2009) suggests that in general, licensing exams do not adequately assess the extent to which teachers know the critical information they must possess in order to effectively teach reading (Wayne & Youngs, 2003).

Where Do We Go From Here?

Having established that many teachers lack the skills to provide quality reading instruction, we must now address the gaps in teachers' content knowledge. Teacher preparation programs are the first line of defense. Highly specified training in empirically based understanding of literacy development, effective literacy instruction, and the elements of the English language should be required components of any teacher preparation program. However, adjustments to professional education programs are only effective in ensuring that future teachers have adequate disciplinary knowledge and concept mastery. In isolation, this approach fails to target current educators, some of whom may be in the classroom for thirty or forty more years. Additionally, simply targeting preservice teachers fails to anticipate that our understanding of the teaching of reading will continue to develop. As a result, we must also rely on the education of teachers through ongoing professional development.

Developing Effective Professional Development Interventions

Professional development refers to the acquisition of skills and knowledge, both for personal development and for career advancement. There are numerous models, including training provided by outside experts such as consultation and coaching, communities of practice (educators working collaboratively to achieve better learning outcome of students), lesson study (small group, inquiry-based approach), mentoring, and reflective supervision (Kennedy, 2005). Because research demonstrates that isolated professional development experiences, such as 1-day workshops taught by outside professionals, do not generally result in lasting change in teacher practices and student achievement (for a review, see Joyce & Showers, 1995; Lonigan, Farver, Phillips, & Clancy-Menchetti, 2011), alternative approaches to professional development must be explored. Although it is beyond the scope of this chapter to address this topic in depth, recent syntheses of adult learning and teacher development have identified several key features of effective professional development. Professional development is most successful when it (1) is intensive and ongoing, (2) includes a sequence of active learning experiences such as explaining to a peer what one has learned or practicing teaching activities that build on each other, (3) emphasizes specific skills and goals rather than general ones, (4) provides opportunities for application and practice of newly acquired knowledge and skills, and (5) incorporates feedback to participants about their errors or successes as well as reflection and self-assessment (Darling-Hammond, Wei, Andree, Richardson, & Orphanos, 2009; Fukkink & Lont, 2007).

As the demand in the field for skilled and responsive teachers grows, systematic and sustained models of professional development are needed that include all of the components just described. An examination of current models to identify the active features involved in teachers' development of targeted competencies, leading to improvements in children's school readiness, is also necessary. Because researchers in the field of teacher professional development have called for a shift away from isolated, single, or 1-day workshops and training as the primary mode of delivery, a movement toward more sustained models as exemplified by *relationship-based professional development* models has grown (Bowman, Donovan, & Burns, 2001; Fukkink & Lont, 2007). Relationship-based professional development refers to using relationships to improve the quality of adult learning and can take the form of mentoring, coaching, professional learning communities, and consultation. The goal of relationship-based professional development is to use the skills of experts to provide support and opportunities for learning to those who are less experienced, to promote change, and to support improvement in professional knowledge and pedagogy (National Council on Compensation Insurance [NCCI], 2008). Differences in the type of relationship-based professional development are based on the form of relationship, the purpose of the activity, and how information is shared between the expert and teachers (NCCI, 2008). Research suggests that relationship-based professional development approaches can increase teachers' knowledge and use of effective classroom practices (e.g., Cunningham et al., 2015 Hepburn et al., 2007; Isner et al., 2011; Neuman & Cunningham, 2009; Powell, Diamond, Burchinal, & Koehler, 2010).

In addition to these general features of effective professional development, research has indicated that there are additional characteristics of professional development designed to promote student's emergent literacy and language skills. In a recent review of thirty-seven studies evaluating professional development programs focused on emergent literacy and language among prekindergarten and kindergarten age students, Zaslow, Tout, Halle, Whittaker, and Lavelle (2010) identified several promising practices. Successful programs provided teachers with recommendations for research-based practices and also encouraged teachers to set their own goals and engage in self-reflection. The provision of instructional resources was another key element. Providing teachers with useful, accessible materials such as activity guides, references for further reading, and summaries of key principles may increase the likelihood of sustainability and fidelity to the approach. Another common thread among effective professional development programs is the notion of establishing a cohort of educators, often from the same school, who collaborate toward a shared long-term goal and learn from each other.

The Teacher Study Group Model

The teacher study group model provides a framework for incorporating the features of effective professional development that we have outlined. This approach is in keeping with principles of relationship-based professional development. In this form of professional development, a small group of teachers meet regularly with a highly trained, knowledgeable facilitator. The goal is to work collaboratively toward deepening content knowledge and integrating research-based practices into teaching. Teachers participating in teacher study groups at the elementary school level reported strong, positive attitudes toward the experience of being included in a supportive, collaborative, and reciprocal professional learning environment and appreciated the opportunity to gain knowledge of research-based strategies to promote children's literacy development (Foorman & Moats, 2004; Gersten, Dimino, Jayanthi, Kim, & Santoro, 2010).

In addition to teachers' positive responses to this approach, teacher study groups have been shown to increase teachers' content and pedagogical knowledge, transform pedagogical practices in the classroom, and positively influence child outcomes (Cunningham et al., 2015; Foorman & Moats, 2004; Gersten et al., 2010; Saunders et al., 2001). For example, Cunningham et al. (2015) conducted a study of the teacher study group model of professional development. Their goal was to support teachers' development of the knowledge and practices that promote children's emergent literacy in the preschool classroom. Three sequential cohorts involving a total of nineteen teachers in a high-need community participated in year-long interventions. There was no comparison group. Two-hour meetings were held twice monthly, for a total of sixteen sessions over the academic year. In the biweekly meetings, disciplinary and pedagogical content knowledge in oral language, phonological awareness, and print knowledge was explored. Outcome measures included teachers' knowledge and observational measures of instructional practice and child outcomes for 101 randomly selected preschool children. Consistent with previous research, teachers demonstrated low initial levels of knowledge of phonological awareness, and phonological awareness activities in classrooms were of low quantity and quality. However, pre-, and post test analyses revealed significant changes in teachers' own phonological awareness ability, content knowledge, and pedagogical content knowledge. Increases were also observed in the quantity and quality of phonological awareness activities in the classroom. The preschool children were assessed on a standardized measure of phonological awareness, the Test of Preschool Early Literacy (TOPEL; Lonigan, Wagner, Torgesen, & Raschotte (2007). Before intervention, 64% of the children scored in the below average range or lower. According to the TOPEL test developers, scores in this range "indicate that a child is below the expected developmental trajectory on at least one of the key skills that predict success in learning to read and write" (Lonigan et al., 2007, p. 20). After the intervention, the number of children who scored below average or lower decreased to 36%. Paired-sample *t*-tests comparing pre- and post test standardized TOPEL scores indicated that the children's phonological awareness abilities improved significantly (pretest $M = 86.42$, $SD = 11.58$, posttest $M = 91.99$, $SD = 11.58$; $t(100) = 5.12$, $p < .001$). The mean change represents a movement from the twenty-third percentile to the thirty-fourth percentile. Although the study lacks a control group, the results offer initial support for the use of relationship-based models of professional development to address many of the challenges inherent in providing teachers with the knowledge needed to affect child outcomes in literacy.

To date, the majority of legislation aimed at educational reform has focused on improving students' outcomes by mandating quality instruction. Policymakers must now turn their attention toward building an infrastructure for supporting the development of teachers at both the preservice and in-service level in order to ensure that they are equipped with the skills they need to provide that quality instruction (Aaron, Joshi, & Quatroche, 2008).

Conclusion

In an appendix to George Bernard Shaw's *Man and Superman* (1903), he wrote, "He who can, does; he who cannot, teaches." It is not known in what fit of pique George Bernard Shaw wrote these words, but they have plagued the teaching profession since they were offered. The statement assumes a complete separation between knowing and teaching, as if the two were somehow irrevocably separated.

Almost a century later, Lee Shulman (1986) reframed Shaw's words to provide a less damning aphorism: "Those who can, do; those who understand, teach" (p. 14). Here the division between knowledge and teaching is restored, for as Shulman suggests, in order to teach, one must know. Shulman (1986) proffers a conception of teaching and teachers' knowledge that includes content knowledge, general pedagogical knowledge, curriculum knowledge, pedagogical content knowledge, knowledge of learners and their characteristics, knowledge of educational contexts (i.e., the workings of the classroom, the district, and the character and culture of the community), and knowledge of educational ends, purposes, and values.

If we consider the movement toward quality teaching and quality educational opportunities, we can see that the split between knowing and teaching has finally gone by the wayside: To teach, one must know. As a research community, we are well on our way to understanding the types of educational experiences that students need to become competent readers. Likewise, we have made great strides in our understanding of what teachers must know in order to provide students with the opportunities that lead to positive outcomes. Despite these advances, there is much work to be done. The 2013 National Assessment of Education Progress report on reading (NAEP, 2013) showed that only 34% of US students scored at the proficient level, suggesting a level of reading skill commensurate with grade level expectations. It appears that we have fallen short in seeing to it that our research serves its ultimate purpose of informing practice.

The research–practice divide may be knitted together by employing a common metric or method for educational decision-making. The first step may be to expose teachers to scientific research on reading in their preservice or in-service preparation programs (Spear-Swerling & Sternberg, 2001). By sharing the values and methods of science—such as gathering evidence through systematic observation and testing and considering alternative explanations—powerful tools for settling disputes and for educational decision-making will be available to teachers. As Stanovich (1993) pointed out, reliance on a political/ideological rather than a scientific model for making decisions has hampered the field of reading education. The controversies that have plagued the field would have been better adjudicated by equipping teachers with a scientific model of decision-making.

The second step in bridging theory and practice is ensuring that all teachers in the United States are exposed to the aspects of language discussed in this chapter. The types of professional development we provide our preservice and in-service teachers should be reconsidered. We recognize that the evidence surrounding the most effective content and methods to teach teachers is far less robust than the base of evidence on reading development and instruction. But an emerging body of research is demonstrating that certain aspects of language related to the teaching of reading are elusive concepts to many teachers. We also recognize that most of the work examining teachers' knowledge of their spelling system has been conducted in the United States (cf. Lopes et al., 2014). The relationship between teachers' knowledge of more shallow orthographies (e.g., Spanish, German, Turkish) and children's reading growth may be different than that observed in the less transparent English spelling system.

Third, methods to impart this knowledge to teachers must be explored. Many research-based professional development programs give teachers the what to do, but often neglect to provide the why or how—the critical background knowledge needed for effective teaching of reading. There is a growing recognition that teachers require mentoring in order to grasp these concepts. Just as teacher–child relationships are integral to children's learning (Curby, Rimm-Kaufman, & Ponitz, 2009; Mashburn et al., 2008), relationship-based models of professional development may serve as a means to address many of the challenges discussed. Recognizing the communal nature of adult learning, and fostering supportive, collaborative, and reciprocal professional learning environments may help teachers gain knowledge of research-based strategies to promote children's literacy development.

As Moats argued in her seminal work *Teaching Reading Is Rocket Science* (1999), the field has come to acknowledge the complexities of what teachers of reading should know and do. Just as we would expect a physics teacher to be successful only with deep content knowledge and training from supportive and knowledgeable mentors, we should expect the same level of knowledge is necessary for teachers of reading. Efforts must now be made to ensure that this critical information makes its way into the hands of those responsible for the provision of quality literacy instruction, that is, teachers.

