ADVANCES IN PSYCHOLOGY

94

Orthography, Phonology, Morphology, and Meaning

Ram Frost
Leonard Katz
Editors

North-Holland

ORTHOGRAPHY, PHONOLOGY, MORPHOLOGY, AND MEANING

ADVANCES IN PSYCHOLOGY

94

Editors:

G. E. STELMACH

P. A. VROON

NORTH-HOLLAND
AMSTERDAM • LONDON • NEW YORK • TOKYO

ORTHOGRAPHY, PHONOLOGY, MORPHOLOGY, AND MEANING

Edited by

Ram FROST

Department of Psychology
Hebrew University
Jerusalem, Israel

Leonard KATZ

Department of Psychology
University of Connecticut
Storrs, CT, USA

1992

NORTH-HOLLAND
AMSTERDAM • LONDON • NEW YORK • TOKYO

NORTH-HOLLAND
ELSEVIER SCIENCE PUBLISHERS B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

ISBN: 0 444 891404

This book is printed on acid-free paper.

Transferred to digital printing 2006

Contents

Acknowledgments vii

Orthography, Phonology, Morphology, and Meaning:
An Overview
Leonard Katz and Ram Frost 1

PART 1

Language and Orthography

1. Linguistic Awareness and Orthographic Form
Ignatius G. Mattingly 11

2. Reading Consonants and Guessing Vowels: Visual Word Recognition in Hebrew Orthography
Ram Frost and Shlomo Bentin 27

3. Basic Processes in Reading: Is the Orthographic Depth Hypothesis Sinking?
Derek Besner and Marilyn Chapnik Smith 45

4. The Reading Process is Different for Different Orthographies: The Orthographic Depth Hypothesis
Leonard Katz and Ram Frost 67

5. Beyond Orthographic Depth in Reading: Equitable Division of Labor
Mark S. Seidenberg 85

6. Automatic Activation of Linguistic Information in Chinese Character Recognition
Daisy L. Hung, Ovid J. L. Tzeng, and Angela K. Y. Tzeng 119

7. Orthographic Neighborhoods and Visual Word Recognition
Jonathan Grainger 131

8. On the Role of Cohorts or Neighbors in Visual Word Recognition
Neal F. Johnson 147

PART 2

Orthography and Phonology

9. The Relation of Speech to Reading and Writing
Alvin M. Liberman 167

10. On the Relations between Learning to Spell and Learning to Read
Donald Shankweiler and Eric Lundquist 179

11. Phonological Awareness, Reading, and Reading Acquisition: A Survey and Appraisal of Current Knowledge
Shlomo Bentin 193

12. Can Theories of Word Recognition Remain Stubbornly Nonphonological?
Claudia Carello, M. T. Turvey, and Georgije Lukatela 211

13. Reading in English and Chinese: Evidence for a "Universal" Phonological Principle
Charles A. Perfetti, Sulan Zhang, and Iris Berent 227

14. "Assembled" Phonology and Reading: A Case Study in How Theoretical Perspective Shapes Empirical Investigation
Guy C. Van Orden, Gregory O. Stone, Karen L. Garlington, Lori R. Markson, Greta Sue Pinnt, Cynthia M. Simonfy, and Tony Brichetto 249

15. Dual-route Models of Print to Sound: Red Herrings and Real Horses
Kenneth R. Paap, Ronald W. Noel, and Linda S. Johansen 293

16. Strategies and Stress Assignment: Evidence from a Shallow Orthography
Lucia Colombo and Patrizia Tabossi 319

PART 3

Orthography and Lexical Structure

17. Morphological Analysis in Word Recognition
Laurie B. Feldman and Darinka Andjelković 343

18. Units of Representation for Derived Words in the Lexicon
Cristina Burani and Alessandro Laudanna 361

19. Representation and Processing of Morphological Information
Cecile Beauvillain and Juan Segui 377

20. Bilingual Lexical Representation: A Closer Look at Conceptual Representations
Annette M. B. de Groot 389

21. Memory-addressing Mechanisms and Lexical Access
Kenneth Forster. 413

Index of Authors 435

Acknowledgments

This volume was produced with the help of facilities at Haskins Laboratories, New Haven, The Hebrew University, Jerusalem, and the University of Connecticut, Storrs. The work was supported in part by a grant from the National Institute of Child Health and Human Development (HD-01994) to Haskins Laboratories. At Haskins, Yvonne Manning-Jones was responsible for computer production of the manuscripts and graphics. All of the chapters of the book were reviewed, in most cases by other authors. Our Haskins colleague, Kenneth Pugh, although not an author, reviewed several chapters, sometimes more than once each. We are particularly grateful to Ken for his unflagging energy and his insightful comments. Finally, we would like to thank Alvin Liberman, in part for his advice on matters related to this book but mostly for his wisdom over the years on all things pertaining to speech and reading.

Ram Frost
Leonard Katz

This Page Intentionally Left Blank

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

Orthography, Phonology, Morphology, and Meaning: An Overview

Leonard Katz
Department of Psychology
University of Connecticut

Ram Frost
Department of Psychology
Hebrew University

The area of research on printed word recognition has been one of the most active in the field of experimental psychology for well over a decade. This is in part because the behavior under scrutiny is seen as complex enough to be interesting but circumscribed enough to make discovery feasible. It contains many of the theoretical concepts that are part of the cognitive psychologist's standard investigative repertoire (form perception, attention, awareness, information, representation, neural networks, theoretical linguistics—to name a few) and an armamentarium of clever experimental techniques. However, notwithstanding the energetic research effort and despite the fact that there are many points of consensus, major controversies still exist. One central matter is the question of whether to view reading primarily as a linguistic activity or, alternatively, as a process that is subject to the same kinds of learning as other visually based, but nonlinguistic, information.

Our stance on this is that it is quite necessary to take spoken language into consideration when attempting to understand the psychological processes by which reading is accomplished. We need to do this for two kinds of reasons. First, it is well known that writing systems are designed primarily to represent spoken language and, therefore, it seems at least plausible that we should find the imprint of spoken language in the processes that lead from the recognition of the printed word to the comprehension of the phrase. Secondly, in the past two decades of research, there has been a wealth of data supporting this claim. However, the fact that these two statements are not unchallenged leads us to many of the issues that involve the contributors to this volume. Nevertheless, the motivation for the present volume derives from this putative relationship between language and reading; the book takes as its primary issue the question of the degree to which basic processes in reading reflect the structural characteristics of language such as phonology and morphology.

In Part 1, *Language and Orthography*, several chapters point out how the phonological and morphological structures of a language have, historically, often determined the kind of orthography that is adopted for a language. The variety that exists in spoken languages has given rise to a variety of orthographies, each orthography reflecting a unique relationship to its language's structural characteristics. A chapter by Mattingly begins Part 1. Mattingly points out the dependence of writing and reading on spoken language. He describes how constraints on human memory and perception together with the requirement that writing should be a productive system have shaped the kinds of orthographies that have been developed. Mattingly explores the idea that once a reader has learned a particular orthography the reader's intuitions (linguistic awareness) about his or her language are shaped by the particulars of that orthography. An issue is thereby raised: does the process by which reading is accomplished—the psychological process—also vary among languages, reflecting, in processing diversity, the orthographic diversity? To address this, we present research from a variety of different languages covering the spectrum of writing systems. Frost and Bentin's chapter presents evidence suggesting that a special kind of flexibility exists for the reading process: that the Hebrew reader can optimize his or her processing for the kind of information being presented by the printed word. Skilled Hebrew readers are able to take advantage of a particular characteristic of that orthography: morphologically related words in Hebrew are normally printed to include their common consonantal root but omit the vowel information that distinguishes the relatives when they are spoken. Readers can read words printed this way quite efficiently, determining what the specific form of the word is from the context. Nevertheless, despite this efficiency, when Hebrew readers read words in the extended Hebrew orthography (which includes all vowel information), they are flexible enough to adopt a different strategy—one that prefers to utilize the available vowel information.

Within the group of alphabetic orthographies there are large differences in the degree to which writing systems adhere to a strict alphabetic principle, i.e., the principle of an isomorphic relation between letter and phoneme. A question of interest is whether these differences in orthography are reflected in differences in the process of printed word recognition. How different is reading in Spanish, which conforms closely to the alphabetic principle, from reading in Hebrew, in whose orthography the spelling reflects the spoken form of the word only incompletely? A common proposal has been that the more closely an orthography conforms to the alphabetic principle, the more efficiently phonological representations will mediate between print and lexicon. A phonological representation is assumed to be assembled by the reader who makes use of the orthography's correspondences between subword spelling and sound. This proposal has been termed the orthographic depth hypothesis. It is explored in several chapters in this volume, but particularly in Part 1, in a paper by Besner and Smith and a paper by Katz and Frost. Besner and Smith point out the evidence (both rational and empirical) that a strong version of the orthographic depth hypothesis, in which phonology is deemed to be the *only* code for word recognition, is not viable. Katz and Frost agree with that point but suggest a weaker version of the hypothesis, which has experimental support. According to Katz and Frost's version of the orthographic depth hypothesis, full recovery of a printed word's phonological structure can involve both prelexical (i.e., assembled) phonology and lexical (i.e., stored) phonology. Their main point is that the relative dependence on prelexical versus lexical sources for phonology is a function of the orthography's depth. For example, in Spanish, a shallow orthography, phonology should be more easily assembled because the letter-phoneme relation is fairly simple and consistent. Because assembled phonology is easily obtained, it should be, more often, the actual mechanism for obtaining a word's phonological representation than in a deeper orthography like English, in which letter-phoneme correspondences are more complex.

Even proponents of the orthographic depth hypothesis have not made claims that word recognition in Chinese is aided by phonological representations mediating between the grapheme and the lexical entry. Chinese seems to be a poor candidate for such a claim because the phonologic morpheme in a Chinese word is typically a less precise, less reliable cue for pronunciation than is an alphabetic spelling. Nevertheless, Hung, Tzeng, and Tzeng presents data suggesting that, even in Chinese, phonology plays a role in word recognition. If their claim continues to be supported, its import cannot be understated: Because a phonological representation in Chinese is so much less specific than an alphabetic prelexical phonological representation in determining a unique word, its use in spite of this ambiguity suggests that phonological representations may be even more pervasive in less phonologically opaque (i.e., syllabic, alphabetic) writing systems.

Seidenberg, while acknowledging that orthographic depth plays some role in word recognition, argues that we ought not to put too much emphasis on it: that reading processes for different orthographies are more alike than they are different. And he emphasizes the need to go beyond the question of whether or not the word recognition process employs assembled phonology or not; Seidenberg feels that we need to understand the nature of the flexibility that the reading process displays with regard to how it distributes its resources for word recognition: to the use of assembled phonology, on the one hand, and to the use of visual-orthographic representations, on the other. In order to understand this flexibility, we need to understand better the cognitive structures involved in reading and how their various limitations and constraints are played off against each other.

The two chapters by Grainger and by Johnson are not concerned with cross-orthography effects or with the question of phonologic involvement. Rather, they use a single orthography and focus on the role of the lexicon's neighborhoods of words in word perception. A neighborhood, in the internal lexicon, consists of a set of words that have, to one degree or another, spellings that are similar. Neighborhood theories view the identification of a particular printed word as a process of differentiating that word from the other words in its neighborhood. Thus, neighborhood theories take *similarity*, usually defined orthographically, as the important determiner of word recognition. Mutual facilitation and/or inhibition between the target word and its neighbors (based on the degree of their similarities) can affect the speed and success of identification. These two chapters, together, review the state of the art in neighborhood theory and also contribute new experimental results.

In Part 2, *Orthography and Phonology*, two major controversial issues are treated. The first continues the concern over the role of phonology in lexical access and focusses on experimental work on this topic: Under what conditions does the reader address the lexicon by assembling a phonological representation from the orthographic form and under what conditions is the orthographic form itself used directly for access? How the lexicon is addressed is a matter of considerable importance because it is vital, in understanding the reading process, to know the nature of the internal representations generated in the information flow. Only then can we attempt to characterize the information theoretic structure of the reading process and to understand its division and allocation of processing resources. There is also a second issue, one of considerable practical importance. The problem of understanding how readers achieve word recognition bears directly on the question of the preferred method of teaching reading. Here an often bitter conflict has raged in the education community for decades. Although it is an oversimplification to say that there are only two opposing points of view, nevertheless, it is fair to say that any of the current approaches to the teaching of reading either emphasizes the importance of decoding skills for the beginning reader (and therefore emphasizes the importance of learning to produce phonological representations) or minimizes it (emphasizing, instead, the importance of visual-orthographic representations such as in "whole-word" reading). While the debate within experimental psychology has been more civil, it has not been less volatile, with evidence accumulating for both positions.

Liberman's chapter discusses the theoretical implications of something that should have been obvious but, unfortunately, has not been: that listening to speech is both easy and universal while reading is neither; reading is, in a real sense, unnatural. One implication of this is that a theoretical explanation of reading must be very different from a theoretical explanation of listening. As Liberman points out, theorists usually fail to make this contrast and theories of reading and theories of speech perception have both suffered as a consequence. Liberman stresses the dependence of the reading process on speech; he raises the often expressed idea that there is a link between a child's success in learning to read and the child's ability to analyze the phonological structure of his or her language—at least that particular phonology that the orthography maps on to. This understanding then allows the child to utilize the correspondence between letter and phoneme to decode the printed letter string: It allows recognition of that word as if it were in the child's own spoken lexicon. A deficiency in this ability should be particularly damaging for children who must learn an alphabetic orthography because the child must first understand that spoken words are composed of phonemes, a requirement that is difficult because it requires a cognitive awareness of phonological representations that are normally processed automatically, without awareness, in speech perception and production.

If children who are unable to learn to read easily have a problem in the phonological domain, might this problem be expressed in other ways as well? Shankweiler and Lundquist's chapter widens the focus of discussion to include spelling ability. Their argument is that spelling ability is largely a function of linguistic skills which are indexed by phonological awareness, the ability to analyze a spoken word into its phonologic components. Although it has been claimed by others that spelling ability reflects visual-orthographic memory, the authors make the argument that the underlying psychological structure for spelling has, at the least, a strong phonological component. The chapter by Bentin also deals with phonological awareness, showing that this ability is responsible for much of the variation in reading ability observed among children. His chapter discusses the causes of phonological awareness. Bentin's data demonstrate that maturation and instruction in reading both affect the development of children's phonological awareness. Thus, he shows us that learning to read is not only influenced by phonological awareness but, as learning occurs, it furthers the growth of phonological awareness.

Three chapters each make the strong claim that phonology plays an obligatory part in the word recognition process. Each uses a different experimental methodology, which furthers the generality of the claim. Carello, Turvey, and Lukatela marshal the arguments and evidence that printed word recognition, as observed in the lexical decision task, inevitably (although not exclusively) involves assembling a phonological representation: that word recognition relies mainly on the information provided via the alphabetic principle. Their evidence comes largely from experiments in the shallow Serbo-Croatian orthography, in which there is a strong correspondence between grapheme and phoneme. Perfetti, Zhang, and Berent; present English data that buttress the claim that there is obligatory assembled phonology in lexical access. Their procedure involves the injection of printed phonological information into the processing stream shortly after the tachistoscopic presentation of the target word: the target is masked by a pseudoword that is either phonologically similar to the target or not. Because the time from target onset to mask onset can be quite brief in their experiments (e.g., 25 to 65 ms), the effects of phonological similarity seem to be attributable to an early phase of processing. The authors argue further that phonological processes play a role throughout the entire reading process—not just in word recognition but in comprehension as well. Finally, they compare reading in English and Chinese and suggest that the similarity between the two orthographies with regard to the role of phonology is greater than their differences. Further support for the position that phonology is an inherent part of the word recognition process appears in the paper by Van Orden, Stone, Garlington, Markson, Pinnt, Simonfy, and Brichetto. Their experimental paradigms include sentence verification, lexical decision, categorization, and proofreading: a mix of techniques that all provide converging evidence. Their paper discusses the notion of covariant learning in which the relation between orthography and phonology is expressed as a statistical relation instead of as a collection of rule based correspondences between letter and phoneme. Covariant structural relations fit more naturally into neural network formulations of the word recognition process, the class of models preferred by the authors. Finally, both this chapter and the chapter by Carello, Turvey, and Lukatela make the point that, in the debate between phonological versus visual-orthographic representations, there has been little effort given to collecting positive evidence in favor of the visual-orthographic hypothesis. Rather, the emphasis has been on showing that phonology does (or does not) have an effect on word recognition. In cases in which there was no evidence of phonological effects, it has often been assumed that the alternative hypothesis—visual-orthographic coding—was proven by default. Thus, decisions have been made based on failures to reject the null hypothesis, a notoriously poor strategy.

Paap, Noel, and Johansen lay out the theory and data for dual route theories of printed word pronunciation. The dual routes are the pathways carrying the phonological and orthographic information used for pronouncing printed words. Although the chapter is oriented toward the response of naming printed words, the authors also discuss dual route theory for tasks in which a silent response is made (e.g., a semantic decision about the target word). They describe their chapter as a tutorial for the newcomer but its clarity will serve the advanced researcher as well by dispelling many non-issues (the authors call them "red herrings") that have cluttered discussion in this area. Several other chapters (e.g., those cited in the previous paragraph) also suggest that lexical access depends on dual codes, i.e., dual representations. If one argues for dual codes, there arises the question of what it is that determines whether the visual-orthographic or phonologic code is used in a particular instance of word recognition. In standard dual route theory, the factors that are said to affect the lexical access code are stable and fixed—word frequency, spelling regularity, and perhaps, characteristics of the orthography (e.g., orthographic depth). However, others suggest that subject strategies may play an additional role in determining which code is used: these strategies are affected by more variable characteristics of the reading context. For example, Colombo and Tabossi demonstrate that subjects are capable of making fine adjustments to their strategy for naming words in Italian. Naming (pronouncing a word aloud as fast as possible) is a task that requires subjects to take into account the word's syllable stress. Colombo and Tabossi's show that subjects' response times are affected by the kinds of word stress that the experimenters build into their lists of stimulus words. Subjects could be induced to use either a sublexical (i.e., assembled) or a lexical strategy, assigning stress one way or the other in agreement with the list bias. It will be of major interest if it turns out that the reader has exquisite control of such presumably low-level components of the reading process. The issue of strategic use necessarily raises questions about the allocation of attentional resources to the various components of the reading process: both where and how attention is distributed. These are likely to be questions researchers will be increasingly concerned with, as the chapters of Paap, et al., Seidenberg, and Van Orden et al. suggest. Finally, although Colombo and Tabossi's data show there is flexibility in subjects' coding strategies, an implication of their experiments is that subjects will not choose to use a phonological strategy under normal circumstances in reading. This poses a challenge to those who argue for the ubiquity of phonological coding: to show that the evidence in favor of phonology is not an artifact of particular experimental paradigms but can be generalized to normal reading.

Orthographies convey not only phonologic but also morphologic information. In fact (we must sometimes remind ourselves), phonological and orthographic information are simply vehicles for the activation of morphological information. Part 3, *Orthography and Lexical Structure*, is concerned, in part, with the representation of morphological information in lexicon. Morphological information includes those parts of a word that convey (1) syntactic inflection (e.g., number, case, gender, tense, mood, etc.) or (2) derivational relations (e.g., nominalizations of adjectives, formation of diminutives, etc.) as well as (3) word roots.

Morphological relatives are different words that have a common root. For example, WALK/WALKED/WALKING all have the same root but different inflectional morphemes and WEIGH/WEIGHT/WEIGHTY are derivationally related. Are relations between morphological relatives represented in the lexicon by connections that are specifically morphological? There is considerable evidence of connections between words in the same morphological family. However, much of this evidence may be confounded by one or more artifacts. As Feldman and Andjelkovic point out, it is quite difficult to demonstrate pure morphological relations among words experimentally. One major experimental design problem is the requirement to separate the effects of the strictly morphological relationships between words from their formal similarity (i.e., related words tend to sound and look alike). Subjects may respond similarly to two words for either of those two reasons, producing the appearance of a morphological effect but not the reality. Additionally, morphological relatives will nearly always have a semantic communality, and experimental effects may be found that are attributable to this factor, further obscuring the observation of purely morphological relationships between words. Feldman and Andjelković address much of their chapter to the solution of these problems.

What is the organization in lexicon of words that are morphological relatives? Despite the evidence that some such organization does exist, the details of morphological organization are far from clear. After briefly discussing the various models of morphological representation, Burani and Laudanna go on to focus on the specific question of how derivational relatives are organized. Beauvillain and Segui are also concerned with the organization of morphological relatives. In addition, they discuss the forms that the lexical representations of derived words may take. Does a derived word's lexical entry consist of distinct components, viz., a root and one or more derivational morphemes? Or, instead, is the word stored "intact," as a whole unit? This question is relevant to the third question they confront: What is the process by which derived words are recognized? If words are stored as a set of morphological components, then a printed target word must first be decomposed into its components (root and derivational morpheme) before recognition can occur. If not, then each relative must be accessed as a whole word, without specific reference to its morphological characteristics. De Groot's chapter also deals with relations between words in the lexicon. Here, the lexicon of interest is the bilingual lexicon. The relation between words in the two languages involves their semantic communality. Given the problem she has set herself, de Groot's focus is not on the lexical level itself and its orthographic, phonologic, and morphologic information but, rather, on conceptual memory. Evidence from several semantic memory paradigms are used. De Groot's work is informative beyond the question of the bilingual lexicon because it raises questions about the nature of the representation of meaning itself.

Although much attention is paid to questions of representation in printed word recognition (i.e., the kinds of codes that occur in the information flow from print to lexicon), there has been relatively little concern about other aspects of the process. In spite of the fact that our primary tool for investigating word recognition is the lexical decision task, our knowledge of the mechanism by which lexical access is achieved is still uncertain. Forster's chapter redresses some of this imbalance with an organized discussion of the classes of models that are capable of accounting for the lexical search process: how the target word is selected from among the tens of thousands of words in the lexicon. His paper describes and evaluates the various models on rational grounds of efficiency and plausibility and in terms of their adequacy for explaining the major phenomena of word recognition: the word frequency effect, the repetition effect, neighborhood effects, etc. The chapter will certainly form the nucleus for a revitalized discussion on lexical search models.

This book presented a unique opportunity to bring together leading researchers who address the question of printed word recognition from a linguistic perspective. The chapters reveal the interactive nature of their work in this field; a measure of this closeness can be found in the high degree of mutual citations. From the mix of tutorial articles, critical articles, data papers, and theory papers, comes a portrait of the field; this includes not only a picture of current theory and data but a view of the directions in which this vigorous research area is moving.

PART 1

Language and Orthography

This Page Intentionally Left Blank

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 1

Linguistic Awareness and Orthographic Form

Ignatius G. Mattingly
Department of Linguistics, University of Connecticut, Storrs

Introduction: The taxonomy of writing systems

To impose some pattern on the vast array of writing systems, present and past,[1] several investigators have proposed typologies of writing (Gelb, 1963; Hill, 1967; Sampson, 1985; DeFrancis, 1989; see DeFrancis for a review). While typology for its own sake may seem a dubious goal, these proposals bring to notice certain interesting questions.

Consider first the problem posed by logograms. It is generally recognized that the signs found in writing fall into two broad categories: logographic and phonographic. Logograms stand for words, or more precisely, morphemes. Thus, in Sumerian writing, there is a logogram that stands for the morpheme **ti**, 'arrow.' Phonographic signs stand for something phonological: syllables or phonemic segments. Thus, in Old Persian, there is a sign for the syllable **da**, and in Greek alphabetic writing, a sign for the vowel **a**. This distinction suggests that writing systems might be classified according to whether they are logographic or phonographic. But the attempt to impose such a classification is embarrassed by the fact that while the many systems in the West Semitic tradition are indeed essentially phonographic and have no logograms, writing systems of all other traditions use both logograms and phonograms. There have been no purely logographic systems: phonographic signs are found in all traditions.

In these circumstances, Gelb sets up a hybrid category "word-syllabic," in which he includes Sumerian, Egyptian (whose phonographic signs he takes to be syllabic[2]), and

[1] It will be assumed here, following Gelb (1963), Jensen (1970), DeFrancis (1989) and others, that there are six major orthographic traditions: (1) Mesopotamian cuneiform, beginning with Sumerian (c. 3100 B.C.) and including Akkadian, cuneiform Hittite, Urartian, Hurrian, Elamite, Old Persian; (2) Cretan, including Minoan Linear A, Mycenaean Greek Linear B, Cypriote, and Hittite hieroglyphics, all probably derived from a common source (c. 2000 B.C.); (3) Chinese, beginning with Chinese itself (c. 1300 B.C.) and including Korean nonalphabetic writing and Japanese; (4) Mayan (c. 300 A.D.); (5) Egyptian (c. 3000 B.C); (6) West Semitic, beginning with Phoenician (c. 1600 B.C.) and including Ras Shamrah cuneiform, Old Hebrew, South Arabic, Aramaic, and Greek alphabetic writing. From Aramaic derive Hebrew, Arabic, and many others; from Greek derive Etruscan, Latin, and many others. Germanic runes and Korean alphabetic writing probably belong in this tradition also, though the derivations are not clear. All but the most dogmatic monogeneticists would agree that the Mesopotamian, Cretan, Chinese, and Minoan traditions are probably independent developments. But some scholars (e.g., Driver, 1976; Ray, 1986) would derive Egyptian writing from Mesopotamian, and some (e.g., Driver, 1976), with somewhat greater plausibility, would derive West Semitic from Egyptian.

[2] Egyptologists and most other students of writing believe that Egyptian phonographic signs stand for consonants, the vowels not being regularly transcribed. But according to Gelb, they stand instead for generalized syllables, e.g., the Egyptian sign usually interpreted as consonantal **w** actually stands for **wa**,

Chinese. Other orthographic taxonomists allow a writing system to belong to two different categories. Thus for Hill, Egyptian is both "phonemic" and "morphemic" and for Sampson, Japanese is both "phonographic" and "logographic." DeFrancis, recognizing that logograms are neither necessary nor sufficient for an orthography, more sensibly treats logography as an optional accompaniment to various phonographic categories. But the question of interest is *why* logograms should play only this secondary role, why there have been no pure logographies.

A second problem arises in sorting out the phonographic categories. Here one might recognize, with DeFrancis, systems like Sumerian or Linear B, in which the phonographic signs stands for syllables; systems like Egyptian or Phoenician, in which they stand for consonants; and systems like Greek or English, in which they stand for both consonants and vowels (*plene* systems).

The distinction between consonantal and *plene* systems, however, proves to be less than rigid. In Egyptian, the letters for **j**, **w**, and **?** are used to write **i**, **u** and **a**, respectively, in foreign names (Gelb, 1963). Phoenician, indeed, is a strictly consonantal, but the other "consonantal" systems deriving from it all have some convention for transcribing vowels when necessary. For example, in Aramaic, the letters *yodh*, *waw*, and *he* (or *aleph*) were used to write final **i**, **u**, and **a**, respectively, and to render vowels in foreign names (Cross & Freedman, 1952). In Masoretic Hebrew, Arabic, and various Indic systems, vowels are regularly indicated by diacritic marks on consonant letters. And, of course, the first clearly *plene* system, the Greek alphabet, is a development from the Phoenician consonantal system. The taxonomist thus has to decide where to draw the line between essentially consonantal systems, hybrid systems, and undoubted *plene* systems. Perhaps the wisest course is the one followed by Sampson: simply to classify all these systems as "segmental."

Syllabic systems, in contrast, are clearly a separate category and present no problem to the taxonomist. There is no writing system that must be regarded as a hybrid between a syllabic and a segmental system. Syllabic systems show no tendency to analyze syllables into segments. What is found, rather, is that when analysis becomes necessary, complex syllables are analyzed into simpler syllables. Thus, neither the Mesopotamian nor the Mayan syllabaries had signs for all possible $C_1V_1C_2$ syllables in their respective languages. Instead, such syllables were written in Mesopotamian as if they were $C_1V_1 + V_1C_2$ (Driver, 1976) and in Mayan as if they were $C_1V_1 + C_2V_1$ (Kelley, 1976)). Similarly, Greek $C_1C_2V_1$...syllables were written in Linear B as C_1V_1+ C_2V_1 +...(Ventris & Chadwick, 1973). Nor, despite suggestions to the contrary by Gelb and DeFrancis, has a syllabic system ever developed into a segmental system, or conversely.[3] It cannot be excluded that the Egyptians may, as DeFrancis says (following Ray, 1986), have gotten the idea of writing from the Sumerians. But there is certainly no reason to believe that they borrowed the idea of *syllabic* writing from the Sumerians and then adapted it to consonantal writing, in the way that the Greeks may be said to have borrowed the idea of consonantal writing

wi, **we**, **wu**, or **wo**, according to context. It is obviously difficult to distinguish these two accounts empirically. The only support Gelb offers for his position is that "the development from a logographic to a consonantal writing, as generally accepted by the Egyptologists, is unknown and unthinkable in the history of writing" (Gelb 1963, p. 78). But this argument is clearly circular (Edgerton, 1952; Mattingly,1985).

[3]Gelb (1952, 1963) proposed some cases in which syllabic systems are supposed to have developed into segmental systems ; but see Edgerton (1952). Ethiopic writing, derived from the West Semitic consonantal tradition, might be viewed as a syllabic system derived from a segmental system, because the signs do correspond to syllables. But, with a few exceptions, each sign actually consists of a consonant letter plus a vowel mark, except that **a** is left unmarked. As in the case of Indic systems, one could argue about whether this is a consonantal or a *plene* system, but it is certainly not a syllabic system (Sampson, 1985).

from the Phoenicians and adapted it to *plene* writing. The various orthographic traditions are remarkably self-consistent in this matter. The Mesopotamian, Chinese, Cretan and Mayan traditions began and remained syllabic; the Egyptian and West Semitic traditions began and remained segmental.

If the main purpose here were to arrive at a taxonomy of writing systems, the conclusion would have to be that there are two primary categories: syllabic and segmental. Either of these may or may not be accompanied by logograms. Transcription of vowels in segmental systems is a matter of degree, with Phoenician at one end of the scale and Greek at the other. The interesting question, however, particularly given the degree of overlap or hybridization that is found between logographic and phonographic categories, and between consonantal and *plene* categories, is why the syllabic and segmental categories have remained so distinct.

In an attempt to answer the questions just posed, it is necessary to consider why an orthography can make reading and writing possible, what constraints there are on the form of orthographies, how orthographies could have been invented, and what happens when orthographies are transmitted from one culture to another.

Why reading and writing are possible[4]

When a listener has just heard an utterance in a language he knows, he has available for a brief time not only his understanding of the semantic and pragmatic content of the utterance (the speaker's *message*), but also a mental representation of its linguistic structure. The basis for this claim is that a linguist, by analyzing the intuitions of informants about utterances in their native language (such as that two utterances are or are not the same word, or that a certain word is the subject of a sentence), can formulate a coherent grammar, consistent with grammars that would be formulated by other linguists working with other informants on the same language. This holds true even if, as is typically the case for a language with no writing system, the informants are quite unaware of the linguistic units into which utterances in their language can be analyzed. Because the informants' intuitions are apparently valid, they must be based on linguistic representations of some kind.

While linguists are not in total agreement about the nature of the linguistic representation of an utterance, it seems reasonably clear that such a representation must include the syntactic structure, the selection of lexical items and their component morphemes, the phonological structure, and the phonetic structure. The linguist's syntactic diagrams and phonological and phonetic transcriptions are formal reconstructions of different levels of the representation. These levels are not independent of one another. Syntax constrains lexical choice, lexical choice determines morphology and phonology, syntax and phonology determine phonetic structure. The representation thus has extensive inherent redundancy.

The linguistic representation is strictly structural rather than procedural. The listener has no access to the many intermediate steps he must presumably go through in the course of parsing the utterance, so that these steps are not represented. Acoustic details such as formant trajectories are not part of the linguistic representation, simply because the listener does not perceive them as such, but only the phonetic events they reflect. Other aspects of the utterance, such as individual voice quality, speaking rate, and loudness, which the listener can hear, must be presumed to be excluded because they are not linguistic at all and never serve to mark a linguistic difference between two utterances.

---

[4]The proposals in this section are developed in more detail in Mattingly (1991).

Access must be distinguished from awareness. All normal language users, it has been claimed, have access to the contents of linguistic representations. This means that they have a potential ability to introspect and report on significant details of the representation, and to regard it as a structure of phrases, words, and segments, not that they can actually do so. The representation is a complicated affair, and a person who is not "linguistically aware" can no more be expected to notice its characteristic units and structure than an electronically naive person can be expected to appreciate the units and structure of a circuit diagram (Mattingly, 1972). Linguistic awareness must in large part be acquired. The principal stimulus for linguistic awareness in modern cultures is literacy (Morais, Cary, Alegria, & Bertelson, 1979). Unlike illiterate adults or preliterate children, those who have learned to read can readily report on and manipulate at least those units of the linguistic representations of spoken utterances to which units of the orthography correspond (Read, Zhang, Nie, & Ding, 1986). However, there must certainly be other sources of linguistic awareness: Long before writing was known, poets composed verse in meters requiring strict attention to subtle phonological details.

It is not agreed how linguistic representations are created. On one view, they are a byproduct of the cognitive processes by which utterances are analyzed. Linguistic information, recovered step by step from the auditory image of the input signal, is temporarily represented in memory until, at a later stage, the speaker's message can be computed (Baddeley, 1986). The difficulty with this view is that, as has been noted, the language user seems to have no access to the supposedly cognitive analytic steps that must precede the formation of the representation or to the subsequent steps by which the message is derived from this representation. An alternative view is that the representation, as well as the message itself, is not a byproduct but a true output of a specialized, low-level processor (the "language module") whose internal operations, being inaccessible to cognition, have no cognitive byproducts (Fodor, 1983). This view implies that the linguistic representation must have some biological function other than communication, for which the message alone would suffice. What this function might be is unclear (but see Mattingly, 1991, for some speculations).

So far, the cognitive linguistic representation has been considered just as the product of the perception of utterances. But such representations are produced in the course of other modes of linguistic processing as well. Thus, a linguistic representation is formed in the production of an utterance, so that the speaker knows what it is he has just said. And when one rehearses an utterance in order to keep it in mind verbatim, what presumably happens is that the linguistic processor uses a decaying linguistic representation to construct a fresh version of the representation, and incidentally, of the message. This seeming defiance of entropy is possible for linguistic representations (as it may not be for mental representations in general) because of their high inherent redundancy.

Consideration of rehearsal also shows that the linguistic representation can be an input to as well as an output from the linguistic processor. Even more significantly, for the present purposes, a representation not originally produced by primary processes of perception or production can be such an input. An introspective, linguistically aware person can readily compose a "synthetic" linguistic representation according to some arbitrary criterion: the first five words he can think of that begin with /b/, for example. This is obviously a very partial representation: just a sequence of phonological forms drawn from the lexicon, without explicit phonetics or syntax. But if this sequence is rehearsed, the phonetic level, together with whatever syntactic structure or traces of meaning may be accidentally implicit in the sequence, will be computed, just as if the sequence were what remained of a natural representation resulting from an earlier act of production, perception, or rehearsal. All that is required for a synthetic representation to serve as input for computing a natural one is that it contain enough information so that the rest of the structure of the utterance is more or less determined.

These various considerations suggest how it is that one linguistically aware language user can communicate with another, not by means of speech, but by means of synthetic representations, provided a way of transcribing such representations, that is, an orthography, is available. The writer speaks some utterance (at least to himself), creating a linguistic representation. The orthography enables him to transcribe this representation in some very partial fashion. From this transcription, the reader constructs a partial, synthetic linguistic representation. Such a representation is enough to enable the reader's linguistic processor to compute a complete, natural representation, as well as the writer's intended message.

If we compare what happens between writer and reader with what happens between speaker and hearer, it can be seen that the difference is much more than merely a matter of sensory modality. In speech perception, there is a natural and unique set of "signs"—the acoustic events that the human vocal tract can produce—and they are already in a form suitable for immediate linguistic processing (Liberman, this volume). Only the output of this processing is a linguistic representation. The input speech signal is in no sense a partial linguistic representation, but rather a complete representation of a very different kind. Moreover, the specification of the complex relation between the phonetically significant events in the signal and the units of the linguistic representation is acquired precognitively (Liberman & Mattingly, 1991); it does not have to be learned. Indeed, as has been remarked, the hearer has no access to the acoustic events, and may have little or no awareness of the units of the linguistic representation. In reading, on the other hand, there is no one, natural set of input symbols. Linguistic processing must therefore be preceded by a stage having no counterpart in speech perception: a cognitive translation from the orthographic signs to the units of the synthetic linguistic representation. The beginning reader must therefore deliberately master the mapping between the signs and the units, and for this he must have an awareness of the appropriate aspects of the linguistic representation.

Constraints on orthographic form

What psychological factors constrain the form of an orthography? Gelb (1963) makes a useful distinction between "outer form"—the shape of the visible symbols and their arrangement in a text—and "inner form"—the nature of the correspondence of the symbols to linguistic units. Beyond the trivial requirement that the symbols be visually discriminable, there appear to be no particular psychological constraints on outer form. The shapes of the signs in the writing systems of the world and the way they are arranged are extremely various, and such limitations as exist are to be accounted for not by cognitive or linguistic factors but by practical ones, such as the nature of the writing materials available and what patterns are easily written by hand, or by esthetic ones, such as the beauty of particular stroke patterns. This variety is possible because, as has just been seen, a cognitive translation is required for reading and writing in any event. This price having been paid, outer form can vary almost without limit.

Inner form, on the other hand, is highly constrained. In the first place, the orthography must correspond to the linguistic representation, because there is no other cognitive path to linguistic processes. This is the reason that proposals to treat spectrographic displays of speech as, in effect, an orthography the deaf could learn to read (Potter, Kopp, & Kopp, 1966) are not likely to succeed. On the one hand, the reader of spectrograms cannot process the visually-presented spectral information as a listener can process the same information in the auditorially-presented and biologically-privileged speech signal. On the other hand, the spectrogram reader has no natural cognitive access to raw spectral events, and, a fortiori, no awareness of them. Therefore, even if he could somehow synthesize a cognitive spectral representation from the visible one, there is no reason to believe it could be an input to linguistic processes. All he can do is to apply his cognitive knowledge of acoustic phonetics to the task of inferring the linguistic representation from the spectrogram. Because the relation between spectral patterns and even the most concrete level of this representation, the phonetic level, is extremely complex, and a great deal of extraneous information is present, "reading" spectrograms is a slow and unreliable process. Analogous observations, obviously, could be made with respect to other records of physical activity in which linguistic information is implicit, such as the speech waveform or traces of articulatory movements. What has to be transcribed, then, is some level or levels of the linguistic representation itself.

However, certain levels of the linguistic representation are seldom or never transcribed in traditional orthographies. For example, syntactic structure is never transcribed. The few features of orthography that might be considered syntactic, such as punctuation and sentence-initial capitalization, are more reasonably regarded as transcriptions of prosodic elements. Why is syntax thus avoided? It is not just that tree diagrams are cumbersome to draw and nested brackets difficult to keep track of, but that the syntactic structure alone would be insufficient to specify a particular sentence: Each possible phrase marker is shared by an indefinitely large number of sentences. It would therefore be necessary that a syntactic orthography also transcribe in some way the particular lexical choices. But if this is to be done, the phrase-marker itself becomes redundant, because (barring some well-known types of structural ambiguity, such as those discussed by Chomsky, 1957) the words, and the order in which they occur, are themselves sufficient to specify syntactic structure.

Again, someone who supposed that speech and writing converged at the lowest conceivable level, given the difference of modality, might expect that the most efficient form of writing would be a narrow phonetic transcription (see Edfeldt, 1960). This transcription would correspond to the output of the phonological component of the grammar, presumably the level of the linguistic representation closest to the speech signal itself. Owing to contextual variation, higher-level units such as phonemes, syllables, morphemes, or words are not consistently transcribed or explicitly demarcated in such a transcription. But, in contrast to the syntactic orthography just considered, more than enough linguistic information to specify the linguistic representation would nevertheless be implicit. Why is such an orthography not found? A partial answer is that because, as has been suggested, writing and speech are not, in fact, so simply related, there is no particular advantage to a low-level, phonetically veridical representation. Moreover, it seems more difficult to attain awareness of phonetic details insofar as they are predictable. Once the language-learner is able to represent words phonemically, the phonetic level seems to sink below awareness. But as will be seen, there is a still more fundamental reason why a narrow phonetic transcription would be impractical.

It is important to distinguish between the linguistic unit used for the actual processing of an utterance by writer and reader, and the linguistic units to which the various graphemic units correspond. Elementary graphemic units correspond to phonemes (English letters or digraphs), syllables (Japanese kana[5]), or morphemes (simple Chinese characters). These are usually organized into complex units that have been called "frames" (Wang, 1981). A spelled word in English, a complex Chinese character, a grouping of Egyptian hieroglyphics are examples. Frames are usually demarcated by spaces in modern writing, but other demarcative symbols have been used. Sometimes the frame is implicit: The structure of the frame itself may be sufficient to demarcate it from adjacent frames, as in Japanese, where a kanji logogram or logograms is regularly followed by kana syllable signs specifying affixes. Some orthographies, such as those early alphabetic orthographies in which there is no demarcative information of any kind, have no frames larger than their elementary signs. Frames often correspond to linguistic words, but not always: In Chinese and Sumerian, they correspond to morphemes.

By "unit of transcription" is meant the linguistic unit that the writer actually transcribes and the reader cognitively translates to form the synthetic linguistic representation. One might expect that the units of transcription for a particular orthography would be those to which its frames corresponded. Thus, in English, the frames are consistent spellings of words, and the experienced reader's intuition is surely that he reads word by word and not letter by letter, as he would if the transcription unit were the segment. This intuition is borne out by demonstrations of "word superiority." In these experiments, it is found, for example, that subjects can recognize a letter faster and more accurately when it is part of a real written word than when it appears alone or in a nonword (Reicher, 1969). This result suggests that in the case of a real word, subjects can use the orthographic information to recognize the word very rapidly, and then report the letters it contains. If the segment were the transcription unit, the letters corresponding to the segments should be recognized and reported faster than the words.

However, it is possible that the unit of transcription does not really depend on the frame used in a particular orthography, but is in fact *always* the word. One reason for believing this is that the word has to be the most efficient unit of transcription, because words are the largest lexical structures. Anything smaller would require processing more units per utterance; anything larger could not be readily coded orthographically.

Chinese writing allows a test of this possibility. A Chinese word consists of one or more monosyllabic morphemes. In the writing, characters are the frames and correspond to these morphemes. Words as such are not demarcated. There is some evidence, however, that the unit of transcription is nonetheless the word. In a recent experiment (Mattingly& Xu, in preparation), Chinese speakers were shown sequences of two characters on a CRT. In half the sequences, one of the characters was actually a pseudocharacter, consisting of two graphic components that in actual writing occur separately as components of other characters, but not together in the same character. Of the sequences in which both characters were real, half were real bimorphemic words and half were pseudowords. The subject's task was to respond "Yes," if both characters in a sequence were genuine and "No," if either was a pseudocharacter. Subjects performed this task faster for words than for pseudowords, and it was possible to show that this was not simply an effect of the higher transitional probabilities of the word sequences, but rather a valid "word superiority" effect. This result, like that of an earlier experiment by C. M. Cheng (1981, summarized in Hoosain, 1991) suggests that despite morphemic framing and the absence of word

---

[5]Japanese kana correspond, strictly speaking, to moras, which are not equivalent to English syllables. But they do belong to a general class of phonological units that can be called "syllables" (see, e. g., Hyman, 1975).

boundaries, the word is the transcription unit for Chinese readers. Other writing systems in which words are not framed remain to be investigated.

But if word-size frames are not essential for reading word by word, why is a narrow phonetic transcription an unlikely orthography? The reason must be that the shapes of words in such a transcription are context-sensitive and thus difficult to recognize. (Notice what happens to **/hænd/**, *hand*, in **[hæntuwlz]**, *hand tools*, **[hæŋgrənejd]**, *hand grenade*, **[hæmpɪkt]**, *hand picked*, etc.). The reader is therefore forced to process the transcription symbol by symbol, a slow and arduous procedure. In Chinese, on the other hand, though word-boundaries are absent, the form of an orthographic word is constant, or at least not subject to contextual variation. It is suggested that this is a minimal constraint that all writing systems must meet, so that words can serve as units of transcription.

Although words are the transcription units, writing always employs graphemic units corresponding to linguistic units smaller than the word. It might seem possible, in principle, to have a pure logographic system, consisting simply of one monolithic symbol for each word. But the difficulty with such a system is that while the lexicon of a language is, in principle, finite, it is in practice, indefinite: New words are continually being coined or borrowed. In some cases—a nonce word or an unusual foreign name, for example—it would make little sense to provide a special logogram. A writer could thus find himself with no means of writing a particular word because no logogram for it existed. Or, of course, he could be stuck simply because he did not know the correct logogram. An actual writing system insures that the writer will never be in this situation by providing a system of spelling units. The availability of the spelling system guarantees that the orthography will be "productive," that is, that the writer who has mastered the spelling rules will always have some way (though it may not be the "correct" or standard way) to write every word in the language (Mattingly, 1985).

The only linguistic units that have served as the basis for spelling units are syllables and phonemes. It might be thought that morphemes could be the basis of a spelling system and some (e.g., Sampson, 1985) have argued that Chinese has such a system, because the characters correspond to morphemes. This is true, but, as has already been noted, these morphemic units are frames: Relatively few of the characters in the inventory are simple logograms. Over 90% are phonetic compounds, each consisting of two graphic components that (in general) occur also as separate logographic characters. One of these, the "phonetic" stands, in principle, for a particular phonological syllable, and the set of phonetics thus constitutes a syllabary. The other, the "semantic," is one of 214 determiners that serve to mitigate the extensive homophony of Chinese: The number of monosyllabic morphemes far exceeds the number of phonologically distinct syllables. The situation is complicated, however, because there is usually more than one phonetic corresponding to a particular phonological syllable (there are about 4000 in all for about 1300 phonologically distinct syllables), and because, through various accidents of linguistic history, a phonetic often has different phonological values in different characters. But these circumstances should not obscure the highly systematic, syllabographic nature of the spelling, any more than the existence of several spelling patterns for one sound, and numerous inconsistencies in letter-to-sound correspondence, should obscure the systematic, alphabetic nature of English spelling (DeFrancis, 1989).

Words can indeed be analyzed into morphemes as well as segments and syllables, but the inventory of morphemes in a language, like the inventory of words itself, is indefinitely large and subject to continual change. While logograms that are morphemic signs can have a valuable supplementary function in orthography, they could not constitute a productive spelling system, and there is no orthography in which they play this role.

Syllables and segments, on the other hand, have several properties that make them suitable as a basis for spelling units. First, a word can always be analyzed as a sequence of phonological elements of either type. Second, the inventory of syllables may be small (and indeed was small in all the languages for which syllabic spelling developed independently) and the inventory of segments is *always* small. Third, the membership of these inventories changes only very slowly. No other linguistic units have these convenient properties, save perhaps phonological distinctive features (Because a diacritic is used to indicate voicing, it could be maintained that features have a marginal role in Japanese spelling).

In sum, every orthography needs to have a spelling system and a spelling system is necessarily phonographic. It is not accidental that all orthographies spell either syllabically or segmentally: there is probably no other way to spell.

The invention of writing[6]

Writing was invented, probably several times, by illiterates. From what has been said already, it follows that what had to be discovered was one or the other of the two possible spelling principles, the syllabic or the segmental, and that this must have required awareness of these units of the linguistic representation. How could the inventors have arrived at such awareness?

Some linguistic units seem to be more obvious than others. Awareness of words can perhaps be assumed for most speakers, even if they are preliterate or illiterate. It probably requires only a very modest degree of awareness to appreciate that an utterance is analyzable as a sequence of syntactically functional phonological strings, if only because sequences consisting of just one such string are quite frequent: Words may occur in isolation. Certainly preliterate children have no difficulty in understanding a task in which they are to complete a sentence with some word, and a linguist's naive informant readily supplies the names of objects. Awareness of syllables as countable units may also be fairly widespread. The syllable is the basis for verse in many cultures; preliterate children can count the number of syllables in a word. This kind of syllabic awareness, however, is probably not the same thing as being aware (if such is indeed the case) that the syllables of one's language constitute a small inventory of readily demarcatable units.

These limited degrees of linguistic awareness are probably readily available to speakers of all languages. But more subtle forms of awareness may well have arisen only because they were facilitated by specific properties of certain languages, including, in particular, those for which writing was originally invented.

Consider, first, Chinese. In the Ancient Chinese language, words were in general monomorphemic, there being neither compounding nor affixation. Morphemes were monosyllabic and a particular morpheme was invariant in phonological form. Because of restrictions on syllable structure, the inventory of syllables was small. Homophony was therefore very extensive, one syllable corresponding to many morphemes (Chao, 1968).[7]

[6]An earlier formulation of some of the proposals in this section can be found in Mattingly (1987).

[7]DeFrancis (1950), protesting against the "monosyllabic myth," has suggested that there actually were many polysyllabic words in Ancient Chinese, just as in Modern Chinese, but that only one of the syllables in a word was transcribed in the writing. Thus, morphemes that appear from the writing to be monosyllabic homophones may actually have been polysyllabic morphemes with common homophonous syllables. Y.-R. Chao's (1968) response was that "so far as Classical Chinese and its writing system is concerned, the monosyllabic myth is one of the truest myths in Chinese mythology" (p. 103). For the present purpose,

The number of different characters in the Chinese writing system sharing a particular phonetic component gives some notion of the degree of homophony in Ancient Chinese, and this number often exceeds twenty. Chinese thus contrasts sharply with English and other Indo-European languages, in which morphemes vary in phonological form, may be polysyllabic, and may not even consist of an integral number of syllables; syllable structure is complex; the number of possible syllables is relatively large; and homophony is therefore a marginal phenomenon.

Since words coincided with morphemes in Chinese, awareness of morphemes required no analysis, and the use of logograms, i.e., morphemic signs, was an obvious move. The extensive homophony made "phonetic borrowing"—using the sign for one morpheme to write another morpheme with the same syllabic form[8]—a strategy that was both obvious and productive; when a writer needed to write a morpheme, a sign with the required sound was very likely to be available. It thus became obvious that the number of different sounds was in fact small, yet every morpheme corresponded to one of them. Awareness of demarcatable syllable units thus developed. Of course, the same extensive homophony that fostered the discovery of these units also meant that their signs had to be disambiguated by the use of logograms as determiners, as in the large class of characters called "phonetic compounds," described earlier.

Chinese morphophonological structure thus encouraged the discovery of the syllable; on the other hand, it did not encourage the discovery of the phonemic segment. There was nothing about this structure that would have served to isolate phonemes from syllables or morphemes.

Sumerian was an agglutinative language. A word consisted of one or two monosyllabic CVC morphemes and various inflectional and derivational affixes. Its phonology had certain properties that imply a preference for a CVCVC...VC syllabification. There were no intrasyllabic consonant clusters; a cluster simplification process deleted the first of two successive consonants across syllable boundaries, resulting in such alternations as **til**, **ti**, 'life'; and final vowels were deleted (Driver, 1976; Kramer, 1963). In other relevant respects, however, Sumerian resembled Chinese and, like Chinese, favored awareness of morphemes and of syllables as demarcatable units. Aside from the effects of the syllable-forming processes just mentioned, a root maintained an invariant phonological form. A root could be repeated to indicate plurality. Because the morphemes were monosyllabic, and because of the restricted syllable structure, the number of possible distinct syllables was small. These circumstances, resulted, again, in extensive homophony.

For a speaker of Sumerian to become aware of morphemes was perhaps not quite as easy as for a speaker of Chinese. He would have had to notice that words with similar meanings often had common components, for the most part corresponding to syllables. This stage of awareness having been achieved, morphemic writing is possible. From this point on, the story is quite similar to that for Chinese, homophony leading to phonetic borrowing, and then to syllable writing supplemented with determiners.

There is, however, one striking difference between the Sumerian and the Chinese writing systems. While Chinese makes no internal analysis of syllables, Sumerian does. A sign for a $C_1V_1C_2$ morpheme could be borrowed to write a $C_1V_1C_3$ morpheme, e.g., the RIM sign was used to write **rin**. A VC syllable sign could be used as a partial phonetic indicator after

---

however, it does not matter whether the myth is true or false. DeFrancis's partial homophony will serve as well as the total homophony more usually attributed to Ancient Chinese.

[8]Or, on DeFrancis' (1950) view, another morpheme having a syllable in common.

a logogram, e.g., GUL + UL. For many of the $C_1V_1C_2$ syllables, as has been mentioned, there was no special sign; instead, such a syllable was written with the sign for the C_1V_1 followed by the sign for V_1C_2. Thus the syllable ral is written RA AL (examples from Gelb, 1963). A possible explanation of these various practices is that in spoken Sumerian, consistent with its preference for CVCVC...VC structure, some form of vowel coalescence took place when two similar vowels came together, so that $C_1V_1 + V_1C_2$ sequences became phonetically $C_1V_1C_2$, and thus homophonous with original $C_1V_1C_2$ syllables. Such homophony could have suggested analyzing and so writing the latter as $C_1V_1 + V_1C_2$. Again CV signs as well as VC signs were used to indicate the endings of $C_1V_1C_2$ morphemes. For example, because of multiple semantic borrowing, the logogram DU could stand not only for **du**, 'leg,' but also for **gin**, 'go,' **gub**, 'stand,' and **tum**, 'bring'. Which of the latter three was intended was indicated by writing DU NA for gin, DU BA for gub, and DU MA for tum (Driver, 1976). This practice perhaps arose because the phonological final vowel deletion made $C_1V_1C_2$ and $C_1V_1C_2V_2$ sequences homophonous, suggesting that what followed C_1V_1 could be written in either case as if it were C_1V_2. Thus the Sumerians may have viewed $C_1V_1C_2$ morphemes either as $C_1V_1 + V_1C_2$ or as $C_1V_1 + C_2V_2$, either of which was entirely consistent with their syllabic phonological awareness.

With Egyptian, in contrast to Chinese and Sumerian, the morphology and phonology of the language of the language favored segmental awareness. In Afro-Asiatic languages, the roots are biconsonantal and triconsonantal patterns into which different vowels or zero (that is no vowel at all) are inserted to generate a large number of inflected forms. Because the vowels of Egyptian are unknown, it is easier to illustrate this point with an example from another Afro-Asiatic language, e.g., Hebrew. From the Hebrew root k-t-b are derived **kātab**, 'he wrote'; **yikkātēb**, 'he will be inscribed'; **kātob** 'to write'; **kātub**, 'written'; **miktab**, 'letter; and many other forms. Because of phonological restrictions, the number of different consonantal patterns in Egyptian was relatively small, and there were consequently numerous homophonous roots, e.g., **n-f-r**, 'good'; **n-f-r**, 'lute' (Jensen, 1970).

It is not difficult to imagine an Egyptian noticing that many sets of semantically similar words in his language had a common consonantal ground and a varying vocalic figure, though at first he may not have individuated the consonants. Accordingly, signs for root morphemes were devised. The homophony of Egyptian then did for phonetic segments what homophony in Chinese and Sumerian did for syllables. A morphemic sign was frequently borrowed to write a homophonous morpheme, e.g., NFR, the sign for **n-f-r**, 'lute', used to write **n-f-r**, 'good,' or WR, 'swallow,' used to write **w-r**, 'big.' The signs were now generalized to stand for consonantal sequences that were not morphemes, e.g., WR < WR was used to write the first part of **w-r-d**, 'weary.' And because in some cases roots were actually uniconsonantal, and in other cases the second consonant had become silent, some signs came to stand for single consonants, and constituted a consonantal alphabet. Thus the **d** in **w-r-d** could written with the sign D < DT, the final consonant in **d-t**, 'hand,' being actually the feminine suffix, not part of the root. Finally, logograms were employed as determiners to clarify ambiguous transcriptions: the spelling MN N H for the word **m-n-h** being followed by the determiner for 'plants' when this word had the sense 'papyrus plant,' the determiner for 'men' when it had the sense 'youth,' and the determiner for 'minerals' when it had the sense 'wax' (examples from Jensen, 1970). In this fashion, the Egyptians arrived at a consonantal spelling system.

If the Egyptians had thus achieved segmental awareness, why did they not transcribe the vowels as well as the consonants? It is not likely that they were unable to hear the different vowels. The explanation is rather that because the vowels ordinarily conveyed only inflectional information, the writing was sufficiently unambiguous without such indications, just as English writing is sufficiently unambiguous without stress marking. But as has already been noted, there was a convention for writing vowels when necessary. Such writing is found very early in the history of Egyptian writing (Gelb, 1963).

The Egyptians could hardly have arrived at a syllabic system instead. Because zero alternated with vowels in the generation of words, there was no obvious correspondence between morphemes and syllables or syllable sequences. And because of such alternations, a syllabic orthography would have resulted in a number of dissimilar spellings for the same morpheme.