Acknowledgments

This research was supported by the Institute of Education Sciences, U.S. Department of Education, through the Goal 2 Development & Innovation Grant R305A090183 to Anne Cunningham, University of California, Berkeley. The opinions expressed are those of the authors and do not represent views of the US Department of Education. We gratefully acknowledge the careful reviews of this manuscript provided by Dr. Ruth Nathan, Dr. Jamie Zibulsky, and Yi-Jui Chen.

References

Aaron, P. G., Joshi, R. M., & Quatroche, D. (2008). *Becoming a professional reading teacher: What to teach, how to teach, why it matters*. Baltimore, MD: Brookes.

Adams, M. (1990). *Beginning to read: Learning and thinking about print*. Cambridge, MA: MIT Press.

Besner, D., & Smith, M. C. (1992). Basic processes in reading: Is the orthographic depth hypothesis sinking? In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 45–66). Amsterdam, the Netherlands: North-Holland.

Binks-Cantrell, E., Washburn, E. K., Joshi, R. M., & Hougen, M. (2012). Peter effect in the preparation of reading teachers, *Scientific Studies of Reading*, *16*, 526–536.

Bond, G. L., & Dijkstra, R. (1967). The cooperative research program in first-grade reading. *Reading Research Quarterly*, *2*, 5–42.

Bos, C., Mather, N., Dickson, S., Podhajski, B., & Chard, D. (2001). Perceptions and knowledge of preservice and inservice educators about early reading instruction. *Annals of Dyslexia*, *51*, 97–120.

Bowman, B. T., Donovan, M. S., & Burns, M. S. (Eds.). (2001). *Eager to learn: Educating our preschoolers*. Washington, DC: National Academy Press.

Brady, S., Gillis, M., Smith, T., Lavalette, M., Liss-Bronstein, L., Lowe E., . . . Wilder, T. D. (2009). First grade teachers' knowledge of phonological awareness and code concepts: Examining gains from an intensive form of professional development and corresponding teacher attitudes. *Reading and Writing: An Interdisciplinary Journal*, *22*, 425–455.

Brady, S., & Moats, L. (1997). *Informed instruction for reading success: Foundations for teacher preparation*. Baltimore, MD: International Dyslexia Association.

Carlisle, J. F., Correnti, R., Phelps, G., & Zeng, J. (2009). Exploration of the contribution of teachers' knowledge about reading to their students' improvement in reading. *Reading and Writing: An Interdisciplinary Journal*, *22*, 457–486.

Carlisle, J. F., Kelcey, B., Rowan, B., & Phelps, G. (2011). Teachers' knowledge about early reading: Effects on students' gains in reading achievement. *Journal of Research on Educational Effectiveness*, *4*, 289–321.

Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In H. W. Catts & A. G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 25–40). Hillsdale, NJ: Erlbaum.

Chall, J. S. (1967). *Learning to read: The great debate*. New York, NY: McGraw-Hill.

Chall, J. S. (1992). The new reading debates: Evidence from science, art, and ideology. *The Teachers College Record*, *94*, 315–328.

Cheesman, E. A., McGuire, J. M., Shankweiler, D., & Coyne, M. (2009). First-year teacher knowledge of phonemic awareness and its Instruction. *Teacher Education and Special Education*, *32*, 270–289.

Connor, C. M., Son, S., Hindman, A. H., & Morrison, F. J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience: Complex effects on first graders' vocabulary and early reading outcomes. *Journal of School Psychology*, *43*, 343–375.

Cummings, D. W. (1988). *American English spelling*. Baltimore, MD: John Hopkins University Press.

Cunningham, A. E., Etter, K., Platas, L., Wheeler, S., & Campbell, K. (2015). Professional development in emergent literacy: An examination of teacher study groups. .

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Stanovich, P. J. (2004). Disciplinary knowledge of K-3 teachers and their knowledge calibration in the domain of early literacy. *Annals of Dyslexia*, *54*, 139–167.

Cunningham, A. E., Perry, K. E., Stanovich, K. E., Stanovich, P. J., & Chappell, M. (2001, June). *Is teachers' knowledge of important declarative knowledge of reading well calibrated?* Poster presented at the 8th Annual meeting of the Society for the Scientific Study of Reading, Denver, CO.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*, 934–945.

Cunningham, A. E., & Zibulsky, J. (2009). Perspectives on teachers' disciplinary knowledge of reading processes, development, and pedagogy. *Reading and Writing: An Interdisciplinary Journal*, *22*, 375–378.

Cunningham, A. E., Zibulsky, J., & Callahan, M. (2009). Starting small: Building preschool teacher knowledge that supports early literacy development. *Reading and Writing: An Interdisciplinary Journal*, *22*, 487–510.

Cunningham, A. E., Zibulsky, J., Stanovich, K. E., & Stanovich, P. J. (2009). How teachers would spend their time teaching language arts: The mismatch between self-reported and best practices. *Journal of Learning Disabilities*, *42*, 418–430.

Curby, T. W., Rimm-Kaufman, S. E., & Ponitz, C. C. (2009). Teacher–child interactions and children's achievement trajectories across kindergarten and first grade. *Journal of Educational Psychology*, *101*, 912–925.

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession*. Washington, DC: National Staff Development Council.

Foorman, B. R., & Moats, L. C. (2004). Conditions for sustaining research-based practices in early reading instruction. *Remedial and Special Education*, *25*, 51–60.

Fukkink, R. G., & Lont, A. (2007). Does training matter? A meta-analysis and review of caregiver training studies. *Early Childhood Research Quarterly*, *22*, 294–311.

Gersten, R., Dimino, J., Jayanthi, M., Kim, J. S., & Santoro, L. E. (2010). Teacher study group: Impact of the professional development model on reading instruction and student outcomes in first grade classrooms. *American Educational Research Journal*, *47*, 694–739.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, *5*, 3–8.

Goodman, K. (1986). *What's whole in whole language?* Portsmouth, NH: Heinemann.

Hepburn, K. S., Kauffman, R. K., Perry, D. F., Allen, M. D., Brennan, E. M., & Green, B. L. (2007). *Early childhood mental health consultation: An evaluation tool kit*. Washington, DC: Georgetown University Press.

Isner, T., Tout, K., Zaslow, M., Soli, M., Quinn, K., Rothenberg, L., & Burkhauser, M. (2011). *Coaching in early care and education programs and quality rating and improvement systems (QRIS): Identifying promising features*. Washington, DC: Children's Services Council of Palm Beach County.

Joshi, R. M., Binks, E., Graham, L., Ocker-Dean, E., Smith, D. L., & Boulware-Gooden, R. (2009). Do textbooks used in university reading education courses conform to the instructional recommendations of the National Reading Panel? *Journal of Learning Disabilities*, *42*, 458–463.

Joshi, R. M., Binks, E., Hougen, M., Dahlgren, M., Ocker-Dean, E., & Smith, D. (2009). Why elementary teachers might be inadequately prepared to teach reading. *Journal of Learning Disabilities*, *42*, 392–402.

Joyce, B., & Showers, B. (1995). *Student achievement through staff development* (2nd ed.). New York, NY: Longman.

Kennedy, A. (2005). Models of continuing professional development: A framework for analysis. *Journal of In-Service Education*, *31*, 235–250.

Lonigan, C. J., Farver, J. M., Phillips, B. M., & Clancy-Menchetti, J. (2011). Promoting the development of preschool children's emergent literacy skills: A randomized evaluation of a literacy-focused curriculum and two professional development models. *Reading and Writing: An Interdisciplinary Journal*, *24*, 305–337.

Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2007). *The test of preschool early literacy (TOPEL)*. Austin, TX: PRO-ED.

Lopes, J., Spear-Swerling, L., Oliveira, C., Velasquez, M. G., & Zibulsky, J. (2014). Actual disciplinary knowledge, perceived disciplinary knowledge, teaching experience and teacher's training for reading instruction: A study with primary Portuguese and American teachers. *Revista de Psicodidactica*, *19*, 45–65.

Lyon, G. R. (1999a, July, 27). Hearing on Title I (Education of the Disadvantaged) of the Elementary and Secondary Education Act and the Workforce United States House of Representative. Electronic mailing list message. http://archives.republicans.edlabor.house.gov/archive/hearings/106th/fc/esea72799/lyon.htm

Lyon, G. R. (1999b). In celebration of science in the study of reading development, reading difficulties, and reading instruction: The NICHD perspective. *Issues in Education: Contributions from Educational Psychology*, *5*, 85–115.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., . . . Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79*, 732–749.

Mather, N., Bos, C., & Babur, N. (2001). Perceptions and knowledge of preservice and inservice teachers about early literacy instruction. *Journal of Learning Disabilities*, *34*, 472–482.

McCutchen, D., Green, L., Abbott, R. D., & Sanders, E. A. (2009). Further evidence for teacher knowledge: Supporting struggling readers in grades three through five. *Reading and Writing: An Interdisciplinary Journal*, *22*, 401–423.

McCutchen, D., Harry, D. R., Cunningham, A. E., Cox. S., Sidman, S., & Covill, A. E. (2002). Reading teachers' knowledge of children's literature and English phonology. *Annals of Dyslexia*, *52*, 207–228.

Moats, L. C. (1994). The missing foundation in teacher education: Knowledge of the structure of spoken and written language. *Annals of Dyslexia*, *44*, 81–101.

Moats, L. C. (1999). *Teaching reading is rocket science: What expert teachers of reading should know and be able to do.* Washington, DC: American Federation of Teachers.

National Assessment of Educational Progress. (1995). *NAEP Facts: Listening to children read aloud: Oral fluency.* Washington, DC: National Center for Education Statistics.

National Center for Education Statistics. (2013). *National Assessment of Educational Progress (NAEP), various years, 1990–2013 Mathematics and Reading Assessments*. http://nationsreportcard.gov/reading_math_2013/#/student-groups.

National Council on Compensation Insurance (NCCI). (2008). *State requirements for minimum preservice qualifications for child care center teachers and master teachers in 2007*. National Child Care Information and Technical Assistance Center. Washington, DC.

National Council on Teacher Quality (NCTQ). (2006). *What education schools aren't teaching about reading and what elementary teachers aren't learning*. Washington, DC: National Council on Teacher Quality.

National Early Literacy Panel (NELP). (2008). *Developing early literacy: Report of the national early literacy panel*. Washington, DC: National Institute for Literacy.

National Reading Panel (NRP). (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institutes of Health.

Neuman, S. B., & Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal*, *46*, 532–566.

No Child Left Behind. (2002). *No child left behind act*. Washington, DC: US Department of Education.

Piasta, S. B., Connor, C. M., Fishman, B. J., & Morrison, F. J. (2009). Teachers' knowledge of literacy concepts, classroom practices, and student reading growth. *Scientific Studies of Reading*, *13*, 224–248.

Powell, D. R., Diamond, K. E., Burchinal, M. R., & Koehler, M. J. (2010). Effects of an early literacy professional development intervention on head start teachers and children. *Journal of Educational Psychology*, *102*, 299–312.