These examples suggest that the phonological awareness required for the invention of writing develops when morphemes have a highly restricted phonological structure—monosyllabic, in the case of Sumerian and Chinese; consonantal in the case of Egyptian—that results in pervasive homophony. Speakers of such languages are naturally guided to the invention of writing by these special conditions. (A corollary is that it is not necessary to propose a derivation of Egyptian from Sumerian to account for parallels in the development of the two systems.) On the other hand, Indo-European languages and many others lack any such restrictions, and would not have favored phonological awareness in this way. Indeed, one has to wonder whether, for such languages, writing could have been invented at all.

In the early discussion of the psychology of reading, the precise role of phonological awareness in learning to read appeared equivocal. Is phonological awareness a prerequisite for reading? Or, on the other hand, does the experience of reading engender phonological awareness (Liberman, Shankweiler, Liberman, Fowler, & Fischer, 1977)? It was later seen, however, that both statements must be true: The beginning reader must, indeed, have some degree of awareness, but this awareness is increased and diversified in appropriate directions as a result of his encounter with the orthography (Morais, Alegria & Content, 1987). In the same way, the invention of writing must have been an incremental process, beginning with an initial awareness of morphemic structure. The experience of working out ways to transcribe morphemes for which there were no logograms led to awareness of the syllabic or phonemic structure of these morphemes, and then to awareness of such structure generally.

To say that the process was incremental is not to say that it was not quite rapid. It is noteworthy that in all three of the writing traditions just considered, evidence of spelling is found very early: in Sumerian writing from the Uruk IV stratum (Gelb, 1963); in Chinese writing of the Shang dynasty (DeFrancis, 1989); in Egyptian writing of the First Dynasty (Gelb, 1963). These facts are consistent with the proposal that for general-purpose writing, a purely logographic system is impractical. As has been argued, an orthography is not productive without a spelling system: The invention of the one requires the invention of the other.

To the extent that this account of the invention of writing is plausible, it supports the dichotomy between syllabic and segmental spelling proposed earlier, for what had to be invented was one or the other of the two spelling principles that provide the basis for the classification. It should also be noted that the segmental principle did not develop in Egypt by elaborating on the syllabic principle, but rather by generalizing from the segmental transcription of morphemes: The syllable played no role. And, conversely, when Sumerians analyzed complex syllables, they did not resolve them into their constituent phonemes, but rather into simpler syllables. The discovery of one method almost seems to have guaranteed that the other would not be discovered. In effect, speakers of these languages come to regard them as as essentially syllabic or as essentially segmental, and their writing systems reflect one of these two phonological theories.

Transmission of writing systems

It has already been noted that orthographic traditions are either consistently syllabic or consistently segmental. Some explanation for this consistency is required. It seems natural enough, perhaps, that a segmental tradition should not become syllabic, for this would appear to be a backward step. But that no syllabic tradition should have become segmental is puzzling, the more so because there have been at least two occasions when such a development might reasonably have been expected. The first was when speakers of Akkadian, an Afro-Asiatic language with consonantal root structure similar to that of Egyptian and Hebrew, borrowed Sumerian syllabic writing. A proper awareness of the morphophonology of their language would have suggested that they convert the Sumerian system into a consonantal system. But instead, the Akkadians preserved the syllabic character of the borrowed writing, even though to write the same triconsonantal pattern in different ways depending on the particular inflectional vowels obscured the roots of native words. Similarly, the Mycenaean Greeks borrowed Minoan syllable writing, and instead of making an alphabet out of it, as would have been sensible, given the extensive consonant clustering in Greek, they continued to write with signs that stood for CV syllables, either ignoring the "extra" consonants or pretending that they were syllables. This resulted in such bizarre transcriptions such as A RE KU TU RU WO for **alektruōn**, 'cock' (Ventris & Chadwick, 1973). What can have happened to linguistic awareness in these cases?

The explanation begins with the observation that the mismatches between language and writing observed for Akkadian and Mycenean Greek are not unparalleled; they are simply fairly extreme cases. While an originally invented writing system clearly reflects the morphophonological structure of the language it was invented to write, this situation is obviously exceptional. In general, the system used at a particular time to write a particular language has been inherited from an earlier stage in the history of that language, or has been adapted from a system (itself perhaps an adaptation) used for some other language, or, most commonly, both. The consequence, in many cases, is that the writing often seems very poorly suited to the spoken language. If Akkadian and Mycenaean Greek illustrate the risks of borrowing, the English writing system is a good illustration of the effects of orthographic inheritance. The phonology of English has changed considerably since the fifteenth century, most notably in consequence of the Great Vowel Shift, but the writing system has remained very much as it was then (Pyles, 1971). As a consequence, the system has a number of features that must seem very peculiar to the foreigner learning English: For example, the same letter is used to write phonetically dissimilar vowels, a tense vowel is denoted by an E after the following consonant, and a lax vowel is denoted by the doubling of this consonant. A similar account could be given for Chinese writing, which corresponds more closely to Classical Chinese than to any modern dialect.

It cannot be doubted, given what has been learned in recent years about the relation between orthographic structure and learning to read in modern languages, that such complications place a heavy burden on the learner (Liberman, Liberman, Mattingly, & Shankweiler (1980). What is surprising, given the close connection between literacy and awareness of linguistic representations, a connection clearly essential in the invention of writing, is that readers and writers have so often happily accepted (once they have learned it) an orthography that seems poorly matched to their language. It might have been expected that Akkadian cuneiform would have been rejected as soon as it was proposed, and that English orthography would by now have been abandoned as obsolete. But, instead, it is reported that the Akkadians believed their writing system to be of divine origin (Driver, 1976), and Chomsky and Halle (1968) say that "conventional [English] orthography is...a near optimal system for the lexical representation of English words" (p. 49).

In the case of inherited orthographies, the explanation may be that the orthography itself may determine not only which aspects of linguistic representations are singled out for awareness, but perhaps, indirectly, the character of these representations themselves. This could come about if the orthographically based, synthetic input representations were taken seriously by the language processor as evidence about the structure of the language, and thus led to adjustments in the beginning reader's morphophonology. It will be recalled that according to the sketch of the reading and writing process given earlier, the processor does not distinguish synthetic representations from natural ones. Consistent with this possibility is the fact that orthographic conventions sometimes mimic phonology: The conventions for marking English tense and lax vowels invite the reader to assume that underlying lax vowels become tense in open syllables and underlying tense vowels become lax before underlying geminate consonants. Such pseudophonological rules, as well as derivational morphological relations as those between *heal, health* or *telegraph, telegraphy,* though at first having merely orthographic status, may acquire linguistic reality for the experienced reader.[9] For such a reader, the orthography corresponds to linguistic representations because the representations themselves have been appropriately modified, and English orthography now indeed seems "near optimal."

In the case of borrowed orthographies, a similar explanation may apply. The phonological awareness of a borrowing group, such as the Akkadians or the Greeks, was not guided by peculiarities of their own spoken language, as was the awareness of the original inventors of writing, but by the writing system they were borrowing. This is hardly surprising: The borrowers were not sophisticated consumers, comparing competing technologies to decide which was better for their particular needs. They did not realize that there was a choice that could be made between the two different spelling principles and the theories of phonology implicit in each. They simply embraced unquestioningly the spelling principle—syllabic in the cases considered above—used by the culture under whose influence they had come, just as beginning readers accept the principle of the writing system they inherit. This principle having been accepted, the morphophonologies of the borrowers adjusted so that their linguistic representations became, in fact, a good match to their syllabic orthographies.

If this account is correct, it has to apply to the transmission of segmental systems, as well. A segmental system has obvious advantages over a syllabary for languages with complex syllable structure. But the spread of the alphabet is perhaps to be explained by an appeal to the forces of tradition rather than to those of reason.

An orthographic tradition can perpetuate itself because it offers a particular brand of morphophonological awareness ready-made. The processes of introspection needed to invent writing in the first place are not demanded. The kind of awareness offered may be poorly matched to a particular language, but this does not impede the process. Whether the

---

[9]These changes in the morphophonologies of individual readers have, by hypothesis, no basis in the spoken language and are transmitted only from writer to reader, and not from mother to child. Thus, though psychologically real, they are not part of the grammar of the language as usually conceived of.

writing system is borrowed or inherited, the morphophonology of the new reader adjusts to meet the presuppositions of the system.

Conclusions

It has for some time been widely agreed that the notion of linguistic awareness is essential for an understanding of the reading process, the acquisition of reading and reading disability. This notion is likewise essential for an understanding of the invention and dissemination of orthographies. There are really only two possible ways to write, the syllabic method and the segmental method, because only by using one of these two methods is the writer assured of being able to write any word in his language. But for an illiterate to discover either of these methods, and thus be in a position to invent writing, requires awareness of the appropriate unit of linguistic representations. Awareness of syllables, or, on the other hand, of segments, is fostered by special morphophonological properties found in those languages for which writing systems were invented, though by no means in all languages. But once it has become established, the writing system itself shapes the linguistic awareness, and even the phonology, both of those who inherit the system and of those who borrow it to transcribe some other language. Thus, in the history of writing, syllabic and segmental traditions are clearly distinguished.

Acknowledgment
Preparation of this paper was supported in part by NICHD grant HD-01994 to Haskins Laboratories. Alice Faber, Leonard Katz, Alvin Liberman, and Yi Xu gave helpful comment and criticism on an earlier draft.

References

Baddeley, A. D. (1986). *Working memory.* Oxford: Clarendon Press.
Chao, Y.-R. (1968). *Language and symbolic systems.* London: Cambridge University Press.
Cheng, C. M. (1981). Perception of Chinese characters. *Acta Psychologica Taiwanica, 23,* 137-153.
Chomsky, N. (1957). *Syntactic structures.* The Hague: Mouton.
Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York: Harper & Row.
Cross, F. M., & Freedman, D. N. (1952). *Early Hebrew orthography: A study of the epigraphic evidence.* New Haven: American Oriental Society.
DeFrancis, J. (1950). *Nationalism and language reform in China.* Princeton, NJ: Princeton University Press.
DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems.* Honolulu: University of Hawaii Press.
Driver, G. R. (1976). *Semitic writing: From pictograph to alphabet.* London: Oxford University Press.
Edfeldt, A. W. (1960). *Silent speech and silent reading.* Chicago: University of Chicago Press.
Edgerton, W. F. (1952). On the theory of writing. *Journal of Near Eastern Studies, 11,* 287-290.
Fodor, J. A. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.
Gelb, I. J. (1952). *A study of writing: The foundations of grammatology.* Chicago: University of Chicago Press.
Gelb, I. J. (1963). *A study of writing* (rev. ed.). Chicago: University of Chicago Press.

Hill, A. A. (1967). The typology of writing systems. In W. M. Austin (Ed.), *Papers in linguistics in honor of Leon Dostert* (pp. 92-99). The Hague: Mouton.
Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese*. Hillsdale, NJ: Lawrence Erlbaum.
Hyman, L. (1975). *Phonology theory and analysis.* New York: Holt-Rinehart-Winston.
Jensen, H. (1970). *Sign, symbol and script: An account of man's efforts to write* (3rd ed.). Tr. G. Unwin. London: G. Allen & Unwin.
Kelley, D. H. (1976). *Deciphering the Mayan script.* Austin. TX: University of Texas Press.
Kramer, S. N. (1963). *The Sumerians: Their history, culture, and character.* Chicago: University of Chicago Press.
Liberman, A. (1992). The relation of speech to reading and writing.
Liberman, A. M., & Mattingly, I. G. (1991). Modularity and the effects of experience. In R. R. Hoffman & D. S. Palermo (Eds.), *Cognition and the symbolic processes: applied and ecological perspectives* (pp. 33-38). Hillsdale, NJ: Lawrence Erlbaum.
Liberman, I. Y., Liberman, A. M., Mattingly, I. G., & Shankweiler, D. (1980). Orthography and the beginning reader. In J. F. Kavanagh & R. Venezky (Eds.), *Orthography, reading, and dyslexia* (pp. 137-153). Baltimore: University Park Press.
Liberman, I. Y., Shankweiler, D., Liberman, A. M., Fowler, C., & Fischer, W. F. (1977). Phonetic segmentation and recoding in the beginning reader. In A. S. Reber & D. A. Scarborough (Eds.), *Towards a psychology of reading* (pp. 207-225). Hillsdale, NJ: Lawrence Erlbaum.
Mattingly, I. G. (1972). Reading, the linguistic process, and linguistic awareness. In J. F. Kavanagh & I. G. Mattingly, *Language by ear and by eye: The relationships between speech and reading* (pp. 133-147). Cambridge, MA: MIT Press.
Mattingly, I. G. (1985). Did orthographies evolve? *RASE remedial and special education, 6(6),* 18-23.
Mattingly, I. G. (1987). Morphological structure and segmental awareness. *CPC cahiers de psychologie cognitive, 7,* 488-493.
Mattingly, I. G. (1991). Reading and the biological function of linguistic representations. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the Motor Theory of Speech Perception* (pp. 339-346). Hillsdale, NJ: Lawrence Erlbaum.
Mattingly, I. G., & Xu, Yi (in preparation). *Word superiority in Chinese.*
Morais, J., Alegria, J., & Content, A. (1987). The relationships between segmental analysis and alphabetic literacy: An interactive view. *CPC Cahiers de Psychologie Cognitive, 7,* 415-438.
Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7,* 323-331.
Potter, R. K., Kopp, G. A., & Kopp, H. G. (1966). *Visible speech* (2nd ed.). New York: Dover.
Pyles, T. (1971). *The origins and development of the English language* (2nd ed.). New York: Harcourt Brace Jovanovich.
Ray, J. D. (1986). The emergence of writing in Egypt. *World Archaeology, 17(3),* 308-316.
Read, C., Zhang, Y.-F., Nie, H.-Y., & Ding, B.-Q. (1986). The ability to manipulate speech sounds depends on knowing alphabetic reading. *Cognition, 24,* 31-44.
Reicher, G. M. (1969). Perceptual recognition as function of the meaningfulness of the stimulus material. *Journal of Experimental Psychology, 81,* 275-280.
Sampson, G. (1985). *Writing systems A linguistic introduction.* Stanford, CA: Stanford University Press.
Ventris, M., & Chadwick, J. (1973). *Documents in Mycenaean Greek* (2nd ed.) Cambridge: Cambridge University Press.
Wang, W. S.-Y. (1981). Language structure and optimal orthography. In O. J. L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology* (pp. 223-236). Hillsdale, NJ: Lawrence Erlbaum.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 2

Reading Consonants and Guessing Vowels: Visual Word Recognition in Hebrew Orthography

Ram Frost
Shlomo Bentin
Department of Psychology
The Hebrew University, Jerusalem

For many years studies in the English language have dominated experimental research in visual word recognition. This state of affairs cannot be accounted for by considering merely geographic reasons. Rather, it was partly due to an underlying belief that English was sufficient because reading processes (as well as other cognitive processes) are universal. In recent years, however, studies in orthographies other than English have become more and more prevalent. These studies have the common view that reading processes cannot be explained without considering the reader's linguistic environment. Moreover, it is assumed that reading strategies in one orthography can be understood better when other orthographies provide additional points of reference. It is in this context that recent research in reading Hebrew should be evaluated. In the present chapter we describe the specific characteristics of Hebrew orthography and discuss their origin with regard to the complex morphology of the Hebrew language. We further examine their possible effects on the reading strategies adopted by beginning and skilled readers. Finally, we discuss the processing of morphologic information conveyed by Hebrew print, a particularly interesting contrast to other writing systems that have been studied.

Characteristics of the Hebrew orthography

The orthography of the Hebrew language should be described in reference to its very complex productive morphology (see Katz & Frost, this volume). In Hebrew, as in other Semitic languages, all verbs and the vast majority of nouns and adjectives are comprised of roots which are usually formed of three (sometimes four) consonants. The three-consonant roots are embedded in pre-existing morphophonological word patterns to form specific words. Phonological patterns can be either a sequence of vowels or a sequence consisting of both vowels and consonants. Thus, in general, Hebrew words can be decomposed into two abstract morphemes, the root, and the phonological pattern. Roots and phonological patterns are abstract structures and only their joint combination (after the application of phonological and phonetic rules) forms specific words. Although these morphemes carry some semantic and morpho-syntactic information, their meaning is often obscure and changes for each root-pattern combination (see Berman, 1980). This is because there are no unequivocal rules for combining roots and phonological patterns to produce specific word meanings. For example, the word KATAVA ("a newspaper article") is composed of the root KTV, and the phonological pattern -A-A-A (the lines indicate the position of the root consonants). The root KTV alludes to anything related to the concept of writing, whereas the phonological pattern A-A-A is often (but not always) used to form nouns that are usually the product of the action specified by the root. It is the combination of both root and word pattern that forms the word meaning "article". Other phonological word patterns may combine with the same root to form different words with different meanings that can be closely or very remotely related to writing. For example, the word KATAV ("press correspondent") is formed by combining the root KTV with the phonologic pattern -A-A-. The phonological pattern -A-A- carries the morpho-syntactic information that the word is a noun which signifies a profession. But this same phonological pattern is also common in adjectives that signify attributes. Unlike KATAV, the word KTOVET ("address") is formed by combining the same root with a phonological pattern that includes consonants as well as vowels. This pattern carries the morpho-syntactic information of that the word is a feminine noun. Note that the same phonologic pattern can be applied to other roots resulting in various different verbs or nouns, each of which is related to its respective root action. Therefore only the combination of both root and phonological pattern specifies the exact meaning of a word.

Although words in Hebrew are composed of two morphemes, the root and the phonologic pattern, the semantic information conveyed by each morpheme is not equally constraining; the semantic information specified by the root is by far more restricted and more specific than that specified by the phonologic word pattern, and it conveys the core meaning of the word. The word pattern, on the other hand, in many cases carries nothing more than word class information. Therefore, one might assume that the understanding of spoken language is based primarily on the identification of the root. Although speculative, it can be reasonably suggested that this morphologic decomposition characteristic of the Semitic languages had directly influenced the development of the Semitic writing systems.

Because of the productive characteristic of Hebrew morphology, Hebrew orthography was designed to convey to the reader primarily the root information (see Katz & Frost, this volume). Hence, the letters in Hebrew represent mainly consonants. The vowels are depicted by diacritical marks (points and dashes) presented beneath (sometimes above) the letters. Although the diacritical marks carry mainly vowel information, they also differentiate in some instances between fricative and stop variants of consonants. In modern Hebrew, we have lost most of the phonetic differentiation between fricative and stop pronunciations, but it is still kept for 3 consonants, in which the letter indicates two different phonetic realizations of these phonemes: /b/→[b] or [v], /p/→[p] or [f], and /k/→[k] or [x]. In these cases a point is inserted inside the letter to indicate the stop pronunciation. Thus the presentation of vowels reduces considerably several aspects of phonemic ambiguity. The diacritical marks, however, are omitted from most reading material, and can be found only in poetry, children's literature, and religious scripts. Although some of the vowels can also be conveyed by letters, these letters are not regularly used, and are considered optional. Thus the most salient characteristic of the Hebrew orthography is that it presents the reader with only partial phonological information. However, incomplete phonologic information is only one specificity of the Hebrew orthography. Because the same root may be combined with different word patterns, frequently the vowel-sequence is the only difference between several words. Therefore, when the vowel marks are omitted, the same string of letters sometimes denotes up to seven or eight different words. Consequently the Hebrew reader is normally exposed to phonological as well as semantic ambiguity. An illustration of Hebrew unpointed and pointed print is presented in Figure 1.

The Figure describes the possible reading of one consonant cluster. The unpointed letter string "דבר" has five meaningful possible readings. The letter "ב" can be read either as [v] or [b] which are distinguished by a dot that appears within the letter, but only in pointed print. The triconsonantal root "דבר" can, thus, be read as /dvr/ or /dbr/ and forms 3 clusters of words: Three words inflected from the root /dbr/ which signifies the action of speaking, and two words /davar/ and /dever/ which share the same consonants, but originated historically from different languages, and therefore do not share any semantic features (the former meaning "a thing" while the latter means "pestilence"). An example of a phonologic pattern which is conveyed by letters in addition to diacritical marks underneath the consonants can be seen in the word /dover/. Note that in the present tense of the root /dbr/ the pronunciation of the middle phoneme /b/ changes into a /v/. These interchanges between fricative and stop pronunciations of consonants are very common in Hebrew. For /dover/ the phoneme /o/ is conveyed by the letter ו. In its unpointed form, this letter can represent the phoneme /v/ as well.

The root דבר
/DVR/

דָּבָר	דַּבֵּר	דִּבֵּר	דֻּבַּר	דֶּבֶר
/davar/	/daber/	/diber/	/dubar/	/dever/
(thing)	*(speak!)*	*(he spoke)*	*(was spoken)*	*(pestilence)*

דוֹבֵר

/dover/
(speaker, he speaks)

Figure 1. **Phonologic ambiguity in unpointed Hebrew print.**

The introduction of vowels marks in printed Hebrew

For the above reasons, Hebrew orthography was designed to provide the reader with the abstract root information, regardless of the possible words that the letter string might represent. Thus, the unpointed orthography served the purpose of denoting in print the optimal amount of phonologic information. The gain in omitting the vowels from the print was multiple: First, the set of letters in the alphabet was smaller—Hebrew has only 22 letters, and the written words were shorter. Second, the presentation of consonants alone made the abstract root more salient. Indeed, the original Hebrew writing system was unpointed. It remained unpointed as long as Hebrew was a living language, that is until the second century.

It was only between the second and the tenth century that the vowel marks were introduced into Hebrew orthography (see Morag, 1972). Since after the second century most of the Jewish nation was dispersed in Europe, Asia, and Africa, they no longer spoke Hebrew as their native language. For the fear that the correct pronunciation of the Hebrew words in the holy scriptures might be forgotten, the vowel marks were introduced. The point of interest in this historical analysis is that the vowel marks were not necessary when Hebrew was a living spoken language. Their function of denoting the specific pronunciation of words became a necessity only when Hebrew ceased to be a naturally spoken language. It is worth noting that the vowel marks were used only for writing holy scriptures or poetry. This is because it is only for poetry and religious scripts that the exact phonemic notation is indeed crucial. Nevertheless, as will become evident in the next section, the importance of vowel marks for both beginning and skilled readers is incontestable.

The use of vowel marks by the beginning reader

Vowel marks aid phonologic recoding

Aside from poetry and religious texts, most children's literature in Hebrew is pointed. Traditionally, most schools in Israel have adopted methods of teaching reading which involve the use of vowel marks at the initial stages of reading acquisition. The purpose of this method is two-fold. First, the vowels convey the unequivocal phonemic structure of the printed word to the beginning reader. It is well established today that beginning readers recognize and name printed words through a process of phonological mediation (e.g., Calfee, Chapman, & Venezky, 1972; Conrad, 1972; Shankweiler & Liberman, 1976). Moreover, decoding skills were shown to be a developmental prerequisite for efficient reading for meaning (e.g., Perfetti & Hogaboam, 1975). Phonemic recoding based on grapheme-to-phoneme conversion rules is very simple in pointed Hebrew. In fact, in its pointed form, Hebrew orthography is almost as shallow as the Serbo-Croatian orthography (Katz & Frost, this volume) and allows a simple use of prelexical phonologic processing. Without the vowel marks the beginning reader in Hebrew would have to rely on the holistic identification of consonant clusters and their correspondence to spoken words, which as mentioned above is extremely ambiguous.

Vowel marks affect phonologic awareness

A second gain in teaching children to read with vowel marks is their beneficial effect on the development of phonological awareness. Phonological awareness is the ability to consciously recognize the internal phonemic structure of spoken words (Bentin, this volume). Several authors reported that the ability to manipulate phonemic segments consciously, develops only around the first grade in elementary school (e.g., Liberman, Shankweiler, Fisher, & Carter, 1974), and has been positively correlated with reading ability (e.g., Bertelson, 1986; Bradley & Briant, 1983; 1985; Liberman & Shankweiler, 1985). This correlation was used to develop methods for predicting in kindergarten, how efficiently would the children acquire the reading skills in school (Lundberg, Olofsson, & Wall, 1980; Mann, 1984). Recently, the importance of phonological awareness for reading was demonstrated also in Hebrew (Bentin & Leshem, in press). In that study, the authors found that children who scored low on a phonemic awareness battery administered in kindergarten scored also low on a reading test in school. However, if those children were trained in kindergarten and improved their segmentation skills, they reached the school standards and read as well as children who had scored highly on the initial tests of phonological awareness.

The relationship between phonological awareness and reading is not, however, unidirectional. Several studies have suggested that, in the absence of reading instruction, the ability to isolate and manipulate single phonemes in coarticulated syllables is obstructed (e.g., Bertelson & de Gelder, 1990). Apparently, by being exposed to the alphabetic principle, children become aware that letters are usually mapped into single phonemes rather than into coarticulated phonological units. The emergence of this revelation should be facilitated when the relationship between letters and phonemes is simple and isomorphic (as in a shallow orthography) than when it is complex or partial (as in a deep orthography). The addition of the vowel marks to the consonants changes the Hebrew orthography from being deep to being almost as shallow as Serbo-Croatian or Italian. Therefore, by using the pointed print, teachers help triggering phonemic awareness that is essential for efficient reading acquisition.

The processing of consonants and vowel marks by the skilled reader

A question of great interest in the study of word recognition in Hebrew is how vowel marks are processed by the skilled reader. From the beginning of the third grade children are gradually exposed to unpointed print and by the sixth grade they encounter unpointed print almost exclusively. What is, then, the possible purpose of vowel marks for the skilled reader? How are they processed in print? This question is of special interest because it is often assumed that mature readers rely on fast visual-orthographic cues rather than on phonologic recoding in word recognition (see McCusker, Hillinger, & Bias, 1981, for a review).

Skilled readers cannot disregard vowel information in print

Navon and Shimron (1981) were the first to examine the use of vowel marks by skilled readers in Hebrew print. Interested to see whether readers can disregard the vowel marks while making lexical decisions, they presented undergraduate subjects with pointed letter strings, and instructed them to ignore the vowel marks while making word/nonword discriminations. Their results showed that positive decisions were slowed when the consonants formed a legal word while the marks underneath the letters suggested an incorrect vowel configuration. Consequently, Navon and Shimron (1981) concluded that although the Hebrew skilled reader does not need the vowel marks for fast lexical decisions he or she cannot ignore them even when instructed to do so (see also Navon & Shimron, 1984).

One point of interest in Navon & Shimron's study relates to the recognition of the correct vowel marks in Hebrew. Although modern Hebrew differentiates only between five vowels (/a/, /e/, /i/, /o/, /u/) it has more than five vowel marks. When the vowel marks were introduced into Hebrew between the second and the tenth century, the vocalization system that became the most influential originated from the Tiberias region. This system had two notations for /a/ (ָ,ַ) and two notations for /e/ (ֶ, ֵ). These notations probably reflected a Hebrew dialect that was spoken in the northern part of the country and had seven rather than five vowels. Although this dialect had become extinct, the printed notations for these vowels are still used in modern Hebrew and used consistently according to orthographic rules (Morag, 1972). Navon & Shimron's results demonstrated that the Hebrew reader is not sensitive to interchanges in the printed forms of the two vowel marks representing /a/ or the two vowel marks representing /e/, as long as the correct phonemic structure of the word is maintained. This is of special interest because similar ambiguity exists with current Hebrew consonants. Hebrew has two letters representing each of the phonemes /t/, /k/, and /kh/. Similar to the vowels, these letters also reflect a historical distinction between phonemes, a distinction without phonetic reality in modern Hebrew. Nevertheless, in contrast to the insensitivity of the reader to the alternative forms of the vowel marks, the skilled reader makes very few errors in lexical decision when the letters representing these consonants are interchanged. This probably reflects the relative importance given by the skilled reader to the consonants as opposed to the vowel marks.

The inability of Hebrew readers to disregard vowel information was further examined in a study that employed the repetition priming paradigm (Bentin, 1989). In this study subjects were required to make lexical decisions to words and nonwords that were either pointed or unpointed. Orthographic, phonemic, and identity repetitions were examined at lags 0 and 15. Orthographic repetition consisted of a second presentation of the consonants but with different vowel marks. Phonemic repetition consisted of repeating the phonemes but with different letters (Hebrew has several pairs of letters that denote the same phoneme). The results showed differential effects of phonemic and orthographic repetition for pointed and unpointed print. For unpointed print, all three forms of repetition affected lexical decisions at lag 0, whereas at lag 15 only identity repetition was effective. With pointed print, on the other hand, phonemic repetition had a significant effect at lag 15, but orthographic repetition did not. These results suggest that the vowel marks indeed attracted subjects' attention and induced phonologic coding of the printed words. Because the same phonemic cluster appeared at the second presentation, it was recognized faster even though the orthographic spelling referred to a different meaning. When the vowels were not presented to the reader, he or she was encouraged to access the lexicon through a visual-orthographic code, and the effects of phonemic repetition disappeared.

Naming unpointed print involves postlexical phonology

Although the vowels convey to the reader unequivocally the phonemic structure of a printed word, for many words the vowel marks are not essential for locating a specific lexical entry. For these words the consonant structure is sufficient for specifying a unique word. This is because in such cases, only one phonologic pattern can be assigned to the letter string to create a meaningful word. But even considering the prevalence of phonologic ambiguity in Hebrew, the skilled reader does not need the vowel marks for fast reading. A comparison of lexical decision time in the deep unpointed Hebrew orthography and in the very shallow Serbo-Croatian orthography revealed similar, almost identical, performance (Frost, Katz, & Bentin, 1987). Being exposed to unpointed print almost exclusively, the skilled reader in Hebrew has developed reading strategies that allow him to generate the missing vowel information in the print using the lexical route following visual lexical access. This hypothesis was confirmed by a cross-language study that compared naming strategies in deep and shallow orthographies (Frost et al., 1987). In this study lexical decisions and naming performance were examined in unpointed Hebrew, in English, and in Serbo-Croatian. The results showed that, in Hebrew, the lexical status of the word (being a high-frequency word, a low-frequency word, or a nonword) had similar effects on naming and on lexical decision, suggesting that pronunciation was achieved by an addressed routine in which the whole word phonology is retrieved from lexical memory. The lexical status of the word had smaller effects on naming in English and even smaller effect on naming in Serbo-Croatian. Similar results were obtained in a second experiment that showed stronger semantic priming effects on naming in Hebrew relative to English and Serbo-Croatian, again suggesting stronger involvement of the lexicon in naming unpointed words.

Lexical decisions in unpointed print are based on fast orthographic recognition
The use of the lexical route in processing Hebrew print was also demonstrated by Koriat (1984), who examined lexical decision latencies for pointed and unpointed letter strings. In his study, Koriat used Hebrew words that had only one meaningful pronunciation in their pointed form, and found almost identical lexical decision latencies for pointed and unpointed words. Moreover, the presentation of vowel marks had similar effects on words of different length. Koriat has therefore concluded that lexical access in Hebrew is probably visual and direct, not involving phonologic mediation. In a subsequent study, however, Koriat (1985) found that the presentation of vowel marks had some beneficial effect on lexical decisions. The advantage of pointed print was larger for low-frequency words than for high-frequency words, suggesting that the use of prelexical phonology is more prevalent for infrequent words. To summarize Koriat's work, it appears that despite his initial conclusions, his data indicate that the presence of vowel marks affects visual word recognition. This evidence, however, was inconclusive.

Additional and more convincing evidence suggesting that lexical decisions in Hebrew do not involve deep phonologic processing of the printed word, emerges from studies that employed words with two meaningful pronunciations (Bentin, Bargai, & Katz, 1984; Bentin & Frost, 1987). Bentin et al. (1984) examined naming and lexical decision for unpointed consonantal strings. Some of these strings could be read as two words whereas some could only be read as one word only. The results demonstrated that phonologic ambiguity affected naming but not lexical decision performance: Naming phonologically ambiguous strings was slower than naming unambiguous ones. In contrast, phonologically ambiguous letter strings were recognized as fast as letter strings with only one meaningful pronunciation. These results suggested that, although the reader of Hebrew is indeed sensitive to the phonologic structure of the orthographic string when naming is required, lexical decisions are based on a fast familiarity judgment of the consonantal cluster and do not require a detailed phonological analysis of the printed word.

These conclusions were further supported by Bentin and Frost (1987). In this study subjects were presented with phonemically and semantically ambiguous consonantal strings. Each of the ambiguous strings could have been read either as a high-frequency word or as a low-frequency word, depending on the vowel configuration which was assigned to it. Lexical decision time for the unpointed ambiguous consonantal string was compared to lexical decision time for the unequivocal pointed printed forms of the high- or the low-frequency phonological alternatives. The results showed that lexical decisions for the unpointed ambiguous strings were faster than lexical decisions for either of their pointed (and therefore disambiguated) alternatives; explicit presentation of vowel marks did not necessarily accelerate lexical decision time. This result suggests that lexical decisions for Hebrew unpointed words may occur *prior* to the process of phonological disambiguation at least when the letter string represents two different words. In this case, the decisions are probably based on the printed word's orthographic familiarity (cf. Balota & Chumbley, 1984; Chumbley & Balota, 1984). On the basis of those studies we suggest that lexical decisions in Hebrew involve neither a prelexical nor a postlexical phonologic code. They are probably based upon the *abstract* linguistic representation that is common to several phonologic and semantic alternatives. Thus, in addition to a phonologic lexicon the Hebrew reader probably develops an "interface" lexical system that is based on consonantal strings common to several words. Whether the entries in this interface lexicon are orthographic (letters) or phonologic (phonemes that represent the consonants) in nature is hard to determine. Nevertheless, lexical processing occurs, at a first phase, at this morphophonological level. The reader accesses the abstract string and recognizes it as a valid morphologic structure. Lexical decisions are usually reached at this early stage and do not necessarily involve further phonological processing. This possibility is depicted in Figure 2.

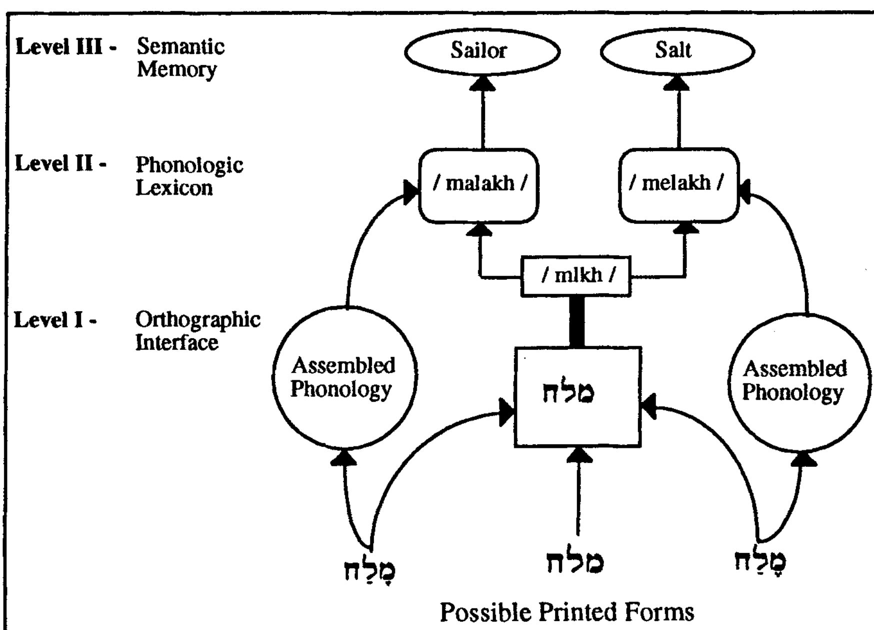

Figure 2. **A model of the lexical structure of the Hebrew reader and the possible processing of pointed and unpointed printed words.**

Naming of phonologically ambiguous words is affected by frequency factors
Although lexical decisions in Hebrew might be based on abstract orthographic representations, there is no doubt that the process of word identification continues until one of several phonological and semantic alternatives is finally determined. This process of lexical disambiguation was more clearly revealed by using the naming task. Bentin and Frost (1987) investigated the process of selecting specific lexical candidates by examining the naming latencies of unpointed and pointed words. The complete phonological structure of the unpointed word that is necessary for naming can only be retrieved postlexically, after one word candidate has been accessed. The selection of a word candidate is usually constrained by context, but we found that in the absence of context it is based on word-frequency. In contrast to lexical decisions, we found that naming ambiguous unpointed strings was just as fast as naming the most frequent pointed alternative, and that the pointed low-frequency alternative was the slowest. In the absence of constraining context, the selection of one lexical candidate for naming seems to be affected by a frequency factor: the high-frequency alternative is selected first.

Naming in pointed Hebrew also involves prelexical phonologic recoding
Another set of experiments recently completed in our laboratory (Frost, forthcoming) provides important insight regarding the use of vowel marks by the skilled reader. In this study subjects were presented with consonantal strings which were followed by vowel marks appearing at different stimulus onset asynchronies (SOA). The vowel marks were superimposed on the consonants at SOAs ranging from 0 ms (simultaneous presentation) to 300 ms from the onset of consonant presentation. In one condition the letter strings represented only one meaningful word, and in another condition the letter strings could represent two meaningful words. Subjects were required either to make lexical decisions or to name the words and nonwords on the computer screen as fast as possible. The aim of this manipulation was to examine whether subjects would be inclined to delay their decisions until the presentation of the vowel marks. The results showed similar decision times for simultaneous presentation of vowel marks and for their very late presentation (300 ms SOA). Thus, lexical decisions were only slightly affected by the delayed presentation of vowels. The effect was especially conspicuous with ambiguous letter strings. These results support the conclusions put forward by Bentin and Frost (1987), suggesting that lexical decisions in Hebrew are based on the recognition of the abstract root or orthographic cluster and do not involve access to a specific word in the phonologic lexicon.

In contrast to lexical decision, a very different strategy was revealed with lagged presentation of vowels in the naming task: the delayed presentation of the vowels delayed naming latencies, and the effects of SOA on RTs were twice as large as the effects found for lexical decisions. Thus, although the phonologic structure of the unambiguous words could be unequivocally retrieved from the lexicon following visual access (postlexical phonology), subjects were more inclined to wait for the vowels to appear in the naming task. Obviously, the longest delays occurred when the words were phonologically ambiguous. Because the correct pronunciation of these words was unequivocally determined only after the presentation of the vowel marks, subjects had to wait for the vowels to appear in order to name those words correctly. Thus, these stimuli provide a baseline for assessing the effect of lagging the vowel marks on naming latencies. When the words were phonologically ambiguous, the effects of lagging the vowel marks on RTs were twice as large as the effects found for unambiguous words, where only one pronunciation was meaningful. These results suggest that subjects adopted two parallel strategies for generating the phonology of the unambiguous printed words: on the one hand they used explicit vowel information using prelexical transformation rules (hence the greater effect of SOA on naming relative to lexical decisions latencies), on the other hand they generated the phonologic structure of the unambiguous words postlexically as well (hence the smaller effects of SOA on naming unambiguous words relative to ambiguous words). These conclusions converge with the results reported by Koriat (1984). Koriat examined the joint effects of semantic priming and vowel mark presentation, and found that semantic priming facilitated naming performance for both pointed and unpointed words, but to the same extent. The presentation of vowel marks speeded naming latencies, but so did a previous presentation of semantic context. Koriat therefore concluded that the pronunciations of unambiguous words are derived both lexically and nonlexically in parallel, and that both processes must be completed and their outcomes compared before the onset of articulation.

Processing lexical ambiguity in Hebrew

Obviously, in the absence of vowel marks, the complete phonemic structure of the letter string in Hebrew cannot be recovered by applying grapheme-to-phoneme conversion rules. Prelexical phonology, therefore, does not appears to be a viable option for the Hebrew reader when presented with unpointed print. He or she is forced to recover the missing phonological information from the lexicon. When the letter string can have only one meaningful pronunciation, the relevant phonologic representation is easy to recover lexically. However, when the letter string has two or more meaningful pronunciations, how does the reader chose among the possible alternatives?

Semantic activation of heterophonic homographs is ordered-accessed

Bentin and Frost (1987) found similar naming latencies for unpointed ambiguous letter strings and for the pointed dominant alternatives. Therefore, they suggested that readers retrieve first the dominant phonological structure of a phonologically ambiguous letter string. The significant delay in naming the subordinate pointed alternatives, relative to the unpointed and the dominant forms of the same letter string, was interpreted as supporting an ordered-access model for the retrieval of phonological information. The naming task, however, cannot disclose covert phonological selection processes. In particular, naming does not reveal whether phonological alternatives, other than the reader's final choice, had been accessed during the process of disambiguation. Although subjects overtly express only one phonological structure, (usually the high-frequency alternative), it is possible that other alternative words were generated but discarded during the output process. Therefore, a more direct measure was necessary to examine whether more than one phonologic alternative of a heterophonic homograph is automatically activated in reading single words.

The possible activation of the two phonologic alternatives related to Hebrew heterophonic homographs was examined by Frost and Bentin (1992) using a semantic priming paradigm. In this study, subjects were presented with isolated heterophonic homographs as primes, whereas the targets were related to only one of the primes' possible meanings. The targets followed the primes at different SOAs ranging from 100 to 750 ms. It was assumed that if a specific meaning of the prime was accessed, lexical decisions for targets related to that meaning would be facilitated. This experimental paradigm is similar to that used by Simpson and Burgess (1985), who examined the processing of English homophonic homographs (letter strings with two meanings but only one pronunciation). Frost and Bentin (1992) reported that, in the absence of biasing context, both meanings of heterophonic homographs were active at SOAs ranging from 250 to 750 ms from stimulus onset, whereas at a short SOA of 100 ms only the dominant meaning was active.

Phonologic disambiguation of heterophonic homographs precedes semantic activation

In another experiment reported in the same study, the processing of heterophonic homographs was compared to the processing of homophonic homographs using an identical technique. It was found that the decay of activation of subordinate meanings of homophonic and heterophonic homographs followed a similar pattern; all meanings remained active as late as 750 ms from stimulus onset. However, when the onset of activation was examined, a different pattern of results was found for heterophonic and homophonic homographs: in contrast to heterophonic homographs, both subordinate and dominant meanings of homophonic homographs were active as early as 100 ms from stimulus onset. Another finding of interest in that study was that across all SOAs, the effects of semantic priming for heterophonic homographs were larger than the effects found for homophonic homographs. Thus, it appears that both the time-course of activating the different meanings, and the amount of activation were influenced by phonological factors.

These results were interpreted to suggest that heterophonic homographs are phonologically disambiguated *before* the semantic network is accessed. Thus, phonologically ambiguous letter strings refer to different lexical entries, one for each phonological realization (see Figure 2). The alternative lexical entries are automatically activated by the unique orthographical pattern, though at different onset times: in the absence of biasing context the order of activation is determined by the relative word frequency; higher-frequency words are accessed before lower frequency words. As a consequence of the multiple-entry structure and the ordered-access process, heterophonic homographs are phonologically disambiguated prior to any access to semantic information. The overall greater priming effects found for heterophonic than for homophonic homographs suggests that when one lexical unit activates two or more semantic nodes, each of these nodes is activated less than nodes which are unequivocally related to phonological units in the lexicon. Thus, in contrast to lexical decisions, the retrieval of meaning requires the activation of the phonological structure to which the unpointed printed word refers. Note that if meaning were retrieved directly from the orthographic input, no difference should be found between processing homophonic and heterophonic homographs.

One intriguing outcome of the study with Hebrew homographs was that subordinate meanings of both heterophonic and homophonic homographs were still available and used 750 ms from stimulus onset. This result contrasts with the relatively fast decay of subordinate meanings of English homographs (Kellas, Ferraro, & Simpson, 1988; Simpson & Burgess, 1985). Because the decay pattern was similar for both types of Hebrew homographs, the divergence from English should be probably accounted for by language-related factors. One possible source of the different results obtained in Hebrew and in English may be related to the homographic characteristics of the Hebrew orthography. The ubiquity of homography might have shaped the reader's reading strategies. Because ambiguity is so common in reading, the process of semantic and phonologic disambiguation is governed mainly by context. However, the disambiguating context often follows rather than precedes the ambiguous homographs. Therefore, an efficient strategy of processing homographs should require maintaining all the phonologic or semantic alternatives in working memory until the context determines the appropriate one. Note that according to this interpretation the subordinate alternatives do not decay automatically, but remain in memory until disambiguation by context has occurred.

Both phonetic alternatives of heterophonic homographs are automatically activated
Frost (1991) presented additional evidence confirming that both phonologic representations of the ambiguous letter string are automatically activated at some stage after the printed word appears. The aim of this study was to examine directly phonologic and phonetic processing of Hebrew heterophonic homographs. Note that the measurement of semantic facilitation, as used by Frost and Bentin (1992), did not indicate directly whether the presentation of the ambiguous letter string caused the activation of the two phonologic structures related to it, or merely the activation of the two semantic meanings which were accessed directly from the print. To solve this problem, Frost (1991) employed a speech detection task and a task consisting of matching simultaneously presented printed and spoken words. These tasks have been previously shown to detect phonetic and phonologic activation that emerges from the visual presentation of meaningful letter strings (Frost, 1991; Frost & Katz, 1989; Frost, Repp, & Katz, 1988).

The speech detection task is based on an auditory illusion previously reported by Frost et al. (1988). When an amplitude-modulated noise generated from a spoken word is presented simultaneously with the word's printed version, the noise sounds more speechlike than when the print is absent. This auditory illusion suggests that subjects automatically detect correspondences between amplitude envelopes of spoken words and printed stimuli. This speech detection task was employed to examine the processing of Hebrew heterophonic homographs. Subjects were presented with speech-plus-noise and with noise-only trials, and were instructed to detect the speech in the noise. The auditory stimuli were simultaneously presented with printed letter strings that represented two phonological meaningful structures (heterophonic homographs), one dominant and the other subordinate. The bias to falsely detect speech in amplitude-modulated noise when matching print accompanies the auditory presentation occurs only when subjects detect a correspondence between the printed and the spoken information. Therefore, Frost (1991) examined whether subjects detected a correspondence between a printed heterophonic homograph and the masked spoken forms of the *two* phonologic alternatives it represents. The results demonstrated that subjects detected a correspondence between the ambiguous letter string and between the amplitude envelopes of *both* dominant and subordinate phonological alternatives. When the homographs were phonologically disambiguated by adding the vowel marks, similar effects were obtained. Moreover, subjects did not detect any correspondence when the printed pointed alternatives did not correspond to the alternative specified by the noise envelope. These results suggest then, that printed heterophonic homographs automatically activate the two alternative words they represent.

These conclusions were supported by additional experiments employing the matching task. In the matching task subjects are simultaneously presented with a printed word on a computer screen and with a spoken word via headphones. The subjects are instructed to decide as fast as possible whether the stimuli presented in the visual and the auditory modalities are the same or different words. In order for the spoken and the printed forms of words to be matched, they both have to converge at an identical lexical entry. Because the transformation of speech into an orthographic representation is by far less practiced than the transformation of spelling into phonology, the common end result of both print and speech processing in the matching task is presumably a phonological representation in the lexicon (see Frost et al., 1990, for a detailed discussion of the matching task). Frost (1991) presented subjects simultaneously with printed heterophonic homographs and with the spoken forms of the dominant and subordinate alternatives. The subjects were instructed to determine whether the printed words and the spoken words were equivalent. In some of the trials the printed homographs were presented in their pointed form and were therefore disambiguated; that is, the vowel marks unequivocally pointed to either the dominant or the subordinate alternative. In these trials the matching of the visual printed words to the spoken words did not require any ambiguity resolution. In other trials the homographs appeared unpointed, and consequently could be read in two ways. In those trials the outcome of matching the visual words to the spoken words was dependent on the specific phonological alternative generated from the ambiguous consonant string. The aim of the experiment was to compare the decision time for pointed and unpointed print. The results demonstrated that matching the unpointed printed forms of heterophonic homographs to the dominant and subordinate spoken alternatives that were presented auditorily was as fast as matching the pointed unambiguous forms to the respective spoken words. Therefore, these results confirm that both phonologic alternatives were automatically generated from the letter string.

In conclusion, the resolution of phonologic ambiguity in unpointed print is a routine procedure for the Hebrew reader. Our findings suggest that the Hebrew reader develops an orthographic lexicon that serves as an interface to the phonologic lexicon. Each orthographic entry is related to one, two, or more phonologic entries. Lexical decisions in Hebrew are given in reference to this orthographic interface prior to the activation of the phonologic lexicon. However, the activation of an orthographic entry results in the automatic activation of all phonologic entries in the mental lexicon. Semantic activation follows the activation of phonologic entries. Since, in general, while reading, the context disambiguates the phonologically abstract letter string, all phonologic and semantic alternatives remain available for relatively longer periods than in other orthographies such as English. Although all phonologic alternatives are activated following the presentation of the unpointed letters, the more frequent alternative acquires dominance when articulation is required.

Morphologic processing in Hebrew

In the present discussion of Hebrew morphology we will limit ourselves to the processing of roots by the reader. Because the root is the most important determinant of meaning in both spoken and written Hebrew, it has a unique status within the word. Both inflections and derivations in Hebrew modify the root by adding to it prefixes, infixes, and suffixes following specific word patterns. As mentioned in the beginning of this chapter, the root usually specifies a constrained semantic field that constitutes the basic information regarding the meaning of the word. Thus it is fairly reasonable to assume that its extraction from the whole word, whether spoken or written, is a primary process in the analysis of spoken or printed words. We cannot report any data regarding the perception of speech. However, the psychological reality of the status of the root in printed words was examined in several studies.

Morphologic relatedness causes long lasting repetition effects

The preferred technique for investigating morphologic processing in Hebrew was to examine the contribution of morphologic relatedness to pattern of facilitation in the repetition priming task (Bentin & Feldman, 1989; Feldman & Bentin, forthcoming). Bentin and Feldman (1989) examined the effects of morphologic repetition at lag 0 and 15 on lexical decision to the target. Specifically, they compared the effects of pure semantic relatedness, pure morphologic relatedness, and combined semantic and morphologic relatedness, on lexical decisions. In the pure semantic relatedness condition primes and targets consisted of words having different roots but related meanings. In the pure morphologic relatedness condition primes and targets shared the same root but had different meanings (as in the example depicted in Figure 1). Finally, in the combined relatedness condition primes and targets shared both root and meaning. The results showed that semantic relatedness facilitated lexical decisions only at lag 0, whereas pure morphologic relatedness exerted its effect on lexical decisions at lag 0 and 15. Semantic facilitation was greater than morphologic facilitation at lag 0. Facilitation of combined relatedness was as strong as semantic relatedness at lag 0 and similar to pure morphologic relatedness at lag 15. This outcome suggests that semantic activation and morphologic activation have different time courses and arise from two different sources. The presentation of a word containing the root has longer lasting beneficial effects on lexical decisions relative to mere semantic relatedness. Thus, it appears that a previous presentation of the abstract Hebrew root aids lexical processes such as the retrieval of related words and word meanings even at long repetition lags.

Roots are extracted by the reader while processing printed words

In another study, Feldman, Frost, & Dar (forthcoming) examined the ability of skilled readers to detach the phonologic patterns from the roots. This study was based on the segment-shifting task proposed by Feldman (1991). In the segment-shifting task subjects are presented with a printed word in which one segment is underlined. The subjects are required to detach the underlined segment from the word and append it to another word presented underneath. The subjects have to pronounce the second word with the new segment (usually appended to its end) as quickly as possible. The experimental conditions typically consist of underlining a segment that is a suffix morpheme in one word, but not in another (e.g., ER is a suffix morpheme in DRUMMER but not in SUMMER). Is it easier to detach ER from DRUMMER than from SUMMER?

The segment-shifting task was originally employed by Feldman in English and Serbo-Croatian. These languages are characterized by concatenative morphology where morphologically complex words are constructed from discrete morphemic constituents that are linked linearly: There is a base morpheme to which other elements are appended so as to form a sequence. In languages with concatenative morphology, suffixes and prefixes are regularly appended to the base morpheme in a manner that preserves its phonological and orthographic structure. In contrast to English and Serbo-Croatian, Hebrew is usually characterized by a nonconcatenative morphology. In Hebrew the phonologic word pattern is an infix, not a prefix or a suffix. It is superimposed on the root and changes both its phonologic and its orthographic structure (see the example of "dover" in Figure 1).

Experiments that employed the segment-shifting task in English and Serbo-Croatian yielded straightforward results: It is easier to detach a segment that serves as a morphemic suffix appended to a base morpheme than to detach the same sequence of letters when it is an integral part of a word that cannot be decomposed into morphemic constituents. This outcome suggests that the processing of morphologically complex words in languages with concatenated morphology entails morphemic decomposition. Such decomposition is relatively easy and straightforward when the morphemic constituents are linked linearly. In contrast to English and Serbo-Croatian, the decomposition of Hebrew derivations and inflections into root and word pattern is not as straightforward. This is because the phonemes of the root and the phonemes of the word pattern are intermixed.

Feldman et al. (forthcoming) took advantage of the fact that although formally all words in Hebrew can be defined as containing roots, not all roots are productive. Roots are considered productive if they can be inflected, and other words can be derived from them. A root is considered nonproductive if it cannot be inflected and is therefore contained in only one Hebrew word. Many words in Hebrew form a unique phonemic sequence that does not lend itself to inflections or derivations. Feldman et al. asked whether a specific phonologic word pattern can be detached more easily from words that contain productive roots than from words that contain nonproductive roots.

The experiment was similar to the typical segment-shifting task experiment. Subjects were presented with pointed words and were required to detach the sequence of vowels from the words, to superimpose them on a nonword consonant cluster, and to name it. The results showed that it was easier to detach the vowels from three consonants that were a productive root than to detach them from three letters that were not. In a second experiment similar and even stronger effects were obtained when the word patterns were not merely vowels but consisted of a sequence of vowels and consonants. These results suggest that productive roots have a special status for the Hebrew speaker and reader. Their psychological reality is reflected by their salience relative to the other letters and phonemes constituting the word. It appears that the presentation of a printed word containing a productive root results in the automatic detection of this root, such that the letters of a word are parsed into letters belonging to the root and letters not belonging to it. The important aspect of this morphologic decomposition is that the root letters do not have to appear in adjacent position (as in the second experiment). Even if they are dispersed within the word they are automatically extracted by the reader. We believe that a similar process can be demonstrated in the recognition and understanding of spoken words as well. That is, the phonemes belonging to the root have a unique psychological reality. However, this suggestion requires further investigation.

Conclusions

The pointed and unpointed Hebrew orthography presents an opportunity to examine reading processes when full or partial phonologic information is conveyed by print. This provides a significant methodological tool for investigating the effects of orthographic depth on visual word recognition, yet avoiding the pitfalls of cross-language designs. Research in reading Hebrew suggests that reading strategies are affected by the presentation or the omission of vowel marks. Efficient reading of unpointed text is based on fast recognition of orthographic clusters that become phonologically and semantically unequivocal given the available context. In contrast, the presentation of vowel marks induces a phonological processing of the printed words, which is often characteristic of shallow orthographies. This suggests that the reader of Hebrew adopts flexible reading strategies that take advantage of all possible phonemic information provided by the print.

Acknowledgment
This work was supported in part by National Institute of Child Health and Human Development Grant HD-01994 to Haskins Laboratories. We thank Avital Deutch for her insightful comments on earlier drafts of this paper.

References

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.

Bentin, S., (1989). Orthography and phonology in lexical decisions: Evidence from repetition effects at different lags. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 61-72.

Bentin, S., Bargai, N., & Katz, L. (1984). Orthographic and phonemic coding for lexical access: Evidence from Hebrew. *Journal of Experimental Psychology: Learning Memory & Cognition, 10,* 353-368.

Bentin, S., & Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology, 42,* 693-711.

Bentin, S., & Frost, R. (1987). Processing lexical ambiguity and visual word recognition in a deep orthography. *Memory & Cognition, 25,* 13-23.

Bentin, S., Hammer, R., & Cahan, S. (1991). The effects of aging and first grade schooling on the development of phonological awareness. *Psychological Science, 2,* 271-274.

Bentin, S., & Leshem, H. (in press). On the interaction between phonological awareness and reading acquisition: It's a two-way street. *Annals of Dyslexia.*

Berman, R. A. (1978). *Modern Hebrew structure.* Tel Aviv: University Publishing Project.

Bertelson, P. (1986). The onset of literacy: Liminal remarks. *Cognition, 24,* 1-30.

Bertelson, P., & de Gelder, B. (1990). The emergence of phonological awareness: Comparative approaches. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the Motor Theory of Speech Perception.* Hillsdale, NJ: Erlbaum.

Bradley, L, & Bryant, P. E. (1983). Categorizing sounds and learning to read: A causal connection. *Nature, 301,* 419-421.

Bradley, L, & Bryant, P. E. (1985). *Rhyme and reason in reading and spelling.* Ann Arbor: University of Michigan Press.

Calfee, R., Chapman, R., & Venezky, R. (1972). How a child needs to think to learn to read. In L. W. Gregg (Ed.), *Cognition in Learning and Memory.* New York: Wiley.

Chumbley, J. I., & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory & Cognition, 12,* 590-606.

Conrad, R. (1972). Speech and reading. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye.* Cambridge, MA: MIT Press.

Feldman, L. B., & Bentin, S. (submitted). Facilitation due to repetition of disrupted morphemes: evidence from Hebrew.