Saunders, B., O'Brien, G., Hasenstab, K., Marcelletti, D., Saldivar, T., & Goldenberg, C. (2001). Getting the most out of school-based professional development. In P. Schmidt & P. Mosenthal (Eds.), *Reconceptualizing literacy in the new age of pluralism and multiculturalism* (pp. 289–320). Greenwich, CT: Information Age.

Shaw, G. B. (1903). *Man and superman: Maxims for revolutionists*. London, England: Penguin

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*, 4–14.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*, 1–21.

Smith, F. (1971). *Understanding reading*. New York, NY: Holt, Rinehart, & Winston.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Spear-Swerling, L., & Brucker, P. (2004). Preparing novice teachers to develop basic reading and spelling skills in children. *Annals of Dyslexia*, *54*, 332–364.

Spear-Swerling, L., & Sternberg, R. J. (2001). What science offers teachers of reading. *Learning Disabilities Research and Practice*, *16*, 51–57.

Spear-Swerling, L., Brucker, P. O., & Alfano, A. M. (2005). Teachers' literacy-related knowledge and self-perceptions in relation to preparation and experience. *Annals of Dyslexia*, *55*, 266–296.

Stanovich, K. E. (1993). Romance and reality. *Reading Teacher*, *47*, 280–280.

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York, NY: Guilford.

Stanovich, K. E., & Stanovich, P. J. (1995). How research might inform the debate about early reading acquisition. *Journal of Research in Reading*, *18*, 87–105.

Stotsky, S. (2009). Licensure tests for special education teachers: How well they assess knowledge of reading instruction and mathematics. *Journal of Learning Disabilities*, *42*, 464–474.

Treiman, R., Kessler, B., & Evans, R. (2007). Anticipatory conditioning of spelling-to-sound translation. *Journal of Memory and Language*, *56*, 229–245.

Venezky, R. L. (1999). *The American way of spelling: The structure and origins of American English orthography*. New York, NY: Guilford.

Walsh, K., Glaser, D., & Wilcox, D. D. (2006). *What education schools aren't teaching about reading and what elementary teachers aren't learning*. Washington, DC: National Council on Teacher Quality.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, *73*, 89–122.

Wong-Fillmore, L., & Snow, C. E. (2003). What teachers need to know about language. In C. T. Adger, C. E. Snow, & D. Christian (Eds.), *What teachers need to know about language* (pp. 7–54). McHenry, IL: Delta System.

Zaslow, M., Tout, K., Halle, T., Whittaker, J. V., & Lavelle, B. (2010). *Toward the identification of features of effective professional development for early childhood educators: Literature review*. Washington, DC: Office of Planning, Evaluation and Policy Development, US Department of Education.

CHAPTER

30 Adolescent Literacy: Development and Instruction

Susan R. Goldman *and* Catherine E. Snow

Abstract

The demands of literacy tasks change appreciably after students have mastered the basics of reading words accurately and with reasonable automaticity. At about age 10, reading becomes a tool for acquiring information, understanding a variety of points of view, critiquing positions, and reasoning. The results of international and US assessments show that many students who succeed at early reading tasks struggle with these new developmental challenges, focusing attention on the instructional needs of adolescent readers. Commonly used approaches to comprehension instruction in the postprimary grades, such as teaching reading comprehension strategies, do not adequately respond to the multiple challenges adolescent readers face. One such challenge is the need to acquire discipline-specific ways of reading, writing, and thinking, often from teachers who are themselves insufficiently aware of how reading literature differs from reading science or history. The chapter argues that appropriate attention in instruction to discipline-specific literacy practices, to maintaining an authentic purpose for assigned literacy tasks, and to the role of focused discussion as a central element in teaching comprehension would improve reading outcomes and would revolutionize current theories about the nature of reading comprehension.

Key Words: adolescent literacy, comprehension, disciplinary literacy, comprehension interventions, history, science, literature, instructional conversations, development, instruction

With new technologies and the World Wide Web, twenty-first-century citizens can now access unprecedented amounts of information on topics ranging from climate change to immigration policy to medical care and financial options. The price of unfettered availability is that information is no longer filtered by teachers, librarians, and traditional publishers. As a result, the burden of determining reliability and relevance falls on the reader, a process made more complex when information sources are inconsistent with, or even contradictory to, one another (Rouet & Britt, 2011; Stadtler & Bromme, 2007). Furthermore, the analysis, synthesis, and evaluation of information is increasingly expected in the workplace, even in jobs that do not require a college education (e.g., driving a truck, repairing ventilation systems). The definition of reading comprehension must include these new, twenty-first-century literacy skills, and literacy instruction for 10- to 18-year-olds should anticipate these realities.

Thus we adopt a conception of reading that aligns with national and international performance standards (National Assessment Governing Board, 2008; Organization for Economic Cooperation and Development [OECD], 2006). These standards define proficiency as reading to acquire content knowledge useful in addressing open-ended questions, solving problems, or making decisions. Conventional notions of comprehension involve locating and identifying facts, combining explicitly stated ideas, and making simple inferences from single text sources. In contrast, new standards specify that students should analyze, interpret, integrate, critique, and evaluate information within single and across multiple sources of information (National Assessment of Educational Progress [NAEP], 2009). Critical literacy processes require attention to text and context; to the use of rhetorical and symbolic devices to convey meaning; and to disciplinary conventions for argumentation, critique, and evaluation. Assessment data indicate that across the globe only about 10% of today's adolescents are mastering such advanced reading comprehension and critical literacy skills (Carnegie Council on Advancing Adolescent Literacy [CCAAL], 2010; NAEP, 2009; OECD, 2013). Thus the need arises for a serious consideration of how to improve adolescent literacy outcomes.

Reading Instruction

Traditionally, most attention in reading research and instruction has been focused on early reading. In fact, only recently has the field called adolescent literacy even been recognized (CCAAL, 2010; Goldman, 2012; Lee & Sprately, 2010; Snow & Moje, 2010). But it is now clear that, although proficiency in word reading and basic fluency are crucial to long-term outcomes (see Connor & Al Otaiba, this volume), these skills are insufficient to ensure success in the more challenging comprehension tasks facing adolescents (Wanzek et al., 2012). Furthermore, where once it was thought that the major change between early and later reading was the complexity of the texts, it is now clear that the tasks change even more radically than the texts. It is widely recognized that reading instruction needs to continue throughout schooling and that the focus of such instruction should be decreasingly the mechanics of reading and increasingly the language, discourse, and argumentation structures of the disciplines in which the reading is located.

The fact that such a high proportion of reading research has emanated from the English-speaking world has inflated attention to early reading instruction, because of the challenges posed by teaching students to read in the deep orthography of English. Students learning to read in more shallow orthographies, those in which phoneme–grapheme mapping is one-to-one rather than many-to-many, master the basic mechanics of word recognition and automaticity within the first year of schooling (Aro & Wimmer, 2003). Such students reveal much earlier the comprehension deficits associated with limited vocabulary, lack of relevant background knowledge, and unfamiliarity with how information is organized in literature versus science versus history. English-speaking readers start to reveal the consequences of such deficits later: US assessments suggest that the percentage of poor readers increases in the 10- to 12-year age range (National Center for Education Statistics, 2013). Poor reading skills will no doubt become even more evident with the introduction of new assessments aligned to the new, higher college and career-ready standards adopted by many of the states in the United States (Council of Chief State School Officers [CCSSO], 2010). Known as the *Common Core State Standards,* they set grade-level expectations for reading, writing, speaking, and listening. They specify that adequate literacy performance requires analysis, synthesis, and critiquing of text, and call for a much higher proportion of nonliterary and complex texts in instruction and in assessment. These specifications are not limited to performance in literature classes but extend to companion standards for reading in history and in science and technical subjects. For perhaps the first time in the history of educational policy and standard setting, a nationwide educational policy calls specifically for attention to the unique reading and writing practices of at least two disciplinary contexts other than literature. These will have the most profound effects on adolescents, since their schooling is typically discipline-based.

Challenges of Postprimary Reading

What are the issues unique to adolescent literacy that require it be dealt with as a topic separate from initial literacy development or instruction? There are several, at different levels of analysis—the learner, the task, the texts, the pedagogy, and the context.

THE LEARNERS

Adolescent readers may need instructional attention for any of a number of reasons. Some still struggle to read words accurately, fluently, and with automaticity. This becomes an insuperable obstacle as they are expected to read longer, more complex texts containing more unfamiliar words after grade three, around age 8. But many adolescents with good basic reading skills need targeted instruction as well, because comprehension of the texts they are asked to read requires a larger vocabulary, more background knowledge, more stamina, and greater motivation than many possess for school tasks. At the same time as school tasks are requiring more knowledge, skills, and persistence, concerns about identity and increased interest in peer relationships compete with academic and cognitive demands, leading to an almost universal decline in intrinsic motivation for reading (Gottfried, 1985).

THE TASKS

Up to about the end of third grade, children are considered good readers if they can read grade-level texts aloud without many errors and answer low-inference comprehension questions such as "What did the wolf want to do to Red Riding Hood?" In the primary grades, children read primarily to learn to read, from texts that use language structures within their oral language repertoires. In the higher grades, students must demonstrate that they can learn new information from reading texts on their own, that they can sift through large amounts of text to find relevant information, and that they can analyze literary texts (Chall, 1983). To reach higher levels of proficiency on tests like the Program for International Student Assessment (PISA; OECD, 2006), students must be able to analyze texts to find evidence for claims, compare texts to find disparities in information and in interpretation, learn word meanings from text, synthesize information from multiple texts, and so on (Goldman et al., 2011; Lee & Spratley, 2010; Shanahan & Shanahan, 2008). Instruction needs to improve and expand to help students succeed at these tasks.

THE TEXTS

Texts presented to older students are, of course, more complex in language structures and vocabulary than those in primary classrooms. In addition, a much higher proportion of them are expository texts, which organize the presentation of information in ways quite different from narrative or literary texts. The genre-specific organization of information is referred to as *discourse structure*. Furthermore, vocabulary and discourse structures are increasingly differentiated by content area so that reading history texts and science texts requires overlapping but not identical skills (Lee & Spratley, 2010; Moje, 2008).

THE PEDAGOGY

The pedagogical focus shifts in the postprimary grades from skills to content. Thus literature, history, science, and mathematics teachers in middle and high schools do not think of themselves as responsible for teaching reading, nor are they for the most part prepared to provide instruction in how to read the texts they assign (Heller & Greenleaf, 2007). Two possible consequences ensue. Either students who struggle with the texts have no chance to learn the content, or content is provided through lectures that supplant the texts and thus unwittingly undermine students' abilities to progress as content area readers.

THE CONTEXT

Because the unique needs of adolescent readers have only recently been widely recognized, most schools and teachers lack the resources to meet those needs (CCAAL, 2010). Teacher preparation programs and professional development efforts pay little attention to the skills required to provide discipline-specific literacy teaching or to incorporate open-ended questions and discussion into classrooms, despite strong evidence that these features support comprehension. Curricula are not in general designed to support such efforts, nor are schools always organized to promote them. In addition, schools serving students at risk of poor reading outcomes (those from low-income homes, from families that do not speak the school language, from primary programs that have not provided adequate instruction) are typically even less likely to have well-prepared teachers and engaging or challenging curriculum and pedagogy.