Feldman, L. B., Frost, R., & Dar, T. (forthcoming). Processing morphemes in different languages.

Frost, R. (1991). Phonetic recoding of print and its effect on the detection of concurrent speech in amplitude modulated noise. *Cognition. 39,* 195-214.

Frost, R. (forthcoming). Prelexical and postlexical strategies in reading: Evidence from a deep and a shallow orthography.

Frost, R., & Bentin, S. (1992). Processing phonological and semantic ambiguity: Evidence from semantic priming at different SOAs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 58-68.

Frost, R., Feldman, L. B., & Katz, L. (1990). Phonological ambiguity and lexical ambiguity: Effects on visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 569-580.

Frost, R., & Katz, L. (1989). Orthographic depth and the interaction of visual and auditory processing in word recognition. *Memory & Cognition, 17,* 302-311.

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,* 104-114.

Frost, R., Repp, B. H., & Katz, L. (1988). Can speech perception be influenced by a simultaneous presentation of print? *Journal of Memory and Language, 27,* 741-755.

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the time course of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 601-609.

Koriat, A. (1984). Reading without vowels: Lexical access in Hebrew. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X.* Hillsdale, NJ: Erlbaum.

Koriat, A. (1985). Lexical access for low and high frequency words in Hebrew. *Memory & Cognition, 13,* 37-44.

Liberman, I. Y., & Shankweiler, D. (1985). Phonology and the problems of learning to read and write. *RASE, 6,* 8-17.

Liberman, I. Y., Shankweiler, D., Fisher, F. W., & Carter, B. J. (1974). Explicit syllable and phoneme segmentation in young children. *Journal of Experimental Child Psychology, 18,* 201-212.

Lundberg, I., Olofsson, A., & Wall, S. (1980). Reading and spelling skills in the first school years, predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology, 21,* 159-173.

Mann, V. A. (1984). Longitudinal prediction and prevention of early reading difficulty. Annals of Dyslexia, 34, 117-136.

McCusker, L. X., Hillinger, M. L., & Bias, R. C. (1981). Phonological recoding and reading. *Psychological Bulletin, 88,* 217-245.

Morag, S. (1972). *The vocalization systems of Arabic, Hebrew, and Aramaic.* Mouton & Co.'s- Gravenhage.

Navon, D., & Shimron, Y. (1981). Does word naming involve grapheme to phoneme translation? Evidence from Hebrew. *Journal of Verbal Learning and Verbal Behavior, 20,* 97-109.

Navon, D., & Shimron, Y. (1984). Reading Hebrew: How necessary is the graphemic representation of vowels? In L. Henderson (Ed.), *Orthographies and reading: Perspectives from cognitive psychology, neuropsychology, and linguistics.* London, Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology, 67,* 461-469.

Shankweiler, D., & Liberman, I. Y. (1976). Exploring the relations between reading and speech. In R. M. Knights & D. J. Bakker (Eds.), *The Neuropsychology of learning disorders: Theoretical approaches.* Baltimore, MD: University Park Press.

Simpson, G. B., & Burgess. C. (1985). Activation and selection processes in the recognition of ambiguous words. *Journal of Experimental Psychology: Human Perception & Performance, 11,* 28-39.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 3

Basic Processes in Reading: Is the Orthographic Depth Hypothesis Sinking?

Derek Besner
University of Waterloo
Marilyn Chapnik Smith
University of Toronto

Overview

Orthographies may be defined as either "shallow" or "deep," depending on the ease of predicting the pronunciation of a word from its spelling. In shallow orthographies, the spelling-sound correspondence is direct: given the rules, anyone can immediately "name" the words correctly. In contrast, in deep orthographies the relationship is less direct, and readers must learn the arbitrary or unusual pronunciations of irregular words such as "yacht." A consequence of this linguistic difference between deep and shallow orthographies is that it is often assumed that the oral reading of shallow orthographies is qualitatively different from the oral reading of deep orthographies. Our goal in this chapter is to provide evidence favoring an alternative viewpoint: Although the linguistic description of "deep" and "shallow" orthographies is quite different, the psychological operations applied to their oral reading share more in common than previously acknowledged. This view is termed the "universal hypothesis." In what follows we first briefly discuss oral reading in alphabetic English, a deep orthography. We then discuss the orthographic depth hypothesis and the evidence that purports to support it. Old and new evidence from studies of Persian, Spanish, Dutch, Italian and Croatian which appears to undermine the essential tenents of the orthographic depth hypothesis is then reviewed and assessed. Finally, a reformulation of the orthographic depth hypothesis is considered but ultimately rejected based on studies of Japanese Kana. Overall, the data are better fit by the universal hypothesis than by the orthographic depth hypothesis. Finally, we consider some meta-theoretical issues concerning multiple routines and local versus distributed representations at "lexical" and semantic levels.

Three ways to convert print into speech in a deep orthography

A widely held assumption is that there are at least three different ways in which print can be transformed into speech when reading alphabetic English (e.g., see Patterson & Coltheart's 1987 review). Our take on this assumption is represented in Figure 1.

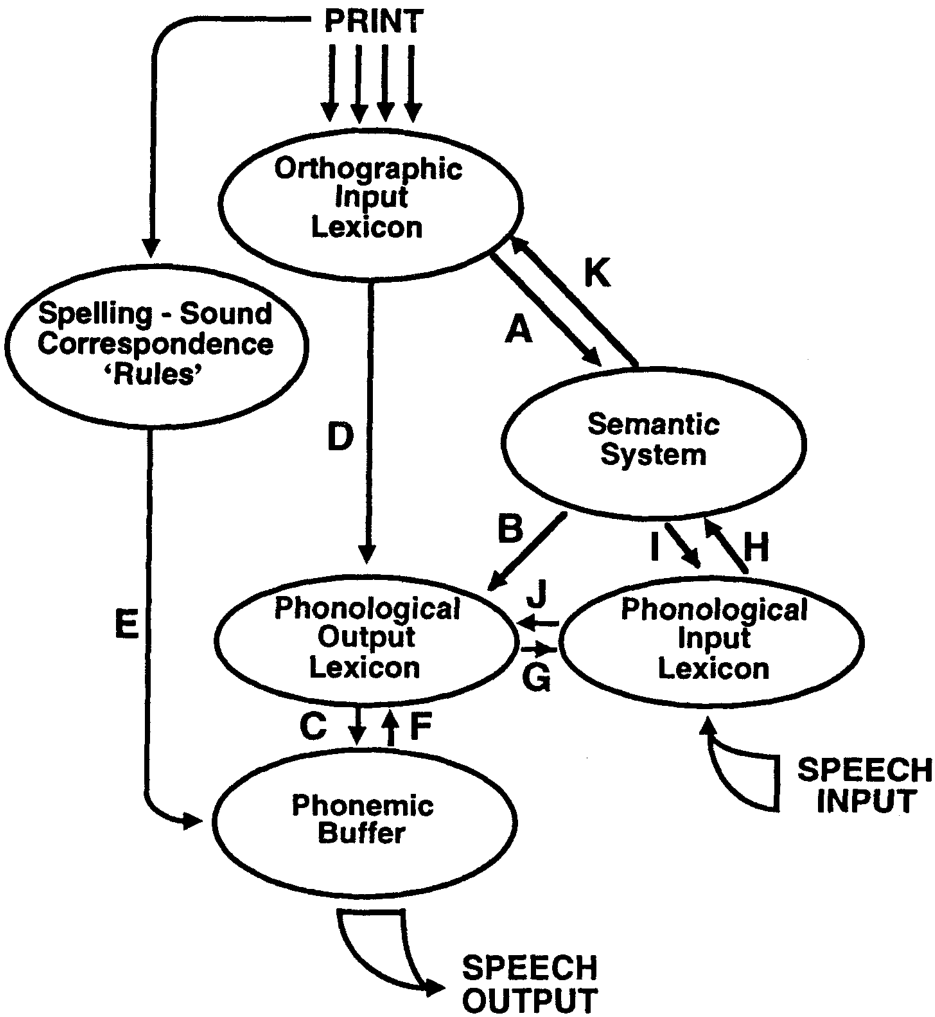

Figure 1.

The orthographic input lexicon consists of lexical entries for all the spellings of words that a reader knows, (neither semantics nor phonology are represented here, and there are no lexical-lexical associations). The semantic system represents meaning (we shall conveniently ignore the issue of how this knowledge is represented). The phonological input and output lexicons consist of lexical entries for all the words that a reader knows, specified in terms of their sounds.[1] The precise nature of this phononological code need not concern us here (but see Besner & Davelaar, 1982 for some comments). Finally, the phonemic buffer holds information about phonemes.

[1] We ignore here the issue of what motivates the distinction between input and output phonological lexicons and whether there is unequivocal evidence to support this distinction.

One way to name a word is via semantics (pathway A-B). When a word is presented it activates its lexical entry in the orthographic input lexicon. This in turn leads to activation in the semantic system and then in the phonological output system. A second way to name words is for the orthographic input lexicon to directly activate the phonological output lexicon (pathway D). A third way is to use the assembled routine.[2] In naming via the assembled routine the reader utilizes spelling-sound correspondence knowledge to translate subword orthographic segments directly into subword phonological segments and then assembles these phonological segments into a speech program. This, in its simplest form, is carried out via pathway E. Although all three ways of naming a word involve activation of the phonemic buffer, pathway E does so directly whereas the other two methods do so via the phonological output lexicon (pathway C).

Neuropsychological evidence from patients reading a deep orthography

Some of the best evidence for the existence of these three different ways of reading a word aloud can be found in single case studies of reading abilities in patients with an acquired dyslexia consequent to brain damage. This single case approach has been adopted in a number of important investigations (e.g., Coltheart, Patterson, & Marshall, 1980; Patterson, Marshall, & Coltheart, 1985).

a) Evidence for the use of Pathway A-B. Some patients read "semantically," employing the A-B pathway at least some of the time. When asked to read single words such as TULIP aloud they make semantic errors (e.g., CROCUS; see Coltheart, 1980 for a review). A standard interpretation here is that pathways D and E are unavailable to the reader because of one or more lesions, and that the semantic route (pathways A-B) though somewhat damaged, will nonetheless support some reading (e.g., Morton & Patterson, 1980).

b) Evidence for the use of Pathway D. Other patient data are consistent with the idea that words can also be read aloud via pathway D. Such patients make no semantic errors when reading either content words or function words. However, they are very poor at reading nonwords. The inability to read nonwords provides evidence that pathway E is drastically impaired. Since a large number of function words typically have no semantic representations, they cannot be read via the semantic route. Hence, pathway D is the only remaining route that can deal with function words which have no semantics (Patterson, 1982; Shallice & Warrington, 1980).

c) Evidence for the use of Pathway E. It is assumed that although words with typical spelling-sound correspondences can be named via simple grapheme-phoneme "rules," words with unusual spelling-sound correspondences require access to lexical knowledge. Some patients are much more successful at reading words such as CAVE that are regular in terms of their spelling-sound correspondences than words such as HAVE whose spelling-sound correspondences are exceptional. This implicates pathway E, since there is no reason that the other two pathways should be differentially sensitive to regularity of spelling-sound correspondences (e.g., see Patterson, Marshall, & Coltheart, 1985). The notion here is that a lesion (or lesions) has either destroyed entries in the orthographic

[2]This is a simplified view. Marcel (1980), Campbell and Besner (1981), Rosson (1983;1985), Brown and Besner (1987), and McCann and Besner (1987) among others, have argued that the assembled routine(s) used when reading english is influenced by lexical/syntactic information. As we shall see below, the assembled route, even in shallow orthographies, is sensitive to lexical influence.

lexicon and/or impaired access to them, thereby making oral reading heavily reliant on the assembled route.

Print to speech in intact readers of a deep orthography

Which routes do normal, intact readers use when reading alphabetic English aloud? According to the "dual route" model often referred to in the word recognition literature, naming can occur either via application of spelling-sound correspondence rules (the assembled route) or via activation of the orthographic input lexicon (the addressed route). Although there is clear neuropsychological evidence for the existence of both pathways A-B and D, we do not yet know whether intact adults use pathway A-B, pathway D, or both. What is important is that both pathways rely on initial activation of the orthographic input lexicon—i.e., they "address" the appropriate lexical entry. Consequently, for present purposes the term "addressed route" is used to indicate the use of either pathway A-B or D, singly or in combination.

Reading via the semantic system (pathways A-B) is hardly disputable, given that readers typically read for meaning. Nonetheless, we are able on occasion to read aloud *without* awareness of meaning (as may occur when reading stories to our children), indicating that lexical information may address output phonology directly via pathway D. Pathway E (the assembled route) is clearly available since people can read nonwords such as ISH, LAR, and FON, even when these letter strings are "hermits" with few if any orthographic neighbors. Further evidence for use of the assembled route, particularly when reading low frequency words aloud, is based on the observation that words with regular spelling-sound correspondences are much less affected by word frequency, a lexical variable, than are words whose spelling-sound correspondences are irregular (e.g., Paap & Noel, 1991; Seidenberg et al., 1984; see also Patterson & Coltheart's 1987 review). The logic is that the assembled routine is by definition not sensitive to word frequency, since it operates at the subword level, whereas the pathways used to read irregular words (A-B and D) *are* sensitive to word frequency. Regular words will therefore show a smaller word frequency effect than irregular words to the extent that they are processed by the assembled routine.

It should be noted that reading via pathways D or E does not preclude semantic activation. As can be seen in Figure 1, activity in the phonemic buffer can ultimately result in activation of the semantic system, which in turn re-activates the phonological output lexicon. On this view the standard finding of "priming" by a related context in oral reading does not demand the interpretation that normal subjects are using the addressed routine. This point is central to the framework entertained here and will be developed in detail in later discussion.

The orthographic depth hypothesis

The orthographic depth hypothesis in its strong form makes a very simple claim: There is no orthographic input lexicon in the minds of readers who process orthographies which consist entirely of words with consistent spelling-sound correspondences.[3] The argument is that orthographic access to semantics and the direct mapping from orthographic input lexicon to phonological output lexicon only exists in scripts with inconsistent spelling-sound correspondences, and does so precisely because of this inconsistency. Consider the

[3]We assume that spelling, as in writing to dictation, is driven by an orthographic output lexicon and/or phoneme-grapheme rules.

segment OU in COUGH, THROUGH, BOUGH, DOUGH and FOUR. If phonology is typically assembled without recourse to whole word knowledge there is no reliable way to know which pronunciation to assign to the segment OU. Hence the need to evolve non-phonologically mediated associations between print and semantics, and print and speech.

To put it another way, the strong form of the orthographic depth hypothesis denies that the normal brain ever develops associations between orthographic patterns and semantics in scripts where spelling-sound correspondences are entirely consistent. Given this assumption it can immediately be seen by reference to Figure 1 that the only way to get from print to speech or from print to semantics *in shallow orthographies* is to rely upon the assembled routine (pathway E). This sentiment is reflected in the writings of several authorities:

> Completely regular languages...are read with strategies that differ from those used with less regular ones. In many regular languages a small set of grapheme-phoneme correspondences can unambiguously define all of the utterances in the language. It is possible that in these languages the lexical route simply does not exist.... (Bridgeman, 1987, p. 331)
>
> To conclude, the Serbo-Croatian orthography is phonologically very regular (permitting a valid prediction of how a word sounds solely on the basis of the letters comprising the word) and as such encourages neither the development of options for accessing the lexicon, nor, relatedly, a sensitivity to the linguistic situations in which one option fares better than another. (Turvey et al. 1984, p. 88)

Criteria for deciding which routine is used in reading shallow orthographies

Katz and Feldman (1983) and Frost, Katz, and Bentin (1987) were among the first to argue that neither word frequency nor priming effects should occur when reading words aloud in a shallow orthography. This is because readers of a shallow orthography rely exclusively on the assembled routine, a routine which operates at subword levels.[4] In contrast, since both word frequency and semantic context are lexical/semantic manipulations,[5] they *are* expected to play a role in the oral reading of a deep orthography in which the addressed routine dominates.[6]

The evidence in support of these contentions comes from studies of word naming in deep and shallow orthographies. In alphabetic English, word naming has repeatedly been demonstrated to be affected by both word frequency (e.g., Balota & Chumbley, 1985; Forster & Chambers, 1973; Frederiksen & Kroll, 1976; McCann & Besner, 1987; Monsell et al., 1989; Scarborough et al., 1977; Waters & Seidenberg, 1985) and by the prior presentation of a related context (see Neely's 1991 exhaustive review). In contrast,

[4] Though note, in a deep orthography, Rosson's (1983) demonstration that the presentation of SOFA primes the pronunciation of a nonword like LOUCH and Lukatela and Turvey's (1991) demonstration that the presentation of TABLE primes the oral reading of a pseudohomophone like CHARE. The assembled route can contribute to oral reading but its products driven through, or influenced by semantics, even in a shallow orthography. This idea is developed further in several later sections.

[5] We ignore here the numerous clever demonstrations of phonological effects in lexical decision by Lukatela, Turvey and colleagues. Our view is that while such demonstrations tell us that phonology is utilized some of the time, they do not inform us as to whether orthographic information is ever a sufficient basis for lexical access, semantic access, or for driving output phonology via one of these systems.

[6] It is often argued that there is little in the way of an assembled routine for reading logographs such as Japanese and Chinese Kanji. Butterworth and Wengang (1991) show that this is not the case in their investigations of Chinese patients with an acquired dyslexia.

Katz and Feldman (1983) and Frost et al. (1987) have failed to find either priming effects (Frost et al., 1987; Katz & Feldman, 1983) or word frequency effects (Frost et al., 1987) in the oral reading of Serbo-Croatian. This absence of priming and word frequency effects is consistent with the view that a purely nonlexical process drives the transformation of print into speech in a shallow orthography. Their observations are particularly interesting given that (a) the same methodology and similar materials *did* give rise to priming and word frequency effects in the naming of two languages with deep orthographies, Hebrew and English, and that (b) priming and word frequency effects were also observed in Serbo-Croatian when the task was changed to *lexical decision*, a task which is often thought to require lexical access.

Falsifying the orthographic depth hypothesis

The results of Katz and Feldman (1983) and Frost et al. (1987) appear to provide strong support for the orthographic depth hypothesis. However, it should be noted that in contrast to the large number of papers showing priming and frequency effects in deep orthographies, the attempt to prove the null hypothesis of no priming and no frequency effects in the oral reading of shallow orthographies rests upon a very narrow data base. There have been only *two* reports that a related context does not facilitate naming relative to an unrelated context (Frost, Katz, & Bentin, 1987; Katz & Feldman, 1983), and only *one* report that word frequency does not affect naming (Frost et al., 1987). Moreover, these null findings are all based on reading Serbo-Croatian. In this paper we will review evidence that, contrary to the results of Katz and Feldman and of Frost et al., it *is* possible to show lexical involvement in the oral reading of many shallow orthographies.

I Evidence for lexical involvement in the oral reading of words in shallow orthographies

A) Observing priming and word frequency effects when reading *words:* The role of nonwords

In the two critical papers of Katz and Feldman and of Frost et al. there is an interesting procedural issue: both words and nonwords served as targets. The problem is that since nonwords can only be read by the assembled route, their presence may bias the word recognition system to take the path of least resistance: if the assembled routine can read both words and nonwords, then why not use it? However, we are more interested in a different question—how does a word recognition system read words aloud in the *absence* of nonwords? Further, in our opinion, attention has been overly focused on how *one* shallow script, Serbo-Croatian, is read. We therefore turn first to a consideration of Persian. This shallow script has special properties which permit a test of the orthographic depth hypothesis *within a script rather than between scripts.*

Oral reading of Persian

Written Persian is transcribed by a modified version of Arabic script, even though Persian is a member of the Indo-European family of languages (unlike Arabic which is a member of the Semitic family of languages). The Persian language has only six spoken vowels, which are represented in script in two different ways. Three of its vowels are each represented by a letter of the alphabet. The other three vowels are represented by diacritics. This diacritic spelling is used only with beginning readers; fluent readers are accustomed to reading script without the diacritics (much like reading vowel free Hebrew).

One of the interesting aspects of written Persian is that although spelling-sound correspondences are always consistent, some words require more lexical processing than others because they vary in terms of their phonological transparency. Words which contain vowel letters are defined as being phonologically transparent, whereas words with diacritics are defined as phonologically opaque when the diacritics are absent (interested readers may consult Khanlari (1979) and Madison (1984) for a fuller treatment). Transparent words are easy to read (at least in principle) by applying subword spelling-sound correspondence knowledge since these correspondences are always consistent. It is theoretically more difficult to read phonologically opaque words (where the vowels are not specified), since lexical information must be recruited in order to identify these words. This difference allows for a within script investigation of the orthographic depth hypothesis. If this hypothesis applies to the oral reading of Persian, then given the criteria adopted by Frost et al. (1987) and by Katz and Feldman (1983) it is expected that phonologically transparent words would always be insensitive to both word frequency and priming effects.

Baluch and Besner (1991) examined the effects of word frequency and priming on the oral reading of Persian by highly literate adult Iranians. In accordance with the orthographic depth hypothesis, phonologically opaque words were sensitive to both word frequency and priming in two experiments, whereas phonologically transparent words were not affected by either word frequency or priming. However, this pattern of data was found *only* if a substantial subset of the items to be named were *nonwords*. When nonwords were deleted from the experiment, the very same set of phonologically transparent words were *also* found to be affected by both word frequency and priming. These results were extended and replicated in an experiment which investigated the oral reading of phonologically transparent words in the absence of any phonologically opaque words. Under these conditions, again, a word frequency effect occurred provided there were no nonwords in the set, and again, the word frequency effect disappeared when nonwords *were* included in the set.

These results suggest that, at least in Persian, the visual word recognition system is much more flexible than the orthographic depth hypothesis would seem to allow. Baluch and Besner concluded that phonologically opaque words are always read by the addressed route, since they are always affected by both word frequency and priming. In contrast, the way in which phonologically transparent words are read depends on the presence or absence of nonwords. If there are no nonwords in the set, they are also read by the addressed route, and consequently are affected by both frequency and priming. However, in the presence of nonwords (which by definition cannot be read by the addressed route given that the orthographic lexicon contains no entries for nonwords) subjects rely on the assembled route, a route which will always provide a correct response for both transparent words and nonwords.

If we adopt the criteria utilized by Katz and Feldman (1983) and by Frost et al. (1987) then we can see that the orthographic depth hypothesis does not fare well in accounting for the data from Persian readers. Rather, a flexible dual route model gives a better account.

It is tempting to suppose that even when orthographies differ quite dramatically at the linguistic level of description there are nonetheless some psychological universals that apply to their reading. We therefore returned to the reading of Croatian with a simple question in mind: would frequency and priming effects emerge in the oral reading of Croatian if only words appeared in the experiment?

Oral reading of Serbo-Croatian and Croatian

Carello, Lukatela, and Turvey (1988) have argued that any priming and word frequency effects in the oral reading of Serbo-Croatian may be attributed to an automatic lexical check. They suggest that this check occurs when reading Serbo-Croatian because there are two scripts, Cyrillic and Roman, which share some orthographic characters that are pronounced differently in each script. To circumvent this argument, the readers tested in the naming experiments described here were all adult Croatians from the Kitchener-Waterloo area who typically read *only* the Roman orthography. Hence the argument that an automatic lexical check forms part of the oral reading process lacks force. Since these experiments will be published elsewhere the details of the methodology are omitted, aside from noting that the general procedures followed those of Besner and Hildebrandt (1987), Besner, Patterson, Lee, and Hildebrandt (1992) and Baluch and Besner (1991).

i) A word frequency effect in Croatian. In Experiment 1, 40 Croatian subjects named, one word at a time, 36 high frequency and 36 low frequency Croatian words that were matched for length and shared the same initial phoneme. The data appear in Table 1. The 50 ms difference between high and low frequency words was reliable in both subject and item analyses.

ii) A word priming effect in Croatian. In Experiment 2, 40 subjects who had not participated in Experiment 1 read a new set of 80 words printed in the Roman alphabet. Half of these words were preceded by a related word and half by an unrelated word. Two subsets of items were employed such that a individual target was seen only once by an individual subject and was seen in a related context by half the subjects and in an unrelated context by the remaining subjects. In the data shown in Table 2, the 16 ms difference between related and unrelated conditions was reliable in both subject and item analyses.

Table 1

RT (ms) and Error rate to name words in the Roman alphabet.

	High Frequency	Low Frequency
RT	696	746
%E	3.3	7.2

Table 2

RT (ms) and Error rate to name target words in the Roman alphabet.

	Related	Unrelated
RT	656	671
%E	4.8	5.2

iii) Conclusions from the Croatian data. The question considered here was whether word frequency and priming effects would emerge when Croatian words were read in the absence of any nonwords. Given the theoretical framework adopted by Katz and Feldman and by Frost et al., they should not. Nonetheless, both effects were evident. If the presence of word frequency and priming effects is a signature of the addressed routine then we must conclude that this route plays a major role in the oral reading of Croatian, just as it does in English and Persian. Hence, the results of these two experiments again suggest that the orthographic depth hypothesis must be rejected or reformulated.

Other evidence

The results reported above do not stand in isolation. In an unpublished study consisting only of words, Seidenberg and Vidanović (1985) found both priming and frequency effects in the oral reading of Serbo-Croatian, and Carello et al. (1988) have reported a reliable priming effect.

Studies of three other shallow orthographies are also relevant. Sebastián-Gallés (1991) observed both priming and word frequency effects in the oral reading of Spanish,[7] while Tabossi and Laghi (1992) found a priming effect in the oral reading of Italian which was eliminated in another condition which inserted nonwords. Finally, Hudson, and Bergman (1985) observed a word frequency effect in the oral reading of Dutch which again was eliminated when nonwords were included in the stimulus set. Again, these results suggest that even when orthographies differ quite dramatically at the linguistic level of description, there are nonetheless some psychological universals that apply to their reading.

We turn now to a consideration of the oral reading of Japanese. These shallow scripts are of special interest to the present discussion because they allow a manipulation that many other shallow scripts do not: Like English, it is possible to hold the phonology constant yet force the reading of words via the assembled routine by making the orthography unfamiliar at the whole word level.

B) Reading orthographically familiar and unfamiliar words: Evidence from Japanese Kana

Written Japanese consists of three scripts: logographic Kanji and two forms of the syllabic Kana script—Katakana and Hiragana. Katakana is used to write foreign loan words (e.g., computer, telephone) while Hiragana is used for grammatical morphemes. Given that the spelling-sound correspondences in Kana are consistent, thereby qualifying Kana as a script with a shallow orthography, do intact readers *always* read aloud by recourse to the assembled route, as the orthographic depth hypothesis would hold, or is there some contribution from the addressed route?

To answer this question, Besner and Hildebrandt (1987) compared naming times for words normally printed in Katakana with words printed in Katakana that were

[7] The data from both Sebastián-Gallés (1991) and Carello et al., (1988) could be taken to warn us that additional factors are likely at play here since they found a priming effect in oral reading despite the presence of nonwords as targets. We have also found a persisting word frequency effect in Croatian despite the presence of nonwords in the list for some readers, but not others. Individual differences will need to be addressed in any comprehensive account, but all the data from Italian, Persian, Spanish and Dutch and Serbo-Croatian available at present can be summarized in the following way: The presence of nonword targets in the stimulus set either eliminates or reduces the magnitude of priming and word frequency effects as compared to when only words serve as targets.

transcriptions from Kanji. Since the latter words are orthographically unfamiliar at the whole word level they *must* use the assembled route. If Japanese readers used *only* pathway E, as predicted by the orthographic depth hypothesis, then orthographic familiarity at the whole word level should not facilitate performance. Hence, words normally written in Katakana should be no faster than words written in Katakana which are transcriptions from Kanji, despite the fact that the latter are orthographically unfamiliar at the whole word level.

In contrast to this prediction, Besner and Hildebrandt (1987) found that words normally written in Katakana were named significantly faster than words in Katakana which were transcribed from Kanji. This finding implies that even when reading the shallow Kana orthography readers make use of pathway A-B and/or D at least some of the time. This result is inconsistent with the orthographic depth hypothesis as currently formulated.

New findings

Converging evidence in support of the claim that the addressed route plays a role in the oral reading of Kana is provided by a recent experiment by Besner et al. (1992). Subjects in this experiment named single words printed in either Katakana or Hiragana. Some words appeared as normally printed in Katakana, and others as normally printed in Hiragana. These same words were also transcribed into the other script. This design thus allows a comparison of orthographic familiarity at the whole word level while holding the target word constant at the phonological level. The results, given in Table 3, are clear-cut.

There is no difference in the time to name words which are normally written in Katakana as compared to words normally written in Hiragana. However, naming time increases reliably when words are rendered orthographically unfamiliar at the whole word level by transcribing them from one script into the other. These results are therefore consistent with those reported by Besner and Hildebrandt (1987). The finding that orthographic familiarity at the whole word level affects performance implies that subjects use pathways A-B and/or D at least some of the time.

Table 3

RT (ms) and Error rate to name words in Katakana and Hiragana.

		Presented in			
		Katakana		Hiragana	
		RT	%E	RT	%E
normally seen in	Katakana	605	3.2	670	3.4
	Hiragana	643	4.6	596	2.2

II Evidence of lexical involvement when subjects are forced to use the assembled route to read aloud

A) Comparison of the oral reading of nonwords and transcribed words

Is there lexical involvement when readers are *forced* to use the assembled route? One way to answer this question is by comparing naming times to *nonwords* printed in Katakana with *words* printed in Katakana which are transcriptions from Kanji. Nonwords, by definition, can only be read via the assembled routine. Transcribed words, because they are orthographically unfamiliar and hence without representation in the orthographic input lexicon, can similarly *only* be read by recourse to the assembled route. If there is no lexical influence on the assembled route, then there should be no difference between these two conditions. However, Besner and Hildebrandt (1987) found clear evidence of lexical involvement: the unfamiliar transcriptions were named significantly more efficiently than the nonwords (although, as described above, slower than the familiar words). Wydell (1991) reports similar results. These experiments thus provide strong evidence that there is lexical involvement in the oral reading of a shallow orthography even when subjects are forced to use the assembled route.

B) Word frequency and priming effects

The question considered in this section is whether frequency and priming effects ever occur under conditions in which readers are *forced* to use the assembled route. Any demonstration that such frequency and priming effects occur when the assembled route is used in a shallow orthography serve to undermine the claim that these effects exclusively signal the action of the addressed route.

i) Word frequency effects. Recently, word frequency effects *have* been observed under conditions which force use of the assembled route. Wydell (1991) found that when subjects named Kana that had been transcribed from Kanji (hence necessitating use of the assembled route because it is orthographically unfamiliar) there was a robust effect of word frequency *provided that there were no nonwords in the list.*

ii) Priming effects. Evidence that priming occurs even when the assembled routine is used is provided by two experiments in which subjects were asked to read nonwords. In one experiment subjects read nonwords that rhymed with real English words, such as LOUCH (Rosson, 1983) while in another experiment subjects read nonwords such as CHARE that are homophonic with English words such as CHAIR (Lukatela and Turvey, 1991). In both experiments nonwords were preceded by words which half the time were semantically related to the rhyming or homophonic English word (e.g., SOFA or TABLE). Clear priming effects were observed in the naming of these nonwords. Given that the targets in these experiments could only be read by using the assembled route (since they are nonwords), these results provide a clear demonstration that the presence of priming effects in English cannot be unequivocally attributed to the action of the addressed route.

Although the above demonstrations of priming when using the assembled route are convincing, it should be noted that there has been considerable debate in some quarters as to whether the addressed and assembled routes are truly independent in English (cf. Humphreys and Evett, 1985 and associated commentaries). A question of interest then is whether such priming effects can be demonstrated in a *shallow* orthography under conditions in which subjects can only use the assembled route. Such evidence follows.

Buchanan and Besner (1992) replicated the essentials of Besner et al. (1992) with an additional manipulation: each Kana target word was preceded by a Kanji word prime, half of which were related to the target. Half the Kana words were orthographically familiar, whereas half were transcriptions and hence orthographically unfamiliar (thereby forcing readers to use the assembled route). If orthographically unfamiliar Kana is named faster in the related as compared to the unrelated condition then this demonstrates that use of the assembled routine in a shallow orthography can produce priming.

Twelve native Japanese readers named target words printed either in Katakana or Hiragana that were preceded by a prime word printed in Kanji. As indicated above, the target words were either orthographically familiar or orthographically unfamiliar because of having been transcribed from Katakana into Hiragana. Character length of the target was also manipulated. The data can be seen in Table 4.

The RT data yielded a reliable three-way interaction between priming, length, and familiarity. Orthographically familiar words were named faster than orthographically unfamiliar words, and there were larger length effects for orthographically unfamiliar words than for orthographically familiar words. Most importantly, *orthographically unfamiliar words showed a significant priming effect* (although this priming effect was restricted to longer words). This result is the first demonstration that forcing subjects to use the assembled routine in a shallow orthography can produce a priming effect.

Surprisingly, orthographically familiar words did not yield any reliable priming. This finding has some interesting implications for the notion that contextual facilitation is an automatic process (cf. Neely, 1991) but will not be treated at length here; interested readers may consult Buchanan and Besner (1992). For present purposes it is sufficient to note that faster naming times to orthographically familiar words as compared to orthographically unfamiliar words has always been attributed to the operation of the addressed route. Despite evidence for the use of this route, there is no priming in this condition of the present experiment.

Table 4

RT (ms) to name Orthographically Familiar and Unfamiliar words in Japanese as a function of Relatedness and Character Length.

	Orthographically Familiar		Orthographically Unfamiliar	
	Related	Unrelated	Related	Unrelated
Short words	522	528	558	560
Long words	546	552	586	642

Error rates were less than 2% in all conditions

Reconsidering the orthographic depth hypothesis

Since we have covered a lot of ground it is useful to review the arguments developed thus far and summarize where we currently stand. The original form of the orthographic depth hypothesis made the straightforward claim that there is no orthographic input lexicon for readers of a shallow orthography. Rather, subword spelling-sound correspondence rules are used to convert print directly to speech via the assembled routine. Since this version of the hypothesis holds that there is no lexical involvement when the assembled routine is used, there can be no effects of either word frequency or context on naming. In contrast, the existence of an orthographic input lexicon for readers of deep orthographies implies that both these factors *do* play a role. In fact, according to the orthographic depth hypothesis, the *presence* of word frequency and priming effects is considered sufficient evidence for use of the addressed routine, whereas the *absence* of these effects is considered sufficient evidence for use of the assembled routine. In this paper we provided evidence which refuted both these claims.

Consider first the presence of word frequency and priming effects. Contrary to the assertion that these factors do not play a role in the reading of shallow orthographies, we reviewed evidence of their existence in a large number of shallow orthographies including Persian, Italian, Dutch, Spanish and Serbo-Croatian. An important moderating factor was whether or not nonwords were included in the target set. If nonwords were included, priming and frequency effects were typically not found. If there were no nonwords both effects emerged. It is therefore not true that the oral reading of shallow orthographies is insensitive to priming and frequency effects. Further evidence for lexical involvement in the reading of shallow orthographies comes from the comparison of naming times to orthographically unfamiliar words and nonwords. The finding of faster naming for words than for nonwords is again evidence of lexical involvement. Hence, one of the cornerstones of the orthographic depth hypothesis, that the oral reading of shallow orthographies is not sensitive to lexical manipulations, is clearly false.

Does the *presence* of lexical factors in oral reading constitute irrefutable evidence for use of the addressed routine; no! In several of the studies reviewed above, manipulations were included which *forced* subjects to use only assembled routine. This was done by having subjects name orthographically unfamiliar words and nonwords. Both types of stimuli must be processed via the assembled routine, but this contrast yielded evidence for lexical involvement. Orthographically unfamiliar words were named faster than nonwords, and both priming and frequency effects were found. Hence a second tenet of the orthographic depth hypothesis was falsified.

Finally, a third tenet of the orthographic depth hypothesis, that use of the addressed routine will always produce priming effects, was not confirmed. It is possible under certain circumstances to show evidence that the addressed routine is being used, yet fail to find contextual priming effects.

In light of these data, the original formulation of the orthographic hypothesis must clearly be rejected. We now turn to a consideration of the model depicted in Figure 1 to determine whether it is possible to revise the orthographic depth hypothesis to accommodate the various effects reviewed above. To anticipate our conclusions, whereas the exclusive use of the assembled route is consistent with some of the data, it is not capable of accommodating the effect of orthographic familiarity seen in the reading of Japanese. We therefore conclude that, overall, the data are best explained by assuming that *both* the addressed and assembled routes are available, and are used in the reading of virtually all shallow orthographies. We turn now to the model.

An important feature of the model depicted in Figure 1 is that the various subsystems are *interconnected* and further, that processing is *cascaded.* More concretely, when a word is presented, spelling-sound correspondences at the subword level (pathway E) activate the phonemic buffer, which is engaged in interactive activation with the phonological output lexicon (pathways C and F). In turn, the phonological output system is engaged in interactive activation with the phonological input system (pathways J and G) which is also engaged in interactive activation with the semantic system (pathways H and I). Finally, the semantic system also has it own direct feedforward link to the phonological output system (pathway B). The important point is that even if access to the word recognition system occurs via the assembled route, lexical and semantic factors can still affect naming since the entire system is interactively engaged.

How might such a model account for the elimination of priming and word frequency effects when nonwords form part of the target set? There are two pieces of evidence to support the claim that the phonological output lexicon still plays a role. First, even in the complete absence of any priming effect, words are still named faster than nonwords (Baluch & Besner, 1991; Frost et al., 1987; Katz & Feldman, 1983). No difference between words and nonwords would be expected if only the phonemic buffer was involved. The second piece of evidence involves information about syllabic stress. If the phonological output system was not involved, word specific information about stress would be unavailable. These data imply that the phonological output lexicon must be involved in naming even if the assembled routine is used.

Although the phonological output lexicon is activated, the *absence* of priming effects when nonwords form part of the target set suggests that one or more of its connecting links are inhibited. One hypothesis is that the link from the phonological output lexicon to the phonological input system (route G) is inhibited, thereby eliminating activation of the semantic system (see Figure 1). Alternatively, input to the phonological output lexicon from the phonological input lexicon or from semantics (pathways J and B) might be inhibited. Either pattern of inhibition would allow for comprehension, yet fail to reveal an effect of context on naming time when there are nonwords in the target set.

b) Word frequency effects when using the assembled route. The model depicted in Figure 1 may again be used to account for the occurrence of word frequency effects when the initial input to the visual word recognition system occurs via the assembled routine. The first question is whether activation of the phonological output lexicon by the phonemic buffer without activation of any other lexical/semantic systems is sufficient to produce a word frequency effect. The answer to this is a clear "No," for several reasons. For example, McCann and Besner (1987) showed that nonwords that sound like real words (e.g., BRANE) are named faster than nonwords which do not sound like real words (e.g., FRANE). This finding implies lexical influence, a fact easily explained by interactive activation between the phonemic buffer and the phonological output lexicon. However, frequency of the real base word (e.g., BRAIN) did not predict naming time to the items which sound like words (e.g., BRANE) suggesting that the phonological output lexicon is not sensitive to word frequency. A second line of evidence which rules out the phonological output lexicon as the source of word frequency effects comes from naming studies which include nonwords as targets. Baluch and Besner (1991) (in Persian), Hudson and Bergman (1985) (in Dutch) and Frost et al. (Serbo-Croatian) all found that although words were named faster than nonwords (thereby implicating the phonological output lexicon), there was no word frequency effect. These considerations suggest that the phonological output lexicon is not the source of the word frequency effect seen in naming. Rather, processing must penetrate deeper into the word recognition system to produce the word frequency effect. As was the case with priming when the assembled route was employed, we suggest that interactive activation between the phonological output lexicon and other lexical/semantic systems is needed to account for word frequency effects in naming (see Besner, 1978; Besner & Smith, 1992; McCann & Besner, 1987; McCann, Besner, & Davelaar, 1988 for discussion of several locii). In the presence of nonwords, output of one or more of the systems is again inhibited.

Although both priming and frequency effects are thus explained in terms of interactive activation, there is an important difference in terms of the time-course of activation in these two effects. In the typical priming paradigm the prime and target are presented sequentially, and there is ample time for the relevant parts of the system to become activated prior to target presentation even if processing were not cascaded. In contrast, frequency effects in word naming must occur on-line. Consequently, it is necessary to assume that activation cascades forward and backwards in various parts of the system (cf. McClelland, 1979; 1987; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). Preparation for speech from the phonemic buffer must also be sufficiently slow so as to allow cascaded processing in the phonological output lexicon to activate other lexical systems and feed back down to the phonemic buffer.

Thus far, it appears that a revision of the orthographic depth hypothesis which permits a spread of activation from the phonemic buffer to various parts of the lexical/semantic system may be sufficient to account for both contextual priming and word frequency effects under conditions in which the initial input is restricted to the assembled route.

Note, however, that we have had to assume *between level inhibition* in order to account for the elimination of frequency and priming effects in the presence of nonwords as targets. This feature distinguishes the present model from that of the general interactive-activation framework postulated by McClelland and his colleagues, since in their framework inhibition only operates *within* levels. If the present problematic data are the only ones that require between level inhibition, this may constitute another reason for preferring an account in which the presence/absence of nonwords leads to strategic shifts in processing that emphasize either the addressed or assembled routes (cf. Baluch & Besner, 1991). Note also that the between level inhibition assumption presupposes some control structure to implement inhibition. Such control structures are a theoretical problem that for the most part has been ignored in the context of models of word recognition, probably because of the difficulty in describing one which is not a disguised homunculus.

We turn now to the third issue which we consider problematic for the orthographic depth hypothesis: orthographic familiarity effects at the whole word level when reading a shallow orthography.

c) Whole word orthographic familiarity effects in a shallow orthography. As described earlier, Besner and Hildebrandt (1987), Besner et al. (1992), and Buchanan and Besner (1992) all reported data which they interpreted as evidence for use of the addressed route in the oral reading of a shallow orthography. More specifically, Japanese Kana was read more efficiently when the words were printed in their familiar form than when they were transcribed into a form which, despite preserving phonology, was orthographically unfamiliar at the whole word level. Since both forms of Kana are perfectly regular in terms of their spelling-sound correspondences and could have been read via the assembled route, faster reading of the familiar form constitutes evidence for the involvement of the orthographic input lexicon. This evidence requires us to reject the orthographic depth hypothesis because the essence of it is the assertion that there is no orthographic input lexicon in the minds of readers of shallow orthographies.

A weaker form of the orthographic depth hypothesis might concede that there is indeed an orthographic input lexicon, but assert that it is not accessed *directly* by print. According to this argument print can only be processed via the assembled routine, and the orthographic input lexicon only becomes activated as a consequence of *this* processing. To use the interactive model described in Figure 1 to account for such activation, we would be forced to postulate that activation of the semantic system via the phonological output lexicon must in turn activate the orthographic input lexicon (pathway K). Again, processing must be heavily cascaded in order to allow output from the orthographic input lexicon to influence the phonological output lexicon (pathways D and/or A-B). Although this sequence of events would provide an account of the data, we can think of no rationale for postulating an orthographic input lexicon which can play a role in the reading of a shallow orthography, but which cannot be *directly* activated by print. Indeed, it involves postulating mental machinery whose only function is to prevent the formation of such direct associations. Why would the mind inhibit associations that could only be helpful to the reader? A more reasonable explanation is that even in a shallow orthography the orthographic input lexicon is directly activated by print at least some of the time, and that this process contributes to reading in shallow orthographies.

Conclusions

In conclusion, we suggest that overall the data are most compatible with the hypothesis that the addressed and the assembled routines are available in virtually all orthographies (see also Patterson, 1990 for similar conclusions). However, we assume that the assembled route makes more of a contribution to word recognition in reading shallow scripts, because the reliable relationship between spelling and sound supports faster computation. Finally, the differential occurrence of both priming and word frequency effects as a function of the presence or absence of nonwords in the target set suggests that readers have some "control"[8] over the mode of processing employed. How such control is achieved awaits further investigation.

Representational issues

As word recognition research matures, representational issues have become more central. Here, we offer some brief speculations concerning the current debate on the utility of distributed representations as they apply to visual word recognition. A very simple question is raised: are there any phenomena that we know about today which are problematic for currently implemented models which are totally distributed ?

---

[8] "Control" is in quotation marks because the term is often taken to imply conscious processing. We have not been able to think of a better term for the notion that processing is quite flexible in response to different stimulus environments, but we most assuredly do not wish to assume that this processing is necessarily tied to "consciousness." What we do want to suggest is that the speed at which these routes operate is not fixed. When the context favours a particular route, more "attention" is paid to that route. One way to implement such an idea is to suppose that "attention" modulates the rate of processing in one or more routes (cf. Baluch & Besner, 1991; Paap & Noel, 1991; Stelmach & Herdman, 1991).

Are parallel distributed processing models of visual word recognition sufficient ?

Until very recently there was a remarkable degree of consensus regarding the way in which knowledge is represented "in" the mental lexicon. Words had "local" mental representations. However, in some quarters there is currently a major conceptual shift towards a view which favors a more "distributed" account in which various kinds of knowledge about words are represented as patterns of activation across a large set of nodes at several different levels (e.g., Hinton & Shallice, 1990; Masson, 1991; Monsell, 1991; Plaut & Shallice, 1991; Seidenberg & McClelland, 1989; Van Orden, Pennington, & Stone, 1990). One consequence of this conceptual shift has been renewed debate as to whether the word recognition system has both a lexical and a nonlexical way of converting print into phonology. Some connectionists argue that there is no genuine distinction between lexical and nonlexical routes (e.g., see Seidenberg, 1989; Seidenberg & McClelland, 1989, 1990). This view is disputed by others (e.g., Baluch & Besner, 1991; Besner, 1990; Besner et al., 1990; Monsell et al., 1992; Paap & Noel, 1991; see also the section on neuropsychological evidence). What is clear is that there are still several empirical problems.

One problem concerns the reading of nonwords. To date, there is no purely distributed model of word recognition in English that pronounces new words (nonwords) with anything like the accuracy that normal subjects do (cf. Besner et al., 1990).

A second difficulty arises from the word frequency data reviewed in the present paper. A central assumption of distributed accounts seen to date is that they are, by definition, sensitive to word frequency.

> "The essence of connectionist learning models is that they learn patterns as a function of their experience with them, and their response to a pattern in a given representational domain is a function of the degree to which they have learned (a) the relationship between its parts, and/or (b) the mapping of it and its parts onto patterns in another domain. Words of high frequency differ from words of low frequency, all other things being equal, in their degree of learning or acquisition. Frequency effects are intrinsic to connectionist learning models."
>
> (Monsell, 1991, p. 155).

However, as discussed in this paper, the presence/absence of a word frequency effect in oral reading is dependent upon the presence/absence of nonwords as targets (Baluch & Besner, 1991; Frost et al., 1987; Hudson & Bergman, 1985; Wydell, 1991). Models such as those of Seidenberg and McClelland (1989, 1990) and Van Orden et al. (1990) are discomfited by such data for the simple reason that they do not have a route which is insensitive to word frequency.

Can this problem be dealt with by simply adding another route which is nonlexical in the sense that it is insensitive to word frequency? We believe not, since a central assumption in distributed accounts to date is that orthographic description at the subword level is intertwined with orthographic description at the whole word level. For example:

> "The pretheoretical distinctions between different types of stimuli are difficult to maintain because several different factors-overall frequency, word body frequency, regularity, orthographic redundancy-are typically confounded in the language. These natural confoundings are neatly handled in the model in terms of the aggregate effects of training on the settings of the weights on connections."
>
> (Seidenberg & McClelland, 1989, p. 546).

Given the claim that these two levels cannot be kept separate in such a framework, it would appear that a purely PDP approach cannot account for the absence of the word frequency effect when readers of shallow orthographies read words in the presence of nonwords.

Semantic level processing

Some PDP models have also been implemented to simulate semantic level processing (e.g., Hinton & Shallice, 1990; Masson, 1991; Plaut & Shallice, 1991). These models have attracted considerable interest, particularly in regard to the apparently successful modeling of semantic errors in deep dyslexia, and because of the suggestion that various kinds of priming effects in word recognition may be easier to reconcile with distributed as opposed to local representations. We merely note here, again, that an explicit account is needed of how the presence of nonwords in the target set modulates the presence/absence of priming effects in shallow orthographies.

A further issue which may yet prove problematic for distributed models of semantic processing is the handling of polysemous words such as "BANK" when they are presented without any accompanying context. We know from a number of studies that both meanings of the word will become activated. Since the different meanings of BANK are in competition with one another it might take longer for the semantic system to settle into a stable pattern of activation as compared to a word with only a single meaning.[9] However, the behavioral data show the opposite effect: polysemous words presented without any context are classified *faster*, than words with a single meaning in lexical decision (e.g., Millis & Buttons, 1989; Kellas et al., 1988; Jastrembski, 1981; see also Balota et al's, 1991 review). Similar results are observed in the naming task (Balota and Ferraro, 1992, personal communication). It is not immediately obvious how to deal with this problem.

Can any guesses be made as to where work on parallel distributed processing in visual word recognition is going? It might turn out that the most viable word recognition model will be a hybrid one which incorporates multiple routes and both distributed and local representations. Given what we currently know there is no evidence that a purely distributed account will suffice.

Concluding Comment

If the past century is any indication, research on visual word recognition will likely be pursued for quite some time to come. Our modest hope is that the review provided here has served to clarify some of the issues and will promote the search for universals in the processes associated with the reading of quite different orthographies.

Acknowledgment

The work reported here was supported by Grants A0998 to the first author and OGPIN012 to the second author from the Natural Sciences and Engineering Council of Canada. We thank Peter Taumofsky for preparing stimulus materials and testing the Croatian readers. This chapter was written while the first author was on leave at the University of Toronto and Carleton University;

---

[9] Indeed, Joordens and Besner (submitted) found that a Hopfield net, while successfully simulating a number of priming phenomena, fails to settle into any of the correct semantic patterns associated with a polysemous word at least 60% of the time. It will be interesting to see if a radically different architecture such as Hinton and Shallice's performs better.

we are grateful to these institutions for facilities, and in particular to Bob Lockhart for the use of his office.

References

Balota, D., & Chumbley, J. (1985). The locus of word frequency effects in the pronunciation task: Lexical Access and/or Production? *Journal of Memory & Language, 24,* 89-106.

Balota, D. A., Ferraro, F. R., & L. T. Conner (1991). On the early influence of meaning in word recognition. In P. Schwanenflugel (Ed.), *The psychology of word meanings.* NJ: Lawrence Erlbaum & Associates.

Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 644-652.

Besner, D. (1978). Output logogens and the locus of the word frequency effect. Unpublished manuscript.

Besner, D. (1987). On the relationship between orthographies and phonologies in visual word recognition. In D. A. Allport, D. MacKay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Shared mechanisms in listening, speaking, reading and writing* (pp. 211-226). San Diego, CA: Academic Press.

Besner, D. (1990). Does the word recognition system need a lexicon? In D. A. Balota, G. B. Flores d'Arcais & K. Rayner (Eds.), *Comprehension processes in reading.* Hillsdale, NJ: Lawrence Erlbaum & Associates.

Besner, D., & Davelaar, E. (1982). Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology, 36(4),* 701-711.

Besner, D., & Hildebrandt, N. (1987). Orthographic and phonological codes in the oral reading of Japanese Kana. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 335-343.

Besner, D., Patterson, K., Lee, L., & Hildebrandt, N. (in press). Two forms of Japanese Kana: Phonologically but NOT orthographically interchangeable. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Besner, D., Twilley, L., McCann, R .S., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review, 97,* 432-446.

Besner, D., & Smith. M. C. (in press). Models of visual word recognition: when obscuring the stimulus yields a clearer view. *Journal of Experimental Psychology: Learning, Memory & Cognition.*

Bridgeman, B. (1987). Is the dual-route theory possible in phonetically regular languages? *Behavioral and Brain Sciences, 10,* 331-332.

Brown, P., & Besner, D. (1987). Assembly of phonology in oral reading: A new model. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 471-489). Hillsdale, NJ: Lawrence Erlbaum Associates.

Buchanan, L., & Besner, D. (in preparation). Reading a shallow orthography: Lexical access in naming consequent to processing via the assembled route.

Butterworth, B., & Wengang, Y. (1991). The universality of two routines for reading: evidence from Chinese dyslexia. *Proceeding of the Royal Society, 245,* 91-95, London.

Campbell, R., & Besner, D. (1981). This and Thap—Constraints on the pronunciation of new, written words. *Quarterly Journal of Experimental Psychology, 33A, 4,* 375-396.

Carello, C., Lukatela, G., Turvey, M. T. (1988). Rapid naming is affected by association but not syntax. *Memory & Cognition, 16,* 187-195.

Coltheart, M. (1980). The semantic error: types and theories. In M. Coltheart, K. E. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia.* London: Routledge & Kegan Paul.

Coltheart, M., Patterson, K. E., & Marshall, J. C. (1980). *Deep dyslexia.* London: Routledge & Kegan Paul.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12,* 627-635.

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 361-379.

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.

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review, 98, 1,* 74-95.

Hudson, P. T. W., & Bergman, M. W. (1985). Lexical knowledge in word recognition: word length & word length in naming and lexical decision tasks. *Journal of Memory and Language, 24,* 46-58.

Humphreys, G., & Evett, L. (1985). Are there independent lexical & nonlexical routes in word processing? An evaluation of the dual-route theory of reading. *The Behavioural and Brain Sciences, 8,* 689-740.

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology, 13,* 278-305.

Joordens, S., & Besner, D. (submitted). *Constraining distributed models of semantic memory: Gaining clarity by examining ambiguity.*

Katz, L., & Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 157-166.

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the time-course of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 601-609.

Khanlari, P. N. (1979). *The history of Persian language, Volume 1.* New Delhi: New Delhi Press.

Lukatela, G., & Turvey, M. T. (1991). Phonological access of the lexicon: evidence from associative priming with pseudohomophones. *Journal of Experimental Psychology: Human Perception and Performance, 17(4),* 951-966.

Maddison, I. (1984). *Patterns of sounds.* Cambridge, England: Cambridge University Press.

Marcel, A. J. (1980). Surface dyslexia and beginning reading : a revised hypothesis of the pronunciation of print and its impairments. In M. Coltheart, K. E. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia,* London: Routledge and Kegan, Paul

Masson, M. E. J. (1991). A distributed memory model of context effects in word identification. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading.* Hillsdale, NJ: Lawrence Erlbaum & Associates.

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 13-24.

McCann, R. S., Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do word frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance, 14,* 693-706.

McClelland, J. L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading.* Hillsdale, NJ: Lawrence Erlbaum Associates.

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.

Millis, M. L., & Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory and Cognition, 17,* 141-147.

Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 148-197). Hillsdale, NJ: Lawrence Erlbaum Associates.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). The effects of frequency upon visual word recognition: Where are they? *Journal of Experimental Psychology: General, 118,* 43-71.

Monsell, S., Patterson, K. E., Graham, A., Hughes, C. H., & Milroy, R. (in press). Lexical and sub-lexical translation of spelling to sound: strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory & Cognition.*

Morton, J., & Patterson, K. (1980). A new attempt at an interpretation, or, an attempt at a new interpretation. In M. Coltheart, K. E. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia.* Routledge and Kegan Paul.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Paap, K., & Noel, R. (1991). Dual route models of print to sound. Still a good horse race. *Psychological Research, 53,* 13-24.

Patterson, K. E. (1982). The relation between reading and phonological coding: Further neuropsychological observation. In A. Ellis (Ed.), *Normality and pathology in cognitive functions.* Academic Press.

Patterson, K. E. (1990). Basic processes of reading: Do they differ in Japanese and English? *Japanese Journal of Neuropsychology, 6(1),* 4-14.

Patterson, K. E., & Coltheart, V. (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Patterson, K. E., Marshall, J. C., & Coltheart, M. (1985). Surface dyslexia. Hillsdale, NJ: Lawrence Erlbaum & Associates.

Plaut, D., & Shallice, T. (1991). Deep dyslexia: a case study of connectionist neuropsychology. *Technical Report CRG-TR-91-3,* Dept. of Computer Science, University of Toronto.