In light of these considerations, we will review what is currently known about literacy development and literacy instruction in the postprimary grades. We focus on a few themes: the inadequacy of just teaching comprehension strategies, which is currently the dominant instructional approach; the need to attend to the ways in which texts and tasks differ across content areas, which we refer to as *disciplinary literacies*; the interdependencies among writing, discussion, and reading in the adolescent years; and the value of attending to purpose and engagement in constructing instructional activities. Note that our use of "text" in this chapter is intended to encompass static and dynamic visual representations of information and oral modes of communication, as well as traditional, printed verbal text.

Outside the scope of our discussion is a burgeoning focus on adolescents' engagement with reading and writing outside of school contexts. Numerous reports chart adolescents' increasing use of social media and participation in affinity groups, online games, and other Internet-enabled hangouts (e.g., Lenhart, 2013; Madden, Lenhart, Duggan, Cortesi, & Gasser, 2013). To date, few positive relationships have been found between participation in these types of out-of-school activities and performance on school-based indices of achievement (e.g., grades, achievement tests) (Moje, Overby, Tysvaer, & Morris, 2008; Purcell, Buchanan, & Friedrich 2013). One educational challenge is to build on these interests and the knowledge they generate to leverage adolescents' participation in school-based activities.

Comprehension Strategy Instruction: A Little Patch on a Big Problem

Chall (1983) identified a major shift in the focus of reading instruction after the first several years of formal schooling from learning to read to reading to learn. By far the most common approach to enacting this shift has been to teach comprehension through the explicit teaching of comprehension strategies (see Oakhill, Berenhaus, & Cain, this volume). Experimental studies have indeed shown positive effects of strategy instruction (National Reading Panel, 2000). However, there are many challenges in teaching comprehension that a simple strategy-based approach cannot resolve.

One challenge relates to the knowledge that readers bring to texts. Comprehension instruction in grades four through twelve is typically applied to fictional narratives. Even young readers have a rich supply of knowledge about the kinds of events and motivations that are central to such stories—what we might think of as background knowledge about human psychology. They can benefit from guiding comprehension questions such as Who are the characters? What is the setting? What happened first? What happened next? Why was she sad/mad/happy? (Duke & Martin, 2008).

Questions like these do not apply to informational texts in science or social studies (nor, in fact, to the full array of literary genres). Generic comprehension strategies (find the main idea, identify the topic sentence, summarize, learn the bold-faced words) are often introduced for informational texts (Alvermann & Wilson, 2011; McKeown, Beck, & Blake, 2009; Palincsar & Brown, 1984; Pressley, 2002). Indeed, they can be helpful in reading textbooks and textbook-like materials in which the conventions match these generic strategies. For example, key vocabulary items are bolded, section headers mark new topics, and the first sentence under the header is often a good summary of the section. Such strategies are very difficult to apply, however, to the full array of texts we hope students are reading—newspaper articles, historical documents, journalistic reports of research, editorials, political speeches. These texts vary in the way information is organized and the conventions used to signal more versus less important information, most of which school-age readers have not been taught (Goldman & Bisanz, 2002; Goldman & Rakestraw, 2000). Lacking these organizational cues to importance, students need the very information they are trying to learn if they are to, for example, evaluate whether or not a summary captures the important ideas. Importantly, it is possible to teach adolescents to use the structure of text to bootstrap comprehension of new information (e.g., Meyer & Wijekumar, 2007).

The set of comprehension strategies that is widely taught was selected because of the empirical evidence that good readers and better learners use them (Goldman & Saul, 1990; Pressley, 2002; RAND, 2002). The assumption was that if poor readers could be taught to do what good readers do, their comprehension and learning would improve. But good readers use strategies like monitoring, self-questioning, and rereading selectively—and they can be selective because they understand some of what they are reading. Good readers are also strategic in their choice of what to read deeply, what to skim, and what to skip. Strategic deployment of reading strategies depends on readers being metacognitively aware of whether, and how, particular efforts—strategic or otherwise—are leading to success in the task for which the reading is being done (cf. Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012; Goldman, Lawless, et al., 2012).

Guided by the findings from good readers, comprehension strategy training over the last forty to fifty years progressed from single to multiple strategy interventions (cf. Pressley, 2002). As well, there has been a shift from autonomous student use of strategies to applying them in pairs or small-group settings, often with strategic roles (questioning, summarizing, monitoring) distributed over different group members. In small groups, students get feedback on their own thinking as well as exposure to alternative approaches to building understanding and to other interpretations of the text. Multiple-strategies interventions include attention to metacognitive strategies for monitoring understanding and for selecting what strategy to use when. Although multiple-strategies interventions have targeted informational text, there has been little attention to the differentiated conceptual skills and knowledge required in specific content areas (e.g., science, history, or literature). Few interventions have targeted multiple content areas simultaneously so as to provide students with contrastive information about discipline-specific reading.

The limited success of generic reading strategies as a basis for comprehension instruction is also related to changes in the tasks posed and the behavior that is taken as indicative of comprehension. Whereas 10-year-olds might be asked only to summarize or to define a novel word after reading an expository text, by early adolescence students should be asked to make inferences, identify the author's point of view, evaluate the credibility of claims and conclusions, and integrate information derived from several sources (Alvermann & Wilson, 2011; Lee & Spratley, 2010; Shanahan & Shanahan, 2008; Snow & Biancarosa, 2004). These skills are particularly important for reading the variety of text genres (articles, blogs, comments, posts, tweets) of varying credibility encountered on the World Wide Web (Stadtler & Bromme, 2007). These tasks require that adolescents reason with and about the information that they read, bringing to bear applicable inferential processes and requisite content knowledge that the texts do not provide. In so doing, students need to coordinate information in one text with information in other texts and with already known information. They also need knowledge of appropriate criteria for evaluating the information and the reasoning they do with it.

Thus both the texts and reading-to-learn tasks adolescents encounter create high prior knowledge demands. It is therefore not surprising that researchers repeatedly find an effect of prior knowledge on new learning—those who start out knowing more about a topic typically perform better on memory and learning tasks than those who know less (Alexander & Jetton, 2002; Kintsch, 1994; McNamara & Kintsch, 1996). Evidence that a major obstacle to successful comprehension is lack of relevant background knowledge has led to approaches that downplay the role of generic reading skills (self-monitoring, drawing inferences) or improving readers' thinking processes. Rather, these approaches emphasize the match between reader knowledge and text demands—building on the familiar observation that even good readers flounder when asked to comprehend texts about unfamiliar topics (reports of cricket matches for Americans or of baseball games for South Africans). Although it would be desirable to have direct comparisons of the impact of exposure to background knowledge versus strategy instruction, we are not aware of any studies of this sort. Rather, prior knowledge is frequently treated as an individual differences variable that is used either to define contrastive groups, to predict performance, or as a covariate to control for differences among participants in prior knowledge.

The question remains, however: How do readers make sense of texts for which they lack requisite prior knowledge? Some interventions address this question by framing reading as an inquiry process and externalizing the processes of making sense of the text; classroom talk focuses on how as well as what understanding is occurring (e.g., Schoenbach, Greenleaf, & Murphy, 2012). In interactions referred to as *metacognitive conversations* (Schoenbach et al., 2012), readers discuss with each other what they do and do not understand, how they figured out the meaning of an unfamiliar word, what parts of the text confused them and why, what questions they have, and so forth. Externalizing the processes of sense-making exposes not only a range of strategies for building understanding but also what was difficult to understand and strategies for dealing with these difficulties. Externalizing reading processes in this way encourages deep engagement with the text even when comprehension difficulties are encountered. These types of interventions encourage readers to engage with rather than disengage from a complex text by focusing on the questions the text raises.

In a review of studies conducted with fifth graders, McKeown et al. (2009) found that text-focused questions generated better comprehension and better recall than strategies instruction. The results of this rigorous study confirmed a wide array of findings showing the importance to successful comprehension of mastering procedures for figuring out precisely what the text says and what to do when you cannot.

Science, History, Math, and Literature: Learning Content and Literacy Together

With few exceptions (cf. Greenleaf et al., 2011), most comprehension instruction approaches have been tested on generic, narrative texts. However, studies of experts in different disciplines make clear they read differently and that the comprehension demands of texts in different disciplines are not the same. Literary experts reading poetry and prose relate what they are reading to other works by the same author and from the same period. They are sensitive to multiple interpretations and explore insights into human experience afforded by the literary work (Graves & Frederiksen, 1996; Langer, 2010; Lee, 2007). In both history and science, experts routinely engage in selection, analysis, and synthesis within and across multiple sources of evidence (Chinn & Malhotra, 2002; Shanahan & Shanahan, 2008; Wineburg, 1991). However, historians and scientists approach texts in their fields differently. For example, chemists use multiple representations to understand what they are reading. In other words, a hydrogen molecule can be represented symbolically (H_2O) or visually as a hydrogen atom connected to two oxygen atoms. When reading, chemists spend much time relating elements of one representational form to another. On the other hand, historians first look at and spend time considering when, why, and by whom a text was created. For the most part, novices, including adolescents, do not engage in the disciplinary processes exhibited by experts (Britt & Aglinskas, 2002; Goldman, Braasch, et al., 2012; Greene, 1994; Rouet, 2006; Rouet, Britt, Mason, & Perfetti, 1996; Seixas, 1994; Smith, 1991; Wineburg, 1991). Interestingly, experts reading outside their field of expertise do not display the same strategies they use when they read within their field of expertise (Bazerman, 1985, 1998). This reinforces the important role of topic knowledge and discipline-specific tasks in guiding reading behavior. It also makes clear that adolescents will need instruction in how to read for each of the content areas in which they are expected to use literacy skills as a route to knowledge.

That disciplinary literacies have been largely absent from middle and high school curricula is not surprising for several reasons. First, in most schools, content learning has been dominated by transmission of what is known rather than how it came to be known. That is, students are typically not engaged in the disciplinary inquiry practices that generate knowledge. Compendia-like textbooks present what is known; students are expected to learn but not necessarily to understand the facts. Criticisms of the transmission model of learning abound. As discussed earlier, contemporary approaches point to the importance of active involvement in learning that engages students in developmentally appropriate forms of disciplinary practices, including disciplinary reading practices (e.g., Bransford, Brown, & Cocking, 2000; Donovan & Bransford, 2005; Duschl, Schweingruber, & Shouse, 2007; Langer, 2010; Lee, 2007; Moje, 2008; Wineburg, 2001). Second, most disciplinary experts, including teachers of adolescents, are typically unaware of how they themselves read in their discipline. Also, they are typically unaware that the way in which they read and interpret texts in their own discipline is not the same as the way in which a teacher from a different discipline would read and interpret texts in that discipline (e.g., Grossman, Wineburg, & Woolworth, 2001). In other words, teachers' knowledge of their disciplinary literacy practices is tacit, so they are unaware of the need to make these practices explicit to students. Finally, we lack information about the time courses and trajectories for the development of disciplinary reading. For example, we know little about how and when to introduce and then deepen disciplinary inquiry practices for building content knowledge from text, either within a school year or across years. We also know little about how to support students' appreciation of the differences across disciplines in the nature of valid arguments and how these are manifest in written (and spoken) discourse.