Rosson, M. B. (1983). From SOFA to LOUCH : Lexical contributions to pseudoword pronunciation. *Memory and Cognition, 11(2),* 152-160.
Rosson, M. B. (1985). The interaction of pronunciation rules and lexical representations in reading aloud. *Memory and Cognition, 13(1),* 90-98.
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.
Scarborough, D. L., Cortese, C., & Scarborough, H .L. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 1-17.
Sebastián-Gallés, N. (1991). Reading by analogy in a shallow orthography. *Journal of Experimental Psychology: Human Perception & Performance, 17, 2,* 471-477.
Seidenberg, M. S. (1989). Visual word recognition and pronunciation: A computational model and its implications. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 25-74). Cambridge, MA: MIT Press.
Seidenberg, M. S., & McClelland, J. L . (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523-568.
Seidenberg, M. S., & McClelland, J. L. (1990). More words but still no lexicon: A reply to Besner et al. (1990). *Psychological Review, 97,* 447-452.
Seidenberg, M. S., & Vidanovic, S. (1985, November). *Word recognition in Serbo-Croatian and English: Do they differ?* Paper presented at the 20th annual meeting of the Psychonomic Society, Boston.
Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tannehaus, M. (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. (1980). Single and multiple component central dyslexic syndromes. In M. Coltheart, K. E. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia.* London: Routledge and Kegan Paul.
Stimach, L. B., & Herdman, C. M. (1991). Directed attention and perception of temporal order. *Journal of Experimental Psychology: Human Perception & Performance, 17(2),* 539-550.
Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory and Cognition, 20,* 315-328.
Turvey, M. T., Feldman, L. B., & Lukatela, G. (1984). The Serbo-Croatian orthography constrains the reader to a phonologically analytic strategy. In L. Henderson (Ed.), *Orthographies and reading. Perspectives from cognitive psychology, neuropsychology, and linguistics* (pp. 81-89). Hillsdale, NJ: Lawrence Erlbaum Associates.
Van Orden, G. C., Pennington, B.F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97,* 488-522.
Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: time course and decision criteria. *Memory and Cognition, 13(6),* 557-572.
Wydell, T. (1991). Unpublished doctoral dissertation. Birkbeck College, University of London.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 4

The Reading Process is Different for Different Orthographies: The Orthographic Depth Hypothesis

Leonard Katz
Department of Psychology
University of Connecticut

Ram Frost
Department of Psychology
Hebrew University

It has been said that most languages get the orthography they deserve and there is a kernel of truth in that statement. There is generally an underlying rationale of efficiency in matching a language's characteristic phonology and morphology to a written form. Although the final product may turn out to be more suitable for some languages than for others, there are certain basic principles of the fit that can be observed. The attempt to make an efficient match between the written form, on the one hand, and morphology and phonology, on the other, typically determines whether the orthography chosen is a syllabary, a syllabary-cum-logography, or an alphabet. Further, within the group of alphabetic orthographies itself, there are varying degrees of dependence on the strict alphabetic principle: the range of correspondence between grapheme and phoneme varies both in consistency and completeness. The degree of this dependence is to some extent a function of a language's characteristic phonology and morphology, just as was the choice of the kind of orthography itself. We discuss here what this varying dependence on the alphabetic principle may mean for the mental processes involved in reading and writing.

Diversity in writing systems

Although writing systems are, in general terms, systems for communication, what they actually communicate is the spoken language—as opposed to communicating nonverbal ideas and meanings. DeFrancis (1989) reinforced this point with his analysis of so-called pictographic languages, writing systems whose elements are pictures and symbols that do not stand for words. DeFrancis argued that true pictographic systems are not, in principle, effective and showed that existing examples of pictographic systems had been designed only as novelties or playful communication systems. In practice, they were never used to communicate without substantial ancillary aid from spoken language. The example of pictographic writing emphasizes the poverty of written communication that is not based on language. Therefore, because writing systems are systems for representing the spoken language, it is reasonable to suggest that an understanding of the psychological processing involved in using a writing system must include an understanding of the processing of the spoken language. Although the former will not necessarily parallel the latter, it will be constrained by it. A major constraint arises from the spoken language's morphemes which are the smallest units that carry meaning. It is these morphemes of speech that will be the focus of communication, both spoken and written. Word stems are all morphemes and so are their derivational and inflectional affixes; these units of the spoken language must be easily recoverable from the written language.

A large number and variety of writing systems have flourished, evolved and developed, and in many cases, died, over the centuries. Each of the known systems can be categorized as either logographic-phonetic, syllabic, or alphabetic (DeFrancis, 1989). These distinctions are made on the basis of how a script (a set of symbols) relates to the structure of its language. This relationship between a script and its language is what is described by the term orthography (Scheerer, 1986). The kind of script system and its orthography are typically not wholly the result of accident. It is not accidental that the Chinese languages, for example, have a logographic-phonetic system. In the Chinese orthography, the typical character has two parts to it: a logographic and a phonetic component, the former providing a visually distinctive cue to the semantics and the latter giving the reader a partial guide to the pronunciation. Together, the two components make a combination that specifies completely a unique spoken morpheme. Words may be mono- or polymorphemic.

Chinese morphemes are mainly monosyllabic and, because the variety of possible syllables is limited, there is a high degree of homophony in the language. Such a language is best served by an orthography that distinguishes between the different meanings of morphemes that sound alike. Instead, if the orthography had represented the spoken form alone (e.g., only the phonetic component in the printed word), the reader would not be able to determine the intended meaning of each homophone except, possibly, from the word or sentential context (many polymorphemic words are unique compounds of homophonous morphemes)—but not without the additional cognitive cost needed to resolve the ambiguity. The homophony problem is more of a problem for a reader than for a listener because the listener has more nonverbal contextual information available to assist in determining word meanings.

But a pure logography would not suffice either, for Chinese. If the orthography had no phonetic component, a reader would have to remember, without a phonetic cue, a pronunciation for each of several thousand logographs. This would have effectively limited the number of printed characters that a reader could remember and name to an unacceptably small number. Instead, in modern Mandarin, it is necessary to remember only a small number of phonetic components together with a smaller number of semantic signs. DeFrancis gives the number of these as 895 and 214, respectively (DeFrancis, 1989). A phonetic and a semantic component are paired to produce a character; the effective set consists of 4300 characters, the approximate number considered necessary for full literacy.

In contrast to Chinese, spoken Japanese is polysyllabic and is composed of regular syllable-like components, called moras. Because the number of syllables is small (fewer than 113), it is feasible to represent them by means of a syllabary. The Japanese orthography called kana was such a system, adapted from Chinese characters. However, because there is a good deal of homophony in Japanese, the use of a syllabary alone would not have been without problems and a logography also came into use. That logography is still in use today and is routinely mixed with the use of the kana, the syllabaries being used primarily for morphological affixes and grammatical function words, foreign loan-words, and words not already covered by the Chinese.

Indo-European languages have less homophony and more polysyllabic morphemes than Chinese and Japanese. In addition, the structure of the Indo-European syllable itself is generally more complex than Chinese or Japanese, containing a larger number of phonologically permissible clusters. English is said to have at least 8000 syllables in its phonology, compared to fewer than 1300 for Chinese (DeFrancis 1989). Eight-thousand is far too large a number for an effective syllabary. For English, an alphabet, representing phonemes, is more efficient for learning to read and write. Similarly, the Semitic languages (which include Arabic and Hebrew) would be less suitably represented by a syllabary than by an alphabet. Like Indo-European languages, they too have complex syllable structures. Historically, a consonantal alphabet that developed for West Semitic was the alphabet from which we trace the evolution of modern alphabets.

One of the characteristics of Semitic languages that may have led to the invention of the alphabet is the Semitic triconsonantal root: Semitic words that are related by derivation or inflection have a common core (usually three consonants). Although the vowels in each of the different relatives may be quite different and although there may be additional consonants in a prefix or suffix, there remains an invariant series of three consonants, the root. This core has a strong linguistic salience—it represents a morphological communality of meaning among its family members (see Frost & Bentin, this volume). One can speculate that several early attempts were made at developing a writing system but only an alphabetic system could have captured this communality of morphology efficiently. Syllabic representations would not be optimal: in Hebrew, morpheme boundaries fail to coincide with syllable boundaries (a condition that is true of Indo-European languages as well). Therefore, syllabic representations would not be appropriate for representing morphological units.

The causes of diversity in alphabetic orthographies

Even among the various alphabetic writing systems themselves there are major differences in the degree to which they mirror the phonemic structure of their respective spoken languages. Again, the reason for the differences is largely accounted for by the particular phonological and morphological characteristics of each language. For example, standard written Hebrew is an orthography in which all diacritics (or *points*) are omitted. These diacritics represent nearly all of the vowels and are also used to disambiguate some of the consonants. Nevertheless, writing without diacritics is usually sufficient to indicate the exact intended (i.e., spoken) word if it is supported by a phrasal or sentential context. Thus, although the printed root may be insufficient to allow an unequivocal identification when presented in isolation, when it is presented in a normal context—even a printed one—the combined sources of information are enough for word identification.

In strong contrast to the Hebrew orthography is the Serbo-Croatian. Serbo-Croatian is a major language of the Balkan Peninsula. Its present alphabet was introduced in the early nineteenth century following the principle, "Spell a word like it sounds and speak it the way it is spelled." Each letter represents only one phoneme and each phoneme is represented by only one letter. Moreover, no phoneme in the spoken word is ever excluded in the spelled word. The relation between letters and phonemes is isomorphic and exhaustive. To this day, the Serbo-Croatian spelling system follows the phonemic structure of spoken words. So regular is the relation between speech and writing that the writings of people with different regional pronunciations will show different spellings, mirroring the minor differences in their spoken language. This simple writing system works well for Serbo-Croatian because morphemic variations in the language due to inflection and derivation do not often produce alterations in the phonemic structure of word stems; word stems largely remain intact.

A different state of affairs exists in English. English is somewhere between Hebrew and Serbo-Croatian in the directness with which its phonology is represented in its spelling; there is a large amount of regular phonologic change among words in the same derivational family. Examples of this are the contrasts between the derivational relatives HEAL and HEALTH, between STEAL and STEALTH, etc. Chomsky and Halle (1968) argue that English spelling represents a morphophonemic invariance common to these word pairs, an abstract phonological communality that is below their surface difference in pronunciation. In reading English aloud, the reader must either remember the pronunciation of such a word as a whole or remember the appropriate context-dependent rule for pronunciation. An alternative writing system might have spelled English in the same way that Serbo-Croatian is spelled: with an isomorphic relation between letter and phoneme. However, that method would rob printed English of the advantageous common spelling for words with common morphology. The words HEAL and HEALTH, for example, might then be spelled HEEL and HELTH, disguising their common meaning. The printed form HEEL would also suffer from a double meaning as a consequence of its being homophonic with the word meaning, "part of a foot." English spelling represents a compromise between the attempt to maintain a consistent letter-phoneme relation and the attempt to represent morphological communality among words even at the cost of inconsistency in the letter-phoneme relation.

Thus, alphabetic writing systems reflect the spoken forms of their respective languages with different degrees of consistency and completeness between letter and phoneme. Some of the differences in writing systems have purely political, cultural, or economic causes. But many differences have been motivated by two factors that are purely linguistic. The first has to do with how complex the spoken language is in the relation between phonology and morphology—only a phonologically complex language can have a deep alphabetic orthography. For example, Serbo-Croatian is not phonologically complex. All morphologically related words have a common phonologically invariant core. Two words that are morphologically related will share a common word stem that will necessarily sound the same in both words. Both instances of that common stem will, of course, be spelled the same. Thus, when evaluated by the characteristic of phonological complexity, a language that is not complex can be written (and generally *will* be written) in a shallow orthography, an orthography that tracks the phonology. Secondly, if a language is one that is phonologically complex then the orthography has the option of representing either morphological invariance (a deep orthography) or following grapheme-phoneme invariance (a shallow orthography). As we suggested above, English qualifies as quite complex, phonologically. In principle, it could have been written either as a shallow or a deep orthography. The advantage to English in choosing a deep orthography is in the consistent spelling of morphemic invariances. However, that choice having been made, there are then different pronunciations of the same spelling on occasion (e.g., HEAL-HEALTH) and, inadvertently, identical pronunciations for different spellings (e.g., PEEL-DEAL).

A different situation exists in Hebrew. Hebrew's phonology is complex; morphemes may undergo considerable sound change under either inflectional or derivational change. On the other hand, because of the pervasiveness of the triconsonantal root in Hebrew, a great deal of morphological constancy exists. Therefore, there was an historical choice, so to speak, for the evolution of the Hebrew orthography: It could have opted for either morphemic or phonemic invariance but, unlike Serbo-Croatian, it could not have contained both in a single orthography because of its phonological complexity. Hebrew initially evolved as an orthography in which the morphology was preserved at the expense of phonological completeness. Vowels were omitted thereby emphasizing the morphologically based consonantal invariance in a given family of word roots. Vowel points were added to the script at a later stage in the orthography's development only because the language was no longer being spoken as a primary language and it was feared that its pronunciation would become corrupted unless vowels were included in the script. Nowadays, the orthography used by adults is the unpointed one, which is graphemically incomplete and somewhat inconsistent to the reader because it omits nearly all of the vowels and makes some of the consonants ambiguous.

In summary, all alphabetic orthographies can be classified according to the transparency of their letter-to-phoneme correspondence, a factor that has been referred to as orthographic depth (Liberman, Liberman, Mattingly, & Shankweiler, 1980). An orthography in which the letters are isomorphic to phonemes in the spoken word (completely and consistently), is orthographically shallow. An orthography in which the letter-phoneme relation is substantially equivocal is said to be deep (e.g., some letters have more than one sound and some phonemes can be written in more than one way or are not represented in the orthography). Shallow orthographies are characteristic of languages in which morphemic relatives have consistent pronunciations.

Differences among alphabetic orthographies in processing printed words: The orthographic depth hypothesis

Our discussion to this point has made the standard argument that there are differences among alphabetic orthographies in orthographic depth and that these differences are a result of differences in their languages' phonology and morphology. In this section, we propose that the differences in orthographic depth lead to processing differences for naming and lexical decision. This proposal is referred to as the orthographic depth hypothesis (ODH). It states that shallow orthographies are more easily able to support a word recognition process that involves the language's phonology. In contrast, deep orthographies encourage a reader to process printed words by referring to their morphology via the printed word's visual-orthographic structure.

We would like to make two points, each independent of the other. The first states that, because shallow orthographies are optimized for assembling phonology from a word's component letters, phonology is more easily available to the reader prelexically than is the case for a deep orthography. The second states that the easier it is to obtain prelexical phonology, the more likely it will be used for both pronunciation and lexical access. Both statements together suggest that the use of assembled phonology should be more prevalent when reading a shallow orthography than when reading a deep orthography. Because shallow orthographies have relatively simple, consistent, and complete connections between letter and phoneme, it is easier for readers to recover more of a printed word's phonology prelexically by assembling it from letter-phoneme correspondences.

Suggested by the above is our assumption that there will always be at least some dependence on phonological coding for the process of reading in any orthography. That is, the processing of (at least) some words will include assembled phonology (at least in part). This assumption can be easily motivated for alphabetic orthographies. The assembling of phonology has a certain precedence in a reader's experience; instruction in reading typically means instruction in decoding, i.e., learning how to use letter-phoneme correspondences. It is well established that beginning readers find it easier to learn to read in shallow orthographies, where those correspondences are most consistent (see, for example, Cossu, Shankweiler, Liberman, Katz, & Tola, 1988). Even in learning to read Hebrew, instruction is typically given in the shallow pointed orthography instead of the deep unpointed one (the transition to the unpointed form beginning in the third grade). In any orthography, after learning to read by using assembled phonology routines, skilled readers may continue its use to the extent that the cost of doing so is low. This will be particularly true when the orthography is shallow. However, given the experimental evidence, some of which we discuss later, it seems certain that assembled phonology is not used *exclusively*. More likely, a mix of both assembled phonology and visual-orthographic codings are nearly always involved, even in shallow orthographies.

Two versions of the orthographic depth hypothesis

Two versions of the orthographic depth hypothesis (ODH) exist in the current literature. What can be called the *strong ODH* states that phonological representations derived from assembled phonology alone are sufficient for naming and lexical decision in shallow orthographies. Thus, according to the strong ODH, rapid naming in shallow orthographies is a result of only this prelexical analytic process and does not involve pronunciation obtained from memory, i.e., the lexicon. However, we submit that the strong form of the ODH is patently untenable when applied to the orthographies that have typically been used in research on word perception. It is insufficient to account for pronunciation even in a shallow orthography like Serbo-Croatian. This is so because Serbo-Croatian does not represent syllable stress and, even though stress is often predictable, it is not always predictable. Because the final syllable is never stressed, stress is completely predictable for two-syllable words but for words of more than two, it is not. However, one- and two-syllable words make up a large part of normal running text so much or most of the words a reader encounters can be pronounced by means of a prelexical subword analysis. But, of course, many words will be greater than two syllables in length and these can be pronounced correctly only by reference to lexically stored information. In addition, there are some exceptions to the rule that a letter must represent only one phoneme; some final consonant voicing changes occur in speech that are not mirrored in the conventional spelling (these changes are predictable, however). Thus, Serbo-Croatian, although it should be considered an essentially shallow orthography, is not the perfect paradigm of a shallow orthography. We should not expect a strong ODH to make sense for such an orthography.

We support the *weak ODH*. In this version, the phonology needed for the pronunciation of printed words comes not only from prelexical letter-phonology correspondences but also from stored lexical phonology, that is to say, from memory. The latter is the result of a visual-orthographic addressing of lexicon, i.e., a search process that matches the spelling of a whole word or morpheme with its stored phonology. The degree to which a prelexical process is active in naming is a function of an orthography's depth; prelexical analytic processes will be more functional (less costly) in shallow orthographies. However, whether or not these prelexical processes actually dominate orthographic processing for any particular orthography is a question of the demands the two kinds of processes make on the reader's processing resources, a question we discuss further below. We proposed (and supported) this weak form of the ODH in Katz and Feldman (1983) and Frost, Katz, and Bentin (1987); further details are given later in this chapter.

With regard to word recognition (as in lexical decision), some of our colleagues have argued that Serbo-Croatian necessarily involves prelexical (i.e., assembled) phonology (Feldman & Turvey, 1983; Lukatela & Turvey, 1990a). Others have made a similar claim for the obligatory involvement of prelexical phonology in English (Van Orden, Pennington & Stone, 1990; Perfetti, Bell, & Delaney, 1988). However, these researchers have not argued for the *exclusive* involvement of assembled phonology. Logically, assembled prelexical phonological information, without syllable stress information, is sufficient to identify the great majority of words in the English lexicon. However, irregularly spelled words, foreign borrowings, etc., would pose problems for an exclusively phonological mechanism, and, therefore, such a view seems less plausible. Finally, note that we are confining this discussion to the problems of naming and lexical decision; it is an entirely different question to ask whether phonological representations are necessary for *postlexical* processes like syntactic parsing and text comprehension.

Evidence on the questions of phonological recoding and the weak ODH

We discuss next the evidence for the hypotheses that the lexicon is addressed by assembled phonology, presumably in combination with visual-orthographic addressing, and that the specific mix of the two types of codes depends on orthographic depth. We show why single-language studies, in general, are not suitable for testing the weak ODH and mention the few exceptions. Experiments that directly compare orthographies with each other provide the most direct evidence. We will argue that these cross-language comparisons are absolutely critical to an investigation of orthographic depth effects.

It is important to realize that, in the controversy over whether visual-orthographic recoding or phonological recoding is used in word perception, there is little direct evidence of visual-orthographic effects. Instead, the burden of proof is placed on assembled phonology; if no effect of phonology is found, then visual-orthographic coding is said to win, by default. The assumption is not unreasonable because a visual-orthographic representation is obviously available in principle and seems to be the only alternative to assembled phonology. However, it should be kept in mind that, because of this, the experimental evidence hinges on the sensitivity of the experimental task and its dependent measures to phonology. If they fail to indicate the presence of phonology, it may not be because phonology is not operative.

In fact, several experiments have demonstrated that phonological recoding effects can be found even in deep orthographies. In Hebrew, Frost has shown that, if available, the full phonology given by the pointed orthography is preferred for naming; this seems to suggest that even if word recognition does not normally proceed via assembled phonology in Hebrew, the recognition process is prepared to default to its use (Frost, forthcoming). In English, Perfetti and his associates and Van Orden and his associates have presented strong evidence for phonological recoding in lexical access (Perfetti, Bell, & Delaney, 1988, Perfetti, this volume; Van Orden, Pennington, & Stone 1990; Van Orden, this volume).

However, it is one thing to find the active presence of phonological recoding, another to determine the conditions under which phonological recoding occurs, and yet another to determine the degree to which naming or lexical access is dependent on it. Is assembled phonology obligatory or is it rarely used; if neither of these extremes, is it the more preferred or the less preferred code? In this vein, Seidenberg (Seidenberg, 1985; Seidenberg, this volume) has argued that word frequency is the primary factor that determines whether or not assembled phonology is used to access the lexicon. His argument is that in any orthography, whether deep or shallow, frequently seen words will become familiar visual-orthographic patterns and, therefore, rapid visual access will occur before the (presumably) slower phonological code can be assembled from the print. Low frequency words, being less familiar, produce visual-orthographic representations that are less functional; lexical activation builds up more slowly. This gives time for phonological recoding to contact the lexicon first. However, we cannot presently answer questions that are concerned with the relative importance of word frequency to orthographic depth or concerned with the relative dominance of the two kinds of representation, visual-orthographic and phonological. But we can meaningfully address the question of whether or not the relative amount of assembled phonological coding decreases with increasing orthographic depth: the orthographic depth hypothesis.

Comparisons across orthographies

Cross-language experimentation, in which different languages are directly compared, are the critical methodology for studying the orthographic depth hypothesis. Single-language experiments are adequate for testing only the strong form of the ODH, in which shallow orthographies are said to never use lexically stored information for naming—but, as we showed, this is a claim that can be rejected on logical grounds. Single-language experiments are not without interest, however; they can be useful in indicating how easy it is to find effects of phonological coding. This may suggest—but only suggest weakly—what the dominant representation is for an orthography. If it is easy to find effects of phonological coding in Serbo-Croatian and difficult to find those effects in Hebrew, using more or less similar experimental techniques, we may suspect that phonological coding is the dominant (preferred) code in Serbo-Croatian but not in Hebrew. However, such experiments can not rule out the additional use of the alternate type of representation for either orthography. In this vein, we know that it is difficult to find effects of phonological coding in English using standard lexical decision paradigms but, nevertheless, phonological effects can be found using Perfetti's backward masking paradigm, which is apparently more sensitive (Perfetti, 1988). Finally, however, an accurate answer to the question of which type of representation is dominant for a particular orthography can not be given by the current experimental paradigms; the results from these experiments may only reflect the adequacy of the paradigms in capturing the true word recognition process.

The weak form of the ODH proposes that (1) both orthographic information and prelexically assembled phonological information are used for lexical access and (2) the degree to which one kind of information predominates is a function of the structural relationship between orthography and the lexical entry. The Serbo-Croatian orthography, with its simple and consistent letter-phoneme relationships, makes it easy for the reader to learn and maintain the use of assembled phonology. This assembled phonology must address, presumably, the same abstract phonology addressed by a listener's spoken language lexicon (although it may be only a subset of the full spoken phonology, because of the absence of stress information, at the least). In contrast, the Hebrew orthography, because it lacks most of the vowels and has many ambiguous consonants, is incapable of providing enough assembled phonology that will consistently identify a unique word in the phonological lexicon (only the consonants can be assembled); therefore, there are fewer benefits in generating phonological information by assembling it from grapheme-phoneme correspondences. These are lessons that the developing reader can learn tacitly, lessons that may lead, eventually, to different dominant modes of printed word processing for Serbo-Croatian and Hebrew readers. For languages that are in between these two extremes, the relative balance of assembled phonology to orthographic representation should reflect the relative efficacy of the two kinds of information in that orthography; some letters or letter-sequences may be simple and consistent and the assembled phonology derived from these may be used along with orthographic information. It is not possible to make a more precise statement without an understanding of the details of the lexical recognition process and the processing resources that are required. For example, for a visual-orthographically coded word to be recognized, a mental representation of that word must have been created previously and stored in the reader's memory. We do not know how to compare the resources needed to create a new orthographic representation with the resources needed to generate assembled phonology; which is more demanding? Neither can we automatically assume that it is easier to access lexicon via a visual-orthographic representation.

Additional complications arise when we try to be more specific about the phonological nature of the information in the lexicon itself; what, exactly, is the information that is represented in lexicon that is, presumably, addressed by assembled phonology? Alphabets mainly represent phonemes but are words in the spoken lexicon to be represented as phoneme sequences? If so, why do syllabic orthographies work at all, since the printed units of syllabaries map onto syllables, not phonemes? In fact, there are several different theoretical descriptions that have been proposed by speech researchers for the lexical representation of a word. Perhaps, the spoken lexicon contains multiple phonological descriptions of a single word, e.g., phonetic, phonemic, syllabic, gestural, etc. The phonology produced by reading may be a subset of the information constituting the spoken lexicon or it may even be different in kind (although related). We do not propose to discuss this problem in detail here, but only wish to point out that there is a companion to the question of how phonological information is used in reading, namely, the question of the nature of the phonological information that is used in spoken word recognition.

Evidence supporting the orthographic depth hypothesis

Some early evidence that there is a relationship between orthographic depth and lexical access was obtained by Katz and Feldman (1981). They compared lexical decision times in Serbo-Croatian and English for printed stimuli that were divided with a slash character. The stimuli were divided either at a syllable boundary (e.g., WA/TER) or one character to either the left or right of the boundary (e.g., W/ATER or WAT/ER). If word recognition involves recoding of the stimulus to a phonological form, and if that phonology includes the syllable as a unit, then division at a syllable boundary—which preserves the syllable units—should be less disruptive than division off the boundary. Pseudowords were similarly divided. Lexical decisions to words and pseudowords that were irregularly divided were slower than lexical decisions to their regularly divided counterparts and the disruptive effect of irregular division was stronger for Serbo-Croatian. The data, then, were consistent with a model of word recognition that assumes the operation of at least some phonological recoding of print prior to lexical access and, further, assumes that phonological recoding is a more consistent determiner of access in shallower orthographies (e.g., Serbo-Croatian) than deeper ones (e.g., English).

In a second study, Katz and Feldman (1983) made a direct test of a second prediction of the orthographic depth hypothesis: that pronouncing a word (naming) depends more on assembled phonology in Serbo-Croatian than in English. A lexical decision experiment and a naming experiment were run in both languages; stimulus words had the same (or similar) meanings in both languages. The subjects were native speakers who were tested in their native countries (Serbia or the United States), in order to avoid subjects who were fluent in both languages. This was done because the experience of a bilingual speaker might have affected his or her strategy for reading.

Each test stimulus (the target), whether it was a word or a nonword, was always preceded by the brief (600 ms) presentation of a real word. On half of those trials when the target was a word, this predecessor was semantically related to the target (e.g., MUSIC-JAZZ); on the other half, it was unrelated. Words also preceded the nonword targets. It is well established, that preceding a stimulus with a semantically related word will facilitate and speed a lexical decision response to the target. Thus, it was expected that reaction time to the target JAZZ would be faster for those subjects who saw it preceded by the word MUSIC than for those subjects who saw it preceded by the word GLASS, which has no strong semantic relation to the target. The critical fact for this kind of experimental technique is that the facilitating effect of MUSIC can only occur by activating the semantic link between it and JAZZ and this linkage necessarily must be within the lexicon. Thus, to the extent that there is facilitation in the subject's recognition of JAZZ, it indicates that recognition is being assisted by activity within the lexicon. Such lexical activity may facilitate whether the lexicon is addressed by orthography or by phonology.

For naming, however, the prediction is different. Naming can be accomplished largely without accessing the lexicon by means of subword letter-to-phonology recoding. Of course, in neither English or Serbo-Croatian can the process be entirely without reference to lexical memory, because the stress of polysyllabic words is not specified in the orthography. Nevertheless, the process of naming can, in principle, be carried out substantially without reference to the lexicon. Thus, if naming in Serbo-Croatian is more dependent on phonological recoding (and less dependent on lexical look-up) than English, naming in Serbo-Croatian ought not to be affected by the semantic priming manipulation, which is necessarily lexical in its locus of operation. Results supported this suggestion: target words that were preceded by semantically related words (e.g., MUSIC-JAZZ) were pronounced faster than target words that were preceded by unrelated words (e.g., GLASS-JAZZ) in the case of English but not in the case of Serbo-Croatian. In contrast, there were equivalent strong effects of semantic priming for lexical decisions, in both languages.

A three-way comparison of Hebrew, Serbo-Croatian, and English increased the range of orthographic depth examined in a single study (Frost, Katz, & Bentin 1987). Necessary to the success of any cross-language experiment is a set of stimulus words that have equivalent critical characteristics in all the languages under consideration: for example, equivalent frequencies of occurrence. Subjectively estimated frequencies of occurrence were obtained and equivalent sets of stimulus words were used in all three test languages. The important comparison, over the three orthographies, was the relation between naming and lexical decision reaction times for words versus nonwords. The importance of this comparison was based on the rationale that in shallow orthographies the lexicon plays only a minor role in the naming process compared to its role in the lexical decision process. The opposite assumption, i.e., that even in shallow orthographies, a skilled reader always employs the orthographic route to accessing the lexicon, predicts that readers in all three orthographies should perform similarly. On the other hand, if the orthographic depth hypothesis is correct, the greatest difference between naming and lexical decision reaction times should be in Serbo-Croatian, which has the shallowest orthography while Hebrew should show the greatest similarity. Results were in line with the orthographic depth hypothesis; naming times were considerably faster than lexical decision times in Serbo-Croatian but, in Hebrew, lexical decision and naming looked quite similar. In Hebrew, it took as long to name a word as to recognize it: a suggestion that naming was accomplished postlexically. In addition, in Serbo-Croatian, the faster responding for naming versus lexical decision was even greater for pseudowords than for words. In these comparisons, English was intermediate. Thus, the results support the hypothesis that the shallower the orthography, the greater the amount of phonological recoding that is carried out in naming. Subsequent experiments in this study, which maximized the potential for lexical processing by semantically priming target words and by varying the relative number of pseudowords further supported this interpretation.