Authentic reading tasks in the various disciplines typically require reading multiple sources to investigate a question or solve a problem. In literature the inquiry might be about comparative styles and themes, in science about verifying claims and constructing explanations of phenomena in the physical world, in history about contrasting perspectives and explanatory accounts of events. In school, adolescents are rarely asked to read multiple sources and synthesize what they read to solve a problem or answer a puzzling question. In this era of unlimited digital access to vast quantities of information, the capacity to sift, sort, and synthesize—achieving what Bråten and Strømsø (2010) refer to as multiple documents literacy—is crucially important (Bromme & Goldman, 2014; Goldman & Scardamalia, 2013; Rouet, 2006). With the recognition of the importance of reading multiple sources using reading practices specific to particular disciplines, researchers have begun to research and develop instructional approaches through laboratory and classroom-based studies. We describe examples of these in history, science, and literature.

Teaching History Content and Literacy Practices of History

Historians construct accounts of historical events by reading traces of the past found in documents and other artifacts produced at the time of the event as well as other historical accounts written subsequent to the event. They engage in critical analyses, syntheses, and evaluations across these sources using three historical inquiry practices: sourcing, contextualization, and corroboration (Wineburg, 1991). Sourcing asks who produced the source (document, artifact), for what purpose, and when, in order to identify the perspective or point of view reflected in the information. Contextualization considers the larger set of circumstances and events that were occurring at the time a particular source was created. Corroboration examines whether and where multiple sources about the same time period or event agree and disagree. Reading like a historian involves employing these practices. For example, documents written by Hitler about the causes of World War II would reflect his particular perspective and the context in which he lived. Churchill's writings about the causes of World War II would likewise reflect his context and disagree in some fundamental ways with Hitler's.

Several classroom-based interventions have attempted to teach adolescents to use the three historical inquiry practices when reading historical documents (e.g., De La Paz, 2005; Nokes, Dole, & Hacker, 2007; Reisman, 2012). Each prompted sourcing, contextualizing, and corroborating through a series of questions that students were to answer based on a set of historical documents. One of these studies, conducted with 13- and 14-year-olds (De la Paz, 2005), also taught students to evaluate what they had read in light of their answers to the sourcing and corroboration questions and how to write a historical argument essay using the information from the documents. Students in the treatment condition produced longer and more compelling essays than control group students. Interviews indicated that the treatment students showed greater understanding than control students of how historians reason and why there might be different opinions about historical events.

Reisman (2012) took a similar approach to that of De La Paz (2005) in a study conducted with 16- and 17-year-olds and found similar positive effects. In addition to questions that promoted the three historical inquiry practices, Reisman launched each topic in the US history survey course with an engaging question that required reconciling multiple accounts. Further, to assist students in understanding the historical argument, Reisman provided students with questions intended to focus them on the claims made and evidence the author was using, as well as the words and phrases that led students to think the author was using the information appropriately. A 6-month implementation of this program in five high schools produced positive effects on knowledge of facts taught, using historical thinking skills, sophistication in applying those strategies to current events, and general reading comprehension.

De La Paz (2005) and Reisman (2012) have demonstrated that developmentally appropriate forms of reading like a historian have a positive effect on adolescents' literacy achievement and content learning. In both cases, these approaches were implemented with substantial support from the researchers. Project READI (Reading, Evidence, and Argument in Disciplinary Instruction; Goldman et al., 2009) is exploring the kinds of experiences and supports that teachers need to adopt this approach. One key issue is building progressions in reading like a historian within and across grade levels, starting at ages 10 through 12 and extending through late adolescence. Another key issue is specifying the performances that would reflect mastery of, for example, sourcing in a 12-year-old as compared with an 18-year-old. Particularly important for this work are teachers' analyses of the potential challenges posed by specific texts and tasks for their students (e.g., unfamiliar concepts, archaic language, contextual information) and the implications of these for the types of instructional supports they need to provide.

Teaching Science and Literacy Practices of Science Together

Recent efforts in the United States to foster greater achievement and better appreciation for science emphasize the practices of science that inform scientists' formulation of timely and interesting research questions, underlie data representations, and communicate science to other scientists and the general public (Achieve, 2013; Bromme & Goldman, 2014). Several instructional interventions exemplify ways in which the literacy practices of science can be integrated with hands-on experiences of doing observational and experimental science. For example, Magnusson and Palincsar (1995) fostered literacy and inquiry-based (hands-on) science to create an instructional approach they called *Scientist's Notebook*. The main feature of the approach was the science notebook of Lesley, a fictitious physical scientist. In this notebook, Lesley (really Magnusson) modeled experimentation. She recorded her questions about the physical world (e.g., about variables affecting motion), data collection plans, different displays of the data she had collected (graphs, tables), notes about patterns she noticed, and conclusions and revisions to her conclusions based on challenges from colleagues. Fourth- and fifth-grade students (9–11 years old) used these notebooks along with additional texts to conduct their own inquiries arising from their analyses and critiques of Lesley's information. Extensive documentation of instructional conversations around the notebook and related texts as well as students' work indicated increases in critical science literacy practices, including coordination of information across different representations and multiple texts (Hapgood, Magnusson, & Palincsar, 2004).

Romance and Vitale (2001) also reported increases in science understanding and reading achievement for 9- to 11-year-olds who participated in their In-Depth Expanded Applications of Science (IDEAS) program. Instruction included hands-on inquiry experiences that tested predictions and observations students generated in response to "What would happen if…" questions. Students learned to carefully read science texts to inform their written explanations of findings from the hands-on inquiry. Students engaged in additional reading and writing activities to compare their current understanding with what they had originally understood and applied their knowledge to new contexts. When compared with students in regular programs that separated reading and science, IDEAS students performed significantly better on standardized measures of science achievement, reading achievement, and attitudes toward learning, with a one-year grade equivalent difference in science and one-third of a year in reading. Romance and Vitale also showed that the gains persisted for three grade levels beyond students' last year of participation in IDEAS.

Cervetti, Barber, Pearson, and Goldschmidt (2012) expanded on the design features of both Scientist's Notebook and IDEAS in *Seeds of Science/Roots of Reading,* a curriculum for 9- to 12-year-olds. Students do their own investigations (observational as well as experimental), but also read to compare their own findings with those of other investigators who collected data under different conditions. They also read texts that explain the mechanisms that underlie causal relationships that are depicted in diagrams. For example, causal relationships are often depicted by directional arrows from cause to effect (e.g., the water cycle, photosynthesis), However, the visual does not explain why, for example in the case of the water cycle, rain forms in clouds and falls to earth. Classroom discussion of data and students' efforts to make sense of discrepant data provide opportunities for students to share their reasoning, experience challenges to data and interpretations, and create revised understandings based on the exchange of interpretations. Findings from a field trial in sixteen school districts favored the treatment over business-as-usual groups for science learning, vocabulary, and science content in written measures.

Integrated science and literacy approaches have also proven effective with adolescents in the 14- to 16-year-old age range. In particular, Reading Apprenticeship (RA) focuses students on careful and thoughtful reading of science texts in support of their inquiry, including their explicit attention to what, how, and why they are reading and the understanding they are achieving (Schoenbach et al., 2012). In a large randomized field trial (approximately 5,000 students), students in tenth-grade biology with teachers who had received the RA professional development outperformed those in business-as-usual biology classes on standardized assessments of English language arts, reading comprehension, and biology. Effect sizes indicated an advantage for the treatment group of about one year at the end of the study (Greenleaf et al., 2011). The fundamentals of the RA model have since been incorporated into Project READI's (Goldman et al., 2009) work in science. Students engage in text-based inquiry for purposes of constructing explanatory models of science phenomena that rely on cross-cutting concepts (e.g., patterns, cause and effect, structure and function) (Greenleaf, Brown, Goldman, & Ko, 2013). For example, in one implementation students used texts to find out why drinking too much water is as dangerous as drinking too little, in the context of studying biological homeostasis.

Teaching Literature and Interpretive Practices

When reading literature, experts in literary analysis construct interpretations—connotations and thematic inferences about the human condition—whereas novices such as high school and college students do not (Graves & Frederiksen, 1996). These findings are consistent with results indicating that few high school students are successful in going beyond the literal meaning of literary texts (NAEP, 2009). Indeed, research indicates that typical literature instruction emphasizes literal comprehension of the plot and some attention to characterization, with high dependence on teacher-directed instruction (Nystrand, Gamoran, Kachur, & Prendergast, 1997), perhaps because teachers find it difficult to help students move from literal to interpretive strategies for literary understanding (Marshall, Smagorinsky, & Smith, 1995). This situation contrasts with an inquiry approach to literature; that is, one that emphasizes the tentative nature of literary interpretation and affords opportunities for adolescents to explore the ideas, possibilities, emotions, and perspectives of others on the human condition and to compare them with their own (Applebee, Burroughs, & Stevens, 2000; Langer, 2010; Lee, 2011; Olshavsky, 1976; Rosenblatt, 1978). Perspectives in literary texts are conveyed through many elements, including the events and sequences of events, the characters, the dilemmas, the solutions, the emotions conveyed in the narrative, and how language and structure are used to convey these elements (Hillocks & Ludlow, 1984; Rabinowitz, 1987; Scholes, 1985).

Instructional approaches that support an inquiry stance to literature emphasize close reading of text in conjunction with classroom discussion in which students do the intellectual work of constructing thematic and symbolic as well as literal meanings (Langer, 2010; Lee, 2006; Schoenbach et al., 2012). Doing so involves putting forth proposals or claims that go beyond the literal actions or events in the story; for example, that the tragic flaw in the main character of *Hunger Games* is her loyalty to her family. Support for claims of this type draws on both the text in question and on knowledge of other texts, personal beliefs, belief systems (social, political, philosophical, or religious), or literary theories (Appleman, 2000; Lee, 2014; Schoenbach et al., 2012). Literary analysis requires combining knowledge of human nature with knowledge about literary and rhetorical communication practices, for example, that authors make intentional choices about plot structure (e.g., story events are told in chronological order or are relayed through flashbacks), character types (e.g., anti-hero, trickster), and rhetorical devices (e.g., irony, dialogue, first-person narration) in order to convey their messages (Applebee et al., 2000; Lee, 2011, 2014; Olshavsky, 1976; Rabinowitz, 1987; Smith & Hillocks, 1988). Thus, fundamental to teaching literary analysis is making students aware of these conventions and providing opportunities for students to argue with one another as well as with the author about the message, using evidence from the text and reasoning that connects the evidence to the claim.

There are a number of descriptive accounts of efforts to create inquiry-focused literature classrooms, including Langer (2010), Lee (2001, 2007) and Smith and Hillocks (1988). Recently, the Project READI literature design team designed and implemented an approach to instruction that involves closely reading literary texts multiple times, carefully analyzing language use (e.g., repetition of particular words or phrases), and applying criteria for specific themes and motifs (e.g., what counts as evidence of heroism? Of cowardice?) as well as for literary conventions and rhetorical devices (e.g., Why do we think this is a symbol? What is it a symbol of?) (Lee et al., 2014; Sosa, Hall, Goldman, & Lee, under review). The design includes Lee's cultural modeling approach as a means of providing adolescents entry points to literary analysis (Lee, 2007). Teachers select texts from popular culture (cultural data sets) that manifest the rhetorical devices students will encounter in a school-assigned story or novel. Students discuss interpretations of the cultural data sets, making explicit how they know that, for example, particular song lyrics are not meant literally (e.g., the song lyrics to the Academy Award winning song "Let It Go."[1]) Having gone "meta" on the cultural data sets makes students aware of the interpretive practices they already use and they can then apply them to school texts.

Summary

In each of the disciplines discussed here, emphasis in the pedagogical approach has been on all forms of representation typical of the discipline and modes of language—listening, speaking, reading, and writing. Although we emphasized discipline-specific features in this section, these efforts also reveal some features of effective instruction that apply across disciplines and that we will expand on later: the need for engaging texts and tasks and the important role of discussion in building adolescents' knowledge of disciplinary literacy practices as well as content. Through discussion, students can try out ideas, hear alternative interpretations or counterarguments, expand their knowledge, and revise their thinking in socially supportive contexts.

Literacy Is More Than Reading: The Role of Discussion

Although the classic portrait of the successful reader shows a solitary person curled up with a book, much reading comprehension and most effective reading instruction integrate and depend on discussion. Discussion is not a frequent feature of US classrooms (Nystrand & Gamoran, 1991), but considerable evidence suggests it is likely to be present in classrooms where students acquire high-level literacy skills (Lawrence & Snow, 2010). Furthermore, discussion skills themselves are identified as a component of being college and career ready (CCSSO, 2010), increasing the likelihood that they will receive focused instructional attention. Classroom discussion is hypothesized to promote students' literacy skills via several routes: increasing engagement, building content knowledge presupposed by the text, revealing to students teachers' and classmates' alternative perspectives and interpretations, and providing opportunities for students to practice orally the language and thinking skills they need to apply in reading and writing.

A recent meta-analysis of interventions focused on classroom discussions by students in kindergarten through fifth grade examined effects of a wide range of discussion-based interventions on comprehension and learning from text (Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009). The approaches reviewed shared an emphasis on classroom discussion that valued the exploration of ideas and development of understanding through discussion, often referred to as a dialogic rather than monologic orientation (Nystrand et al., 1997). The Murphy et al. (2009) meta-analysis showed that in classrooms where these particular approaches were used the ratio of student to teacher talk increased and students were more engaged. The meta-analysis reveals the feasibility and utility of discussion even for preadolescents.

With a relatively large sample of low- and high-achieving middle and high school adolescents (across approximately 80 schools), Applebee, Langer, Nystrand, and Gamoran (2003) found that dialogic classroom discussion was significantly related to performance on tasks requiring students to adopt interpretive as opposed to literal stances in literature. Langer (2010) stressed that this type of discussion needs to move students from looking for the point of a story to exploring the possible through literary works. In other words, literature can be a vehicle for exploring what might be rather than what is. Engaging adolescents in these conversations requires that teachers invite students to develop their ideas, listen carefully to the ideas of others, and use multiple perspectives to enrich interpretation of literary works. Prompts for discussion are designed to move students through a series of orientations, or stances, toward text: initial understanding (e.g., What images catch your attention as you read?), developing ideas and multiple perspectives (e.g., What are you noticing about the ideas?), learning from the text (e.g., What does this story do to help you understand about the character's culture? How does it help you understand your world?), taking a critical stance (e.g., What are you noticing about the style of the text?), and going beyond (e.g., How does this story help you understand what is and what might be?). It should be evident that these types of discussions require literary works that are sufficiently complex, challenging, and interesting to the students.

Classroom discussion is the primary means of implementing Lee's (2007) cultural modeling approach. As discussed earlier, cultural modeling is designed to make students explicitly aware of the processing they are doing to understand text; its goal is to have students externalize how they know that some object or phrase is to be interpreted beyond the literal meaning. Making this process explicit allows students to apply it to their comprehension and interpretation of school texts. Classroom discussion is initially led by teachers but gradually gets taken over by students (Lee, 2001).

Classroom discussion has also significantly increased conceptual skills and knowledge in mathematics and science. In mathematics, O'Connor and colleagues examined the impact of introducing a conceptually based mathematics program along with the type of dialogic discourse that Langer and Lee used in their interventions. Of course, the prompts were appropriate to mathematics thinking and to the curricular content being taught to the participants (10- to 14-year-olds) (Chapin & O'Connor, 2012; O'Connor, Michaels, & Chapin, in press). For example, students are encouraged to provide multiple answers to a problem along with an explanation of how they got the answer and why that method is reasonable. If answers agree but were arrived at using different methods, students are asked to think about why the methods agreed. If answers conflict, teachers elicit comments from students about the mathematical reasonableness of the answers. Teachers deepen the mathematics of conversations by revoicing students' contributions, sometimes introducing math-appropriate language (e.g., associative and distributive principles). Over the course of instruction, students take up these forms of mathematical reasoning and speaking. O'Connor and colleagues found that student gains on standardized achievement tests exceeded those in comparison classrooms, as well as those in one of the highest scoring districts in the state (Chapin & O'Connor, 2012; O'Connor & Michaels, 2011).

Similar characteristics of classroom talk are found in efforts to promote inquiry-oriented science in elementary and middle school classrooms. When these norms for interacting and conversational routines are established, there is visible development of student-generated scientific argumentation (Osborne, Erduran, & Simon, 2004; Ryu & Sandoval, 2012).

A program designed specifically to improve students' argumentation skills across the content areas is *Word Generation.* Word Generation introduces engaging topics and supports small- and large-group discussion about those topics as a stimulus to more purposeful reading and writing for authentic audiences. Texts provided to start Word Generation discussions provide models of the academic language forms (precise word choice, compact sentence structure, use of nominalizations, avoidance of evaluation, etc.; see Snow, 2010) that students will be expected to use in their own writing. In addition, Word Generation incorporates activities requiring students to analyze texts to understand speakers' or characters' perspectives. The program is effective in supporting vocabulary learning (Snow, Lawrence, & White, 2009), with larger-than-expected effects for low-scoring students in low-scoring schools on reading comprehension as well as vocabulary (Lawrence, Snow, & Francis, under review). Most saliently, the effects of the program are significantly mediated by the quality of the discussion in which the class engages (Lawrence, Paré-Blagoev, Crosson, & Snow, in press). This result supports the claim that discussion is the active ingredient accounting for outcomes.

It is worth noting that the instructional approaches highlighted in our earlier discussion of disciplinary literacy rely heavily on text-focused discussion. For example, classroom discussion plays a key role in the Reading Apprenticeship model of integrating biology and literacy for 14- to 18-year-olds (Greenleaf et al., 2011; Schoenbach et al., 2012), Project READI's designs of instruction, Scientist's Notebook, and in the approaches to history highlighted in that section.

Efficacy data on classroom discussions is often difficult to obtain because discussions are typically part of more complex interventions that involve multiple pedagogical strategies, a variety of tasks, and texts. Discussions are, however, important in their own right because they make thinking visible. They create opportunities for students to juxtapose their own understanding to that of their classmates and for teachers to assess student understanding.

Purpose and Engagement

The approaches to integrating literacy instruction with disciplinary subject matter instruction discussed thus far incorporate design elements that are intended to actively engage students in the learning process, consistent with contemporary views of effective learning environments (e.g., Bransford et al., 2000). As we have discussed, successful comprehenders actively engage with text, relying on multiple types of knowledge (e.g., of words, concepts, sentence structures, text structures, genres) as they try to interpret print. They monitor their comprehension and use a range of strategies in response to failures to understand what they are reading (Goldman & Saul, 1990; Palincsar & Brown, 1984; Pressley, 2002; RAND, 2002). They connect ideas within a text with each other and with relevant prior knowledge, ask questions, and explain the ideas and connections (Coté, Goldman, & Saul, 1998; Magliano & Millis, 2003).

The active engagement in making meaning described in the previous paragraph contrasts sharply with the default reading activity in many classrooms, where the recurrent student questions are "Why do we have to learn this?" and "Will this be on the test?" These questions reflect the often purpose less nature of many school tasks and learning experiences, along with the disengagement on the part of all but the most dedicated students. Students view history and science as lists of facts to be memorized, static bodies of information that have little bearing on the present and that are encapsulated in thick textbooks with questions at the end of each chapter. Although the movement to introduce hands-on science into kindergarten through twelfth grade classrooms reflects efforts to reform textbook-based science instruction, these efforts have been criticized for encouraging minds-off science: Students carry out sequences of procedures to get the right answer with little understanding of why. The situation is similar for other areas of the curriculum. Thus, it is little wonder that learning school subject matter commands little interest on the part of youngsters in the twenty-first century.

A wide array of approaches has been designed to counter the demotivating practices typical of American classrooms, and many of these approaches are starting to show promising results. The specifics of the approaches vary, and one of the research tasks of the next decade should be to explore their relative merits.

One set of approaches focuses on ensuring students are interested in a topic before starting to teach about it. Concept-Oriented Reading Instruction, for example, has shown positive effects on reading and on science content learning by building engagement and interest with introductory videos, using direct observation prior to reading, and offering students a choice of texts (Guthrie et al., 1999, 2004). The CREATE team developed a similar approach to science as well as to social studies instruction specifically for 12- to 14-year-olds from homes where the school language is not spoken (August, Branum-Martin, Cardenal-Hagan, & Francis, 2009). Lessons designed to cover district-defined topics were launched with videos, incorporated collaborative learning, and explicitly taught second-language structures in the context of the content instruction. Students showed significantly better performance on assessments of language outcomes and content knowledge than those participating in typical science (August et al., 2009) or social studies instruction (Vaughn et al., 2013).

A second set of approaches capitalizes on students' interest in intriguing questions and effective self-expression to place discussion at the center of the instruction. Discussion-based approaches are designed to address the default in classrooms that teachers talk a lot and students rather little. In addition, comprehension-oriented discussion-based approaches such as Word Generation reflect the theory that students' interest in defending their point of view about a topic or dilemma will provoke information-seeking through reading (Snow, Lawrence, & White, 2009). Word Generation has been effective in promoting high engagement and supporting vocabulary learning among middle grades students, especially when the discussion element is well implemented (Lawrence, Paré-Blagoev, Crosson, & Snow, in press). Effects on reading comprehension have not yet been confirmed for this particular program, but other discussion-based programs show significant effects on comprehension (cf. Murphy et al., 2009).

A third set of approaches starts from the assumption that authentic, discipline-specific inquiry tasks that also make contact with the concerns and interests of adolescents will pique the curiosity of students and draw them into rigorous, challenging, adult-like tasks. As illustrated in the section on integrating literacy and content, in these approaches, reading to learn is embedded in inquiry—a need-to-know situation. Learning is directed toward solving some problem or answering some question appropriate to the content area but formulated so that students are genuinely interested in addressing it. Reading thus becomes a tool for knowing, and motivation derives from engagement with authentic disciplinary tasks. Such practices assume that challenge and rigor motivate rather than discourage.

These various approaches embed principles learned from research on strategy instruction and text processing, as well as from small-scale classroom-based research studies. It is clear that adolescent readers benefit from instruction that exposes them to accessible topics, engaging questions, and authentic tasks. Those authentic tasks should incorporate processes akin to those in which disciplinary experts engage in the process of doing their work (Gee, 1992; Lave & Wenger, 1991), albeit in developmentally appropriate forms. Effective instruction intertwines content and communication (Moje, 2008). For example, when adolescent students are given tasks requiring the construction of historical narratives from information found in multiple documents, they learn to think more critically about what they read and engage in deeper processing of text sources (Hartman & Hartman, 1993; VanSledright, 2002; Wolfe & Goldman, 2005). When fourth- through eighth-graders are given social-science or science dilemmas and a variety of texts that offer evidence supporting different resolutions, they read purposefully to find the evidence that will support their own position in debates and in writing (Snow et al., 2009). In science inquiry environments where adolescent students (12–15 years of age) learn to create arguments that support claims with evidence from multiple sources of information, they show improvements in their reasoning and science content knowledge (Geier et al., 2008; Greenleaf et al. 2013; Linn, Clark, & Slotta, 2003). In literature, when adolescents are made aware of interpretive processes they already use to understand texts from their everyday worlds (e.g., rap songs) and are shown how these processes are relevant to particular literary problems (e.g., symbolism), many experience success at interpreting complex literary works (Lee, 2001, 2007; Levine & Horton, 2013; Sosa et al., under review).

Conclusions

In this chapter we have sketched a rationale for treating adolescent literacy as a separate topic, both because of the developmental challenges that are unique to adolescents in academic settings and because of the new tasks and texts that modern schooling presents to adolescents. We have certainly not covered the entire landscape of work in the field of adolescent literacy. We have chosen instead to focus on four specific points:

1. Comprehension strategy instruction is widely used, often forming the primary (or sole)

focus of reading instruction after third grade. Although it can be useful, strategy instruction falls far short of being a solution to the challenges of adolescent literacy.

2. A major new challenge in adolescence is learning to navigate the distinctive literacy demands of different content areas, often with insufficient guidance from content area teachers about how to do that. Nonetheless, instructional programs that embed literacy instruction deeply into disciplinary inquiry have been shown to be effective both for content area learning and for literacy development.

3. Literacy develops not just through reading but also through discussion. Well-structured and well-focused classroom discussion helps prepare students for many of the demands of literacy by confronting them with different perspectives on a topic and providing opportunities to practice academic language and reasoning. Discussion also helps teachers by making their students' thinking public and accessible.

4. As the intrinsic motivation to read declines during adolescence, it becomes vitally important to attend to student engagement and to provide an explicit purpose for reading in the design of instructional activities. Students learn more, read more widely, and write better when they are engaged in the content and understand how the school tasks set for them relate to their lives.

Proper attention to these four points would revolutionize our theories of literacy development and of reading comprehension processes. Full attention to the challenges of disciplinary literacy would shift our theories of literacy development from assuming a single pathway from word reading to successful text comprehension, and substitute instead an image of a branching in middle childhood, with later growth varying across topic, discipline, and task. Recognizing the role of discussion in promoting comprehension and learning would undermine our view of literacy as a purely cognitive, inside-the-head skill, adding the sociocultural and affective dimension. By studying adolescent readers and how they learn, we discover how closely aligned reading is to thinking and to knowing, and the degree to which success at comprehension requires understanding the purposes behind the cultural and disciplinary practices that are being taught.

Acknowledgments

The writing of this chapter was supported in part by the Institute of Education Sciences, US Department of Education, through the following grants: Reading for Understanding across grades 6 through 12: Evidence-based Argumentation for Disciplinary Learning (R305F100007), Word Generation: An Efficacy Trial (R305A090555), and Catalyzing Comprehension through Discussion and Debate (R305F100026). The opinions expressed are those of the authors and do not represent views of the Institute or the US Department of Education.

Note

1 My power flurries through the air into the ground; My soul is spiraling in frozen fractals all around; And one thought crystallizes like an icy blast; I'm never going back, the past is in the past (Lopez & Lopez, 2013). www.metrolyrics.com/let-it-go-lyrics-idina-menzel.html.

References

Achieve (2013). *The next generation science standards.* Washington, DC: Achieve.

Alexander, P. A., & Jetton, T. L. (2002). Learning from text: A multidimensional and developmental perspective. In M. L. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 285–310). Mahwah, NJ: Erlbaum.

Alvermann, D., & Wilson, A. (2011). Comprehension strategy instruction for multimodal texts in science. *Theory Into Practice*, *50*, 116–124.

Applebee, A. N., Burroughs, R., & Stevens, A. S. (2000). Creating continuity and coherence in high school literature curricula. *Research in the Teaching of English*, *34*, 396–429.

Applebee, A. N., Langer, J. A., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. *American Educational Research Journal*, *40*, 685–730.

Appleman, D. (2000). *Critical encounters in high school English: Teaching literacy theory to adolescents.* New York, NY: Teachers College Press and National Council of Teachers of English.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics*, *24*, 621–635.

August, D., Branum-Martin, L., Cardenas-Hagan, E., & Francis, D. (2009). The impact of an instructional intervention on the science and language learning of middle grade English language learners. *Journal of Research on Educational Effectiveness*, *2*, 345–376.

Bazerman, C. (1985). Physicists reading physics: Schema-laden purposes and purpose-laden schema. *Written Communication*, *2*, 3–23.

Bazerman, C. (1998). Emerging perspectives on the many dimensions of scientific discourse. In E. J. Martin & R. Veel (Eds.), *Reading Science* (pp. 15–30). New York, NY: Routledge.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn.* Washington, DC: National Academy Press.

Bråten, I., & Strømsø, H. I. (2010). When law students read multiple documents about global warming: Examining the role of topic-specific beliefs about the nature of knowledge and knowing. *Instructional Science*, *38*, 635–657.

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to use source information. *Cognition and Instruction*, *20*, 485–522.

Bromme, R., & Goldman, S. R. (2014). The public's bounded understanding of science. *Educational Psychologist*, *49*, 59–69.

Carnegie Council on Advancing Adolescent Literacy (CCAAL). (2010). *Time to act: An agenda for advancing adolescent literacy for college and career success*. New York, NY: Carnegie Corporation of New York.

Cervetti, G., Barber, J., Dorph, R., Pearson, P. D., & Goldschmidt, P. (2012). The impact of an integrated approach to science and literacy in elementary school classrooms. *Journal of Research in Science Teaching*, *49*, 631–658.

Chall, J. (1983). *Stages of reading development.* New York, NY: McGraw-Hill.

Chapin, S., & O'Connor, C. (2012). Project Challenge: Using challenging curriculum and mathematical discourse to help all students learn. In C. Dudley-Marling & S. Michaels, (Eds.), *Places where all children learn* (pp. 113–127). New York, NY: Teachers College Press.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic reasoning in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, *86*, 175–218.

Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes*, *25*, 1–53.

Council of Chief State School Officers (CCSSO). (2010). *Common Core State Standards.* Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers.

De La Paz, S. (2005). Effects of historical reasoning instruction and writing strategy mastery in culturally and academically diverse middle school classrooms. *Journal of Educational Psychology*, *97*, 39–156.

Donovan, S. M., & Bransford, J. D. (Eds.) (2005). *How students learn: History, mathematics and science in the classroom.* Washington, DC: National Academies Press.

Duke, N. K., & Martin, N. M. (2008). Comprehension instruction in action: The elementary classroom. In C. C. Block & S. Paris (Eds.), *Comprehension instruction: Research-based best practices* (pp. 241–257). New York, NY: Guilford.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press.

Gee, J. P. (1992). *The social mind: Language, ideology, and social practice.* New York, NY: Bergin and Garvey.

Geier, R., Blumenfeld, P. C., Marx, R. W., Krajcik, J. S., Fishman, B., Soloway, E., & Clay-Chambers, J. (2008). Standardized test outcomes for students engaged in inquiry-based science curricula in the context of urban reform. *Journal of Research in Science Teaching*, *45*, 922–939.

Goldman, S. R. (2012). Adolescent literacy: Learning and understanding content. *Future of Children*, *22*, 89–116.

Goldman, S. R., & Bisanz, G. (2002). Toward a functional analysis of scientific genres: Implications for understanding and learning processes. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 19–50). Mahwah, NJ: Erlbaum.

Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C., & Brodowinska, K. (2012). Comprehending and learning from internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly*, *47*, 356–381.

Goldman, S. R., Lawless, K. A., Pellegrino, J. W., Braasch, J. L. G., Manning, F. H., & Gomez, K. (2012). A technology for assessing multiple source comprehension: An essential skill of the 21st century. In M. Mayrath, J. Clarke-Midura, & D. H. Robinson (Eds.), *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research* (pp. 171–207). Charlotte, NC: Information Age.

Goldman, S. R., Ozuru, Y., Braasch, J., Manning, F., Lawless, K. Gomez, K., & Slanovits, M. (2011). Literacies for learning: A multiple source comprehension illustration. In N. L. Stein & S. W. Raudenbush (Eds.), *Developmental science goes to school: Implications for policy and practice* (pp. 30–44). New York, NY: Routledge.

Goldman, S. R., & Rakestraw, J. A. Jr. (2000). Structural aspects of constructing meaning from text. In M. L. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 311–335). Mahwah, NJ: Erlbaum.

Goldman, S. R., & Saul, E. U. (1990). Flexibility in text processing: A strategy competition model. *Learning and Individual Differences*, *2*, 181–219.

Goldman, S. R., & Scardamalia, M. (2013). Managing, understanding, applying, and creating knowledge in the information age: Next-generation challenges and opportunities. *Cognition and Instruction*, *31*, 255–269.

Goldman, S. R., Britt, A., Greenleaf, C., Griffin, T., Lawless, K., Lee, C. D., . . . Wiley, J. (2009). *Reading for understanding across grades 6 through 12: Evidence-based argumentation for disciplinary learning.* Funded July, 2010 by the Institute of Education Sciences, US Department of Education, Grant # R305F100007.

Gottfried, A. E. (1985). Academic intrinsic motivation in elementary and junior high school students. *Journal of Educational Psychology*, *77*, 631–645.

Graves, B., & Frederiksen, C. H. (1996). A cognitive study of literary expertise. In R. J. Kruez & M. S. MacNealy (Eds.), *Empirical approaches to literature and aesthetics* (pp. 397–418). Norwood, NJ: Ablex.

Greene, S. (1994). The problems of learning to think like a historian: Writing history in the culture of the classroom. *Educational Psychologist*, *29*, 89–96.

Greenleaf, C. L., Brown, W., Goldman, S. R., & Ko, M. L. (2013, December). *READI for science: Promoting scientific literacy practices through text-based investigations for middle and high school science teachers and students*. Presented at Workshop on Literacy for Science. Washington, DC: National Research Council.

Greenleaf, C. L., Litman, C., Hanson, T. L., Rosen, R., Boscardin, C. K., Herman, J., & Schneider, S. A. (2011). Integrating literacy and science in biology: Teaching and learning impacts of Reading Apprenticeship Professional Development. *American Educational Research Journal*, *48*, 647–717.

Grossman, P., Wineburg, S., & Woolworth, S. (2001). Toward a theory of teacher community. *Teachers College Record*, *103*, 942–1012.

Guthrie, J. T., Anderson, E., Alao, S., & Rinehart, J. (1999). Influences of concept-oriented reading instruction on strategy use and conceptual learning from text. *Elementary School Journal*, *99*, 343–366.

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., . . . Tonks, S. (2004). Increasing reading comprehension and engagement through Concept-Oriented Reading Instruction. *Journal of Educational Psychology*, *96*, 403–423.

Hapgood, S., Magnusson, S. J., & Palincsar, A. S. (2004). Teacher, text, and experience mediating children's learning of scientific inquiry. *Journal of the Learning Sciences*, *13*, 455–506.

Hartman, D. K., & Hartman, J. M. (1993). Reading across texts: Expanding the role of the reader. *The Reading Teacher*, *47*, 202–211.

Heller, R., & Greenleaf, C. (2007). *Literacy instruction in the content areas: Getting to the core of middle and high school improvement*. Washington, DC: Alliance for Excellent Education.

Hillocks, G., & Ludlow, L. H. (1984). A taxonomy of skills in reading and interpreting fiction. *American Educational Research Journal*, *21*, 7–24.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*, 294–303.

Langer, J. A. (2010). *Literature: Literary understanding and literature* . New York, NY: Teachers College Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* New York, NY: Cambridge University Press.

Lawrence, J., Paré-Blagoev, J., Crosson, A., & Snow, C. (in press). Word generation randomized trial: Discussion mediates the impact of program treatment on academic word learning. *American Educational Research Journal.*

Lawrence, J., & Snow, C. (2010). Oral discourse and reading. In M. Kamil, P. D. Pearson, E. Moje, & P. Afflerbach (Eds.), *Handbook of reading research* (Vol. 4, pp. 320–338). New York, NY: Routledge.

Lee, C. D. (2001). Is October Brown Chinese? A cultural modeling activity system for underachieving students. *American Educational Research Journal*, *38*, 97–141.

Lee, C. D. (2006). "Every good-bye ain't gone": Analyzing the cultural underpinnings of classroom talk. *International Journal of Qualitative Studies in Education*, *19*, 305–327.

Lee, C. D. (2007). *Culture, literacy, and learning: Taking bloom in the midst of the whirlwind.* New York, NY: Teachers College Press.

Lee, C. D. (2011). Education and the study of literature. *Scientific Study of Literature*, *1*, 49–58.

Lee, C. D., Goldman, S. R., Levine, S., Sosa, T., George, M. A., Magliano, J., . . . Burkett, C, (2014, April). *Literary reasoning and argumentation: Reconceptualizing pedagogical implications of the domain.* Presented in the Symposium *Expanding pedagogical understandings of literary reasoning and argumentation in middle and high schools: Project READI*. Philadelphia, PA: American Educational Research Association.

Lee, C. D., & Spratley, A. (2010). *Reading in the disciplines: The challenges of adolescent literacy.* New York, NY: Carnegie Corporation of New York.

Lenhart, A. (2013). *Teens, smartphones & texting.* Washington, DC: Pew Research Center. http://pewinternet.org/Reports/2012/Teens-and-smartphones.aspx.

Levine, S., & Horton, W. S. (2013). Using affective appraisal to help readers construct literary interpretations. *Scientific Study of Literature*, *3*, 105–136.

Linn, M. C., Clark, D., & Slotta, J. D. (2003). WISE design for knowledge integration, *Science Education*, *87*, 517–538.

Madden, M., Lenhart, A., Duggan, M., Cortesi, S., & Gasser, U. (2013). *Teens and technology 2013.* Washington, DC: Pew Research Center. www.pewinternet.org/Reports/2013/Teens-and-Tech.aspx.

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction*, *21*, 251–283.

Magnusson, S., & Palincsar, A. S. (1995). Learning environments as a site of science education reform: An illustration using interdisciplinary guided inquiry. *Theory Into Practice*, *34*, 43–50.

Marshall, J., Smagorinsky, P., & Smith, M. W. (1995). *The language of interpretation.* Urbana, IL: National Council of Teachers of English.

McKeown, M. G., Beck, I. L., & Blake, R. G. K. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading Research Quarterly*, *44*, 218–253.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, *22*, 247–288.

Meyer, B. J. F., & Wijekumar, K. (2007). A web-based tutoring system for the structure strategy: Theoretical background, design, and findings. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 347–375). Mahwah, NJ: Erlbaum.

Moje, E. B. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: A call for change. *Journal of Adolescent and Adult Literacy*, *52*, 96–107.

Moje, E. B., Overby, M., Tysvaer, N., & Morris, K. (2008). The complex world of adolescent literacy: Myths, motivations, and mysteries. *Harvard Educational Review*, *78*, 107–154.

Murphy, P. K., Wilkinson, I. A. G., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' high-level comprehension of text: A meta-analysis. *Journal of Educational Psychology*, *101*, 740–764.

National Assessment of Educational Progress (NAEP). (2009). *NAEP 2008 trends in academic progress (NCES 2009–479).* Prepared by Rampey, B. D., Dion, G. S., & Donahue, P. L. for the National Center for Education Statistics, Institute of Education Sciences. Washington, DC: US Department of Education.

National Assessment Governing Board. (2008). *Reading Framework for the 2009 National Assessment of Educational Progress.* Washington, DC: American Institutes for Research.

National Center for Education Statistics. (2013). *The nation's report card: A first look: 2013 mathematics and reading* (NCES 2014-451). Washington, DC: Institute for Education Sciences.

National Reading Panel. (2000). *Teaching children to read: Anevidence-based assessment of the scientific research literatureon reading and its implications for reading instruction*. Washington, DC: National Institutes of Health.

Nokes, J. D., Dole, J. A., & Hacker, D. (2007). Teaching high school students to use heuristics while reading historical texts. *Journal of Educational Psychology, 99*, 492–504.

Nystrand, M., & Gamoran, A. (1991). Instructional discourse, student engagement, and literature achievement. *Research on the Teaching of English*, *25*, 261–290.

Nystrand, M., Gamoran, A., Kachur, R., & Prendergast, C. (1997). *Opening dialogue: Understanding the dynamics of language and learning in the English classroom.* New York, NY: Teachers College Press.

O'Connor, C., & Michaels, S. (2011, September). *Scaling back to look forward: Exploring the results of an in vivo study of Accountable Talk*. Socializing Intelligence Through Academic

Talk and Dialogue—an AERA Research Conference. LRDC, University of Pittsburgh.

O'Connor, C., Michaels, S., & Chapin, S. (in press). "Scaling down" to explore the role of talk in learning: From district intervention to controlled classroom study. In L. B. Resnick, C. Asterhan, & S. N. Clarke (Eds.), *Socializing intelligence through talk and dialogue*. Washington, DC: American Educational Research Association.

Olshavsky, J. E. (1976). Reading as problem solving: An investigation of strategies. *Reading Research Quarterly*, *12,* 654–674.

Organization for Economic Co-operation and Development (OECD). (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006.* Paris, France: Author.

Organization of Economic and Cultural Development. (2013). *PISA 2012: Results in focus.* Paris, France: Organization for Economic Co-operation and Development.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argument in school science. *Journal of Research in Science Teaching*, *41*, 994–1020.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension- fostering and monitoring activities. *Cognition and Instruction*, *1*, 117–175.

Pressley, M. (2002). Comprehension strategies instruction. In C. Collins Block & M. Pressley (Eds.), *Comprehension instruction: Research-based practices* (pp. 11–27). New York, NY: Guilford.

Purcell, K., Buchanan, J., & Friedrich, L. (2013, July). *The impact of digital tools on student writing and how writing is taught in schools.* Washington, DC: Pew Research Center. www.pewinternet.org/2013/07/16/

Rabinowitz, P. J. (1987). *Before reading: Narrative conventions and the politics of interpretation.* Ithaca, NY: Cornell University Press.

RAND Reading Study Group (Snow, C., Chair). (2002). *Reading for understanding, Toward an R&D program in reading comprehension.* Santa Monica, CA: RAND.

Reisman, A. (2012). Reading like a historian: A document-based history curriculum intervention in urban high schools. *Cognition and Instruction*, *30*, 86–112.

Romance, N. R., & Vitale, M. R. (2001). Implementing an in-depth expanded science model in elementary schools: Multi-year findings, research issues, and policy implications. *International Journal of Science Education*, *23*, 373–404.

Rosenblatt, L. (1978). *The reader, the text, the poem: The transactional theory of the literary work.* Carbondale, IL: Southern Illinois University Press.

Rouet, J.-F. (2006). *The skills of document use: From text comprehension to web-based learning.* Mahwah, NJ: Erlbaum.

Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Relevance instructions and goal-focusing in text learning* (pp. 19–52). Greenwich, CT: Information Age.

Rouet, J.-F., Britt, M. A., Mason, R. A., & Perfetti, C. A. (1996). Using multiple sources of evidence to reason about history. *Journal of Educational Psychology*, *88*, 478–493.

Ryu, S., & Sandoval, W. A. (2012). Improvements to elementary children's epistemic understanding from sustained argumentation. *Science Education*, *96*, 488–526.

Schoenbach, R., Greenleaf, C. L., & Murphy, L. (2012). *Reading for understanding: How reading apprenticeship improves disciplinary learning in secondary and college classrooms.* New York, NY: Jossey-Bass.

Scholes, R. (1985). *Textual power, literary theory and the teaching of English*. New Haven, CT: Yale University Press.

Seixas, P. (1994). Students' understanding of historical significance. *Theory and Research in Social Education*, *22*, 281–304.

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, *78*, 40–59.

Smith, M. (1991). *Understanding unreliable narrators: Reading between the lines in the literature classroom*. Urbana, IL: National Council of Teachers of English.

Smith, M. W., & Hillocks, G. (1988). Sensible sequencing: Developing knowledge about literature text by text. *English Journal*, *77*, 44–49.

Snow, C. E. (2010). Academic language and the challenge of reading for learning. *Science*, *328*, 450–452.

Snow, C. E., & Biancarosa, G. (2004). *Reading next: A vision for action and research in middle and high school literacy.* A report to the Carnegie Corporation of New York. Washington, DC: Alliance for Excellent Education.

Snow, C., Lawrence, J., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness*, 2, 325–344.

Snow, C. E., & Moje, E. (2010). What is adolescent literacy? Why is everyone talking about it now? *Phi Delta Kappan*, *91*, 66–69.

Sosa, T., Hall, A, Goldman, S, R., & Lee, C. D. (under review). Developing symbolic interpretation through literary argumentation. *Journal of the Learning Sciences.*

Stadtler, M., & Bromme, R. (2007). Dealing with multiple documents on the WWW: The role of metacognition in the formation of documents models. *International Journal of Computer Supported Collaborative Learning*, *2*, 191–210.

VanSledright, B. A. (2002). Confronting history's interpretive paradox while teaching fifth graders to investigate the past. *American Educational Research Journal*, *39*, 1089–1115.

Vaughn, S., Swanson, E. A., Roberts, G., Wanzek, J., Stillman-Spisak, S. J., Solis, M., & Simmons, D. (2013). Improving reading comprehension and social studies knowledge in middle school. *Reading Research Quarterly*, *48*, 77–93.

Wanzek, J., Vaughn, S., Scammacca, N. K., Metz, N., Murray, C. S., Roberts, G., & Danielson, L. (2012). Extensive reading interventions for students with reading difficulties after grade 3. *Review of Educational Research*, *83*, 163–195.

Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, *83*, 73–87.

Wineburg, S. (2001). *Historical thinking and other unnatural acts: Charting the future of teaching the past*. Philadelphia, PA: Temple University Press.

Wolfe, M. B., & Goldman, S. R. (2005). Relationships between adolescents' text processing and reasoning. *Cognition and Instruction*, *23*, 467–502.

NAME INDEX

Page numbers in italics appear in reference lists.

SUBJECT INDEX