In all the experiments we have discussed to this point, the experimental paradigms used have been naming and lexical decision tasks. These tasks have a disadvantage as methodologies because the phonological variation that is used to affect the subject's response (e.g., the consistency of the grapheme - phoneme relation) is obtained through manipulating the orthography (e.g., different alphabets) and not by manipulating the putative phonology directly. The experimenter never observes any phonologic recoding; its presence is only inferred. Thus, one can not be certain that the differences that are observed are true effects of phonological recoding or, instead, are only the result of orthographic effects which happen to be correlated with phonology. Frost and Katz (1989) addressed this issue by introducing a paradigm in which subjects had to compare a spoken word and a printed word. This paradigm requires subjects to perceive and use phonology in their task processing. Subjects were required to simultaneously read and listen to two words presented by computer and judge whether or not they were the same (i.e., represented the same lexical item). In order to make the comparison, the subject had to mentally place both spoken and printed stimuli into a common representation. This could have been done, in principle, in several ways, although only two possibilities seemed reasonable. The spoken word could have been imagined as a spelled (printed) word or subjects could have generated the phonology of the printed word. The evidence indicated that subjects chose the latter. This was not surprising: subjects have had far more practice reading than spelling. After converting the printed stimulus to a phonological representation, both phonological representations could then have been compared in order to determine if they matched. Over a list of 144 or more trials, subjects made the judgment "Same" or "Different" about each pair of printed and spoken words on each trial. There were three conditions: clear speech and clear print, degraded speech (noise added) and clear print, and clear speech and degraded print (visual noise added). Serbo-Croatian and English native speakers were tested on comparable materials. The effects of degrading were marked; when either the print or the spoken word was degraded, performance declined sharply. However, the difference in latency between the slower responses to print or speech that had been degraded compared to clear print or speech was four times greater in the orthographically deep English than the shallower Serbo-Croatian; degradation had a much stronger deleterious effect in English.

An interactive activation network model can be extended easily to account for these results. The model contains parallel orthographic and phonologic systems that are each multilevel with lateral connections between the two systems at every level. In particular, the sets of graphemic and phonemic nodes are connected together in a manner that reflects the correspondences in a particular orthography: mainly isomorphic connections in a shallow orthography and more complex connections in a deep orthography. The simple isomorphic connections in a shallow orthography should enable subjects to use the printed graphemes to activate their corresponding (unambiguous) phonological nodes, supplementing weaker activation generated more directly by degraded speech. This higher, aggregated, activation should reach threshold fast compared to a network representing a deep orthography with its weaker grapheme-phoneme connections.

Evidence against the orthographic depth hypothesis

We mentioned above the study by Seidenberg (1985) who studied word naming in Chinese and English. In English, he found no difference between regularly spelled words and exception words as long as their word frequency was high, suggesting that phonologic representations play no role in naming frequent words. Differences were found, however, for low frequency words, exception words having the longer latencies. In Cantonese, an analogous pattern was found: There was no significant latency difference between phonograms (compound characters that contain both a phonetic component and a logographic signific) and non-phonograms (characters that contain no phonetic)—as long as they were high frequency. For low frequency items, phonograms were named faster. However, what seemed to drive naming latency most strongly in both languages was word frequency. The results suggest that the effect of frequency, an effect that was similar in both orthographies, may be of overriding importance in determining which kind of lexical access code is successful; differences between orthographies that can affect the coding of phonology may be irrelevant when frequent words are compared.

Besner and Hildebrandt (1987) capitalized on the fact that Japanese is written in three script systems. One of the scripts is a logography that is derived from the Chinese but one in which the characters are pronounced differently. Two of these scripts are essentially syllabic orthographies, katakana and hiragana. Historically, their graphemes evolved from separate sources but they both address the same phonology, similar in this regard to the dual Cyrillic and Roman alphabets of Serbo-Croatian. Unlike Cyrillic and Roman, however, the Japanese scripts are rarely used to write the same words; instead, they "specialize," being used for mutually exclusive vocabularies. In Japanese, then, the pronunciations of those words that are logographic must be recalled lexically. However, those words that are normally printed in katakana and those words that are normally printed in hiragana can, in principle, be pronounced via grapheme-syllable correspondences. In a simple but direct experiment that compared only the katakana and hiragana scripts, subjects named words that were printed either in their normal script or in the other script. When printed in the normal script, naming times were 47 to 65 milliseconds faster than when printed in their atypical script. Besner and Hildebrandt interpreted the results to mean that subjects were not using grapheme-syllable correspondences in order to pronounce the normally printed stimuli because changing the visual-orthographic form had been detrimental; if subjects had been assembling the phonology for naming, they would not have been slowed by the change in grapheme-syllable script system. Thus, there is reason to suspect that Japanese readers always adopt a visual-orthographic mode for naming no matter the depth of the script system they are reading: the deep logography or the shallow syllabaries. In his chapter in this volume, Besner offers further details.

Additional evidence against the orthographic depth hypothesis was reported by Baluch and Besner (1991). They studied naming in Persian, an orthography which offers a comparison between words that omit the vowels, like all words in Hebrew (opaque words) and other words that are spelled with a relatively full phonological specification, like Serbo-Croatian (transparent words). The difference lies in the representation of the vowels; opaque words have one or more of certain specific vowels that can be written as diacritics but, instead, are typically omitted from the spelling (as in Hebrew) while transparent words contain only those vowels that are never omitted. The authors found that semantic priming equally facilitated both transparent words and opaque words; the weak orthographic depth hypothesis would predict less facilitation for the transparent words, which can be pronounced largely via assembled phonology, needing lexical information only for syllable stress. Differences between opaque and transparent words did appear when pseudowords were included in the list of words to be pronounced; then, only the opaque words were facilitated. The inclusion of pseudowords presumably biased the subject toward the use of addressed phonology as the default because the pseudowords would have had no addressed phonology and grapheme-phoneme correspondence rules were therefore an effective alternative. Apparently, when the recognition process is biased in this way, transparent words are no longer processed via visual-orthographic coding and they are no longer facilitated by semantic priming. Their results suggest that in normal reading, where there are no pseudowords, subjects may always use the direct lexical route, without the use of assembled phonology. Baluch and Besner note that in all the studies in which subjects used phonological recoding, pseudowords had been included in the stimulus list of words, perhaps biasing subjects toward an atypical processing strategy (e.g., (Katz et al., 1983; Frost et al., 1987; etc.).

A similar point was made by Tabossi and Laghi (1992). In a clever series of experiments, they showed that semantic priming effects in naming, which are indicative of lexical processing, disappeared or were attenuated when pseudowords were introduced. The implication is, then, that visual-orthographic coding was the preferred strategy for their subjects. The authors interpret their results to suggest that assembled phonology is produced only under artificial conditions such as when pseudowords are present. Because their experiments used a shallow orthography (Italian), the authors suggest that all orthographies, shallow and deep, use the same mechanism for processing print, i.e., the visual-orthographic route.

However, the alternative explanation, the ODH, is not directly addressed by their study (with one exception, discussed shortly). As we suggest above, no standard experiment on a single script system will be able to test the claim of the ODH that the amount of lexical involvement is greater in shallow than in deep orthographies. The ODH is a statement about relationships *among* orthographies; it does not categorically disallow the use of either assembled phonology or visual-orthographic processing for any orthography (except for the strong form of the hypothesis, which is clearly unacceptable on rational grounds). Thus, it is still not known if some phonological processing is occurring in Italian but is not being observed because of special characteristics of the experiment itself (e.g., the task, stimuli, etc.). As in all situations in which the experimenter is pressing the null hypothesis ("Can we show that *no* phonological processing is occurring?"), the better the set of alternative models for comparison, the more convincing the outcome. In the case of the ODH, the more convincing argument against it would be to show that manipulations that affect semantic facilitation cause identical effects (do not cause differential effects) between Italian and some orthography that is deeper than Italian. Tabossi and Laghi (1992) do, in fact, make this test; when they do, they find evidence that is consistent with the ODH. Semantically primed words were named faster than controls in English but not in Italian, suggesting that naming involved the lexicon more in the deeper English than in Italian. However, both word lists contained pseudowords which would tend to increase the amount of phonological processing for both.

A similar demonstration was made in the Frost, Katz, and Bentin (1987) study. In Experiment 3, the authors explicitly examined the same hypothesis that put forward by Baluch and Besner (1991). In this experiment the ratio of words to pseudowords in the stimulus list was manipulated and its effect on naming was measured in Hebrew, English, and Serbo-Croatian. The results showed marked differences in the pseudoword ratio effect in the three different orthographies. Whereas in Serbo-Croatian the inclusion of pseudowords had almost no effect on naming latencies (consistent with the notion that assembled phonology is the preferred strategy for that orthography), much larger effects were found in English and Hebrew. The point raised by Baluch and Besner is indeed important; pseudowords can, in fact, affect naming strategies. However, this issue has no direct relevance to the ODH. The ODH suggests that the relative effect of pseudoword inclusion should be different in deep and in shallow orthographies. The results of both Frost et al. (1987), and Tabossi and Laghi (1992) are compatible with this notion.

Sebastián-Gallés (1991) presented evidence that was said to be inconsistent with theories that propose different mechanisms for shallow and deep orthographies. Spanish subjects pronounced pseudowords that had been derived from real words; each pseudoword was orthographically similar to its counterpart, except for one or two letters. In some pseudowords, the correct pronunciation of a critical letter (*c* or *g*) changed from its pronunciation in the real word according to grapheme-phoneme correspondence rules, because of a change in the following vowel. In other pseudowords, there was no such change in the vowel (and, therefore, no change in the pronunciation from the real word model). Subjects pronounced about 26% of the change pseudowords contrary to the correspondence rules while only about 10% of the no-change pseudowords were pronounced in that way. Sebastián-Gallés interpreted this result to mean that subjects were using a lexical strategy for pronouncing pseudowords. But this interpretation is warranted only if the theory being tested allows *no* lexical involvement at all in naming: The author was attacking the strong form of the ODH. A closer look at the evidence suggests that the data are, in fact, consistent with the weak ODH. Seventy-four percent of the change pseudowords were pronounced in accordance with spelling-to-sound correspondence rules and only 26% were pronounced "lexically." Thus, a mix of the two processes may have been at work. In a second experiment, comparing lexical decision and naming times, Sebastián-Gallés found a moderate correlation (.455) for latencies between the two tasks. When the latency on each task was correlated with word frequency (presumed to be an index of lexical involvement), the correlation was of greater magnitude for lexical decision (-.497) than for naming (-.298). This result is consistent with a continuity of lexical involvement, naming being under weaker lexical control than lexical decision. A final experiment in this series showed semantic priming for naming under conditions where it is usually not found in shallow orthographies, viz., when pseudowords are included in the list of stimuli. Such results do suggest more lexical involvement in naming for a shallow orthography than other research has suggested. Nevertheless, in the author's conclusion, Sebastián-Gallés interprets the sum of the results to mean that "... lexical access from print in Spanish involves the use of orthographic information during at least some of the processing time," a statement that is consistent with the weak ODH.

Concluding remarks

Our working hypothesis has been that all alphabetic orthographies make some use of assembled phonology for word recognition. The proposal that a mixture of prelexical and visual-orthographic information is used for word recognition is consistent with the weak form of the ODH. The approach we suggest is in line with those models that contain dual phonological and orthographic representations, such as the dual route models (see Paap, this volume) and network models (see Seidenberg, this volume). The question of just how prevalent the use of phonology is, relative to the use of visual-orthographic coding, for any given orthography is an open one and is not addressed by the ODH itself. It could even be the case that the predominant lexical access code for frequent words in the shallow Serbo-Croatian is actually visual-orthographic or, on the other hand, that the predominant code in Hebrew is based on the partial phonological information that can be assembled from the unpointed letters (although either possibility seems unlikely from the present evidence). The ODH does not specify what degree of orthographic depth determines predominance for, say, visual-orthographic coding. Of course, orthographic depth will not be the only determiner of dominance; the reader's experience should play a major role as well.

Which of the codes is dominant might be determined by a mechanism like the following. Assume that the processing system is capable of using either code: a dual-code model. Suppose that a phonological representation is the *default* code for any given word but processing of a word via its phonological code can be replaced by processing via its visual-orthographic representation when the word has been experienced by the reader a sufficient number of times: a word frequency criterion. The premise that phonology is the default code is based on the fact that it is typically the code of instruction and the beginning reader receives much practice in its use. (One piece of evidence supporting the notion that the default code is phonological is that even adult readers of Hebrew prefer to use phonological information if it is available; Frost, forthcoming). The criterion word frequency that is required in order to replace processing by assembled phonology with processing by visual-orthographic representations should be a function of the costs involved—in part, the cost for assembling the phonological representation and using it to access the lexicon. A higher replacement criterion will obtain in a shallow orthography where assembled phonology is easy to generate than in a deeper orthography.

What other factors affect the cost? Besides the ease of assembling phonology, a second factor is the ease with which phonology can be used in the lexical search process itself. Likewise, the ease of generating a visual-orthographic representation that is suitable for lexical search, the ease with which *that* information can address the lexicon, and, of course, the cost involved in establishing the visual-orthographically coded lexical representation in the first place, all need to be evaluated in establishing the criterion. The tradeoff between the generation of information that can be used for lexical access and the access process itself is important. Visual-orthographic codes may be costly for both the beginning and skilled readers to generate, particularly if a word is morphologically complex (it may require decomposition, in that case). However, the search process based on a visual-orthographic representation may be rapid for the skilled reader once he or she has a well-established visual-orthographic representation in lexical memory. Phonological codes may be difficult to generate but, once obtained, may be a fairly natural (i.e., low cost) way of addressing lexicon; after all, our primary lexicon, the speech lexicon, is based on phonology. Although the phonological lexicon may have taken years for a child to develop, it is thereafter available free to the reading process. However, it is obvious that we know very little about this tradeoff in terms of processing costs. The answer to the question about which of the two representations, phonological or visual-orthographic, is dominant depends on knowing more than we presently do about the perceptual and cognitive resources involved in word recognition. This question is discussed in some detail by Seidenberg, (this volume). Nevertheless, we repeat that even in the absence of a fuller understanding of how the word recognition process draws on these resources, it is still a plausible (and testable) hypothesis that word recognition in shallow orthographies will depend more on phonological representations simply because such information is available at less cost in those orthographies. Much of the evidence we have presented here is consistent with that hypothesis.

There has been substantial progress over the past decade in understanding the mechanisms behind naming and recognizing printed words. Importantly, research has involved an increasing variety of languages and writing systems, forcing theory to be general enough to encompass this wider scope. The vitality of this research is great and even shows signs of increasing in its pace. We hope that the main import of this article will be to clarify some issues in this area on the differential effects of writing systems on word perception. If there is presently significant (although perhaps not universal) agreement among researchers that visual-orthographic and assembled phonological representations may both play roles in word perception, then the next phase of research activity must include ways of assessing the conditions under which they are active, the relative contributions of each, and the mechanisms of their action. This requires a switch from research designs that address qualitative questions (e.g., "Is the lexicon accessed phonologically or not?") to designs that address the relative balance of phonological and visual-orthographic coding.

The best candidate for a heuristic framework for this research may be network modeling which offers a natural way of simulating the relationships between orthography and phonology, orthography and morphology, and phonology and morphology (at one level) and between these coded representations and the lexicon (at another). Likewise, differences between and within orthographies concerning the consistency, regularity, and frequency of these relationships can be implemented as initial constraints on such networks. Seidenberg (this volume) suggests that network architectures offer ways of modeling how the various general cognitive resources involved are adapted to the processing of printed words: how the system assembles itself under the constraints of language, orthography, and memory. Implicit in this characterization is the additional possibility of modeling the historical evolution of a writing system. While the resources of memory, perceptual discrimination, and the like, may be constant, languages and their orthographies have not been immutable and their histories of change are well known in many cases. The orthographic depth hypothesis itself is a statement about a part of this larger issue of the fit between writing systems and human capabilities. The hypothesis embodies the assumption of covariant learning (Van Orden et al., 1991). That is, the structure and operation of the network should reflect the contingencies among phonology, morphology, and orthography that exist for printed words and, therefore, the contingencies that will be experienced by a reader. Each orthography, shallow or deep, defines its own pattern of contingencies. Additional progress in this area should come from requiring our ideas about the differences among orthographies to be made precise enough to be modeled.

Acknowledgment
Work on this paper was supported by National Institute of Child Health and Human Development Grant HD-01994 to Haskins Laboratories. The authors wish to thank I. G. Mattingly for comments on an earlier draft.

References

Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 644-652.

Besner, D., & Hilderbrandt, N. (1987). Orthographic and phonological codes in the oral reading of Japanese kana. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 335-343.

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

Cossu, G., Shankweiler, D., Liberman, I. Y., Katz, L., & Tola, G. (1988). Awareness of phonological segments and reading ability in Italian children. *Applied Psycholinguistics, 9,* 1-16.

DeFrancis, J. (1989). *Visible Speech: The diverse oneness of writing systems.* Honolulu: University of Hawaii Press.

Feldman, L. B., & Turvey, M. T. (1983). Word recognition in Serbo-Croation is phonologically analytic. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 288-298.

Frost, R. (forthcoming). Prelexical and postlexical strategies in reading: Evidence from a deep and a shallow orthography.

Frost, R., & Katz, L. (1989). Orthographic depth and the interaction of visual and auditory processing in word recognition. *Memory & Cognition, 17,* 302-310.

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.

Katz, L., & Feldman, L. B. (1981). Linguistic coding in word recognition: Comparisons between a deep and a shallow orthography. In A. M. Lesgold &. C. A. Perfetti (Eds.), *Interactive processes in reading.* Hillsdale, NJ: Erlbaum.

Katz, L., & Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 157-166.

Liberman, I. Y., Liberman, A. M., Mattingly, I. G., & Shankweiler, D. L (1980). Orthography and the beginning reader. In J. F. Kavanagh &. R. L. Vanezky (Eds.), *Orthography, reading and dyslexia* (pp. 137-153). Baltimore: University Park Press.

Lukatela, G., & Turvey, M. T. (1990a). Phonemic similarity effects and prelexical phonology. *European Journal of Cognitive Psychology, 18,* 128-152.

Perfetti, C. A., Bell, L. C., & Delaney, S. M. (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Memory and Language, 27,* 59-70.

Scheerer, E. (1986). Orthography and lexical access. In G. Augst (Ed.), *New trends in graphemics and orthography* (pp. 262-286). Berlin: De Gruyter.

Sebastián-Gallés, N. (1991). Reading by analogy in a shallow orthography. *Journal of Experimental Psychology: Human Perception and Performance, 17(2),* 471-477.

Seidenberg, M. S. (1985). The time-course of phonological code activation in two writing systems. *Cognition, 19,* 1-30.

Seidenberg, M., Waters, G., Barnes, M. A., & Tannenhaus, M. K. (1984). When does irregular spelling and pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23,* 383-404.

Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory & Cognition, 20,* 315-328.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review 97,* 488-522.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 5

Beyond Orthographic Depth in Reading: Equitable Division of Labor

Mark S. Seidenberg
University of Southern California, Los Angeles

Introduction: The orthographic depth hypothesis

Research on how orthography influences visual word recognition is important because it has the potential to reveal both universal and orthography-specific aspects of reading. By universal aspects, I mean commonalities among the readers of different orthographies in terms of basic processing mechanisms. We might expect these to exist because people share basic perceptual and cognitive capacities. At the same time, many writing systems have evolved and they differ in ways that could influence the reading process.

The world's orthographies represent different solutions to the problem of representing spoken language in written form. One of the fundamental ways in which they differ is in the extent to which the written and spoken forms correspond. The evolution of writing systems has been marked by a trend toward more direct representation of sound (Hung & Tzeng, 1981). For example, the symbols in logographies typically encode aspects of meaning, whereas the symbols in alphabetic orthographies, which emerged later, encode phonemes. Within the alphabetic orthographies, the consistency of the mapping between orthography and phonology varies. In "shallow" alphabetic orthographies (Turvey, Feldman, & Lukatela, 1984), the correspondences between graphemes and phonemes are entirely consistent. Serbo-Croatian provides a well-studied example. The Serbo-Croatian language has two orthographies, Roman and Cyrillic. Both are shallow in the sense that each letter corresponds to one phoneme. The only complication—at the level of graphemes and phonemes—is that a small number of letters appear in both orthographies with different pronunciations. There is minimal ambiguity, however, because the two orthographies are not mixed in ordinary texts.

Written English also provides systematic information concerning pronunciation, but the correspondences between graphemes and phonemes are notably inconsistent. A large number of studies have examined how the inconsistencies illustrated by words such as GIVE-DIVE, PAID-SAID, DOSE-POSE-LOSE, and COUGH-ROUGH-DOUGH-PLOUGH-THROUGH-THOUGH influence reading (e.g., Jared, McRae, & Seidenberg, 1991). These inconsistencies derive from several sources, including the fact that the orthography also encodes morphological information (Chomsky & Halle, 1969), diachronic changes in pronunciation, and periodic spelling reforms (see Seidenberg & McClelland, 1989, for discussion).

There are many degrees of variation along the dimension of orthographic depth. Chinese, for example, employs a nonalphabetic orthography in which the symbols provide partial cues to either meaning or pronunciation (Wang, 1979). The Hebrew orthography is shallow, but the vowels are indicated by diacritical symbols that are usually omitted, introducing a high degree of spelling-sound ambiguity. Japanese employs two orthographies, one of which (Kanji) consists of logographic characters borrowed from Chinese, and the other of which (Kana) consists of symbols encoding syllables or moras. Whereas the former provide partial information about pronunciation at best, the latter are strictly regular.

This variation among orthographies has led to what has been termed the orthographic depth hypothesis (Katz & Feldman, 1981; Katz & Frost, this volume; Turvey et al., 1984). It has long been recognized that a word could in principle be recognized in two ways in alphabetic orthographies. First, the reader could attempt to recognize words on a visual basis, ignoring the fact that the symbols encode information about pronunciation. In this case, word recognition would be like other pattern recognition processes used in recognizing objects or nonalphabetic symbols. Having recognized the letter string as a token of a particular lexical type, the reader could then access its meaning, which is stored in a mental dictionary, an outcome termed "direct access" (Baron, 1973; Barron, 1978; Rubenstein, Lewis, & Rubenstein, 1971). Second, recognition could be based on a phonological code derived on the basis of the reader's knowledge of the correspondences between spelling and pronunciation. The meaning of a word would then be accessed using this derived phonological code, an outcome termed "phonologically-mediated access" (ibid.). The orthographic depth hypothesis is that the extent to which one or the other of these processes is employed depends on properties of the writing system. "Deep" orthographies are thought to discourage the use of phonological recoding because the correspondences between spelling and pronunciation are inconsistent; hence the orthographic process is more efficient. "Shallow" orthographies afford a phonological recoding strategy, however, because the correspondences are consistent. Thus, readers adapt their processing strategies to the demands of the orthography. This view was succinctly summarized by Katz and Feldman (1981, pp. 85-86):

> [T]he kind of code that is used for lexical access depends on the kind of alphabetic orthography facing the reader. Specifically, it depends on how directly the orthography reflects the phonetic surface. Languages in which the spelling-to-sound correspondences are simple and invariant (as in Serbo-Croatian) will readily support information-processing structures for reading that utilize the language's surface phonological features. On the other hand, in an orthography that bears a complex relation to speech (a deep orthography such as English), phonologically structured mechanisms for processing words will be less developed.

A large amount of evidence, derived from studies of several diverse orthographies, including those for English, Serbo-Croatian, Chinese, Japanese, and Hebrew, is consistent with this general hypothesis. Much of the critical evidence has been provided by the Haskins Laboratories group and their Yugoslavian and Israeli colleagues (e.g., Frost, Katz, & Bentin, 1987; Katz & Feldman, 1981; Lukatela & Turvey, 1980; Turvey et al., 1984), who have engaged in a sustained effort to understand how differences among the

English, Serbo-Croatian, and Hebrew orthographies influence processing. Findings of the following sort have been reported. When subjects who are skilled readers of Serbo-Croatian perform the lexical decision task, factors related to the pronunciations of the stimuli affect performance. For example, many of the studies have exploited the ambiguity between (but not within) the Roman and Cyrillic alphabets mentioned above. A stimulus such as BETAP contains letters that appear in both alphabets. How it is pronounced depends on whether it is interpreted as Roman or Cyrillic, owing to the ambiguity of the letters B and P. Lexical decision latencies for these ambiguous strings are reliably longer than latencies for unambiguous strings such as VETAR (Lukatela, Popadić, Ognjenović, & Turvey, 1980; Lukatela, Turvey, Feldman, Carello, & Katz, 1990). This *phonological ambiguity effect* is taken as strong evidence that phonological mediation is the rule in reading Serbo-Croatian. Only if subjects were computing phonological codes would the ambiguity have an effect. Lexical decisions are assumed to be performed by determining if a stimulus string has a meaning or not. Therefore, if phonological factors influence decision latencies, it must be that subjects derive the phonological code as part of the process of accessing meaning. This basic phenomenon has been replicated many times and is discussed further below. These effects have led some researchers to conclude that phonological recoding is obligatory in Serbo-Croatian, as indicated by the titles of their papers: "The Serbo-Croatian orthography constrains the reader to a phonologically analytic strategy" (Turvey et al., 1984); "Visual word recognition in Serbo-Croatian is necessarily phonological" (Feldman, 1981).

Where does the English orthography fit in light of the orthographic depth hypothesis? Noting the inconsistencies of its grapheme-phoneme correspondences, Turvey and colleagues present English as an example of a "deep" orthography. This implies that words should be recognized on a visual basis. However, the grapheme-phoneme correspondences of English are by no means arbitrary, as indicated by the fact that readers can reliably assign pronunciations to novel stimuli such as NUST or RANE. The issue as to whether word recognition in English is direct or phonologically-mediated is a vexed one that has been the subject of extensive investigation (for reviews, see McCusker, Hillinger, & Bias, 1981; Carr & Pollatsek, 1985). Many researchers have held the view that both processes are available as alternative strategies for reading English (see, for example, Henderson, 1982). Research has focused on identifying the factors that determine which process is used in a given case; for example, phonological mediation was thought to be used by certain readers (i.e., good readers—Barron, 1981; poor readers—Doctor & Coltheart, 1980), or with certain types of words (e.g., regular words—Coltheart, 1978; lower frequency words—McCusker et al., 1981; Seidenberg, 1985).

Both the cross-linguistic studies and the studies of English, then, converged on a "dual-route" conception of reading (Coltheart et al., 1977; Meyer, Schvaneveldt, & Ruddy, 1974; Paap, this volume). There are both direct and phonologically-mediated routes to meaning; deep orthographies such as unvoweled Hebrew afford only the direct route; shallow orthographies such as Serbo-Croatian afford the phonologically-mediated route; English makes use of both, but for different words (Katz & Feldman, 1981).

This picture is quite pretty, and the essential idea—that it is easier to derive phonological codes in some orthographies than in others, and that this variable should influence the extent to which phonological codes activate meaning—seems correct to me. However, I think that emphasis on the orthographic depth hypothesis has tended to obscure similarities in how different writing systems are processed. These similarities may outweigh relatively minor differences between orthographies owing to the depth factor. Moreover, there are questions as to whether the extent to which the activation of meaning is phonologically mediated varies across orthographies in the manner predicted by the ODH. On the one hand, there is reason to question whether the studies of Serbo-Croatian demonstrate that word recognition in this orthography is "necessarily phonological." Many of the arguments are based on interpretations of data from naming and lexical decision experiments that are problematical in light of current theories of word recognition and of these tasks. If word recognition is not necessarily phonologically-mediated in Serbo-Croatian, there may be greater similarities between the processing of this orthography and the processing of deeper orthographies than have been recognized to this point. On the other hand, recent studies of Hebrew and English suggest more reliance on phonological mediation in these "deep" orthographies than suggested by the ODH. Indeed, it has recently been argued that accessing the contextually-appropriate meaning of an unpointed word in Hebrew, the quintessential "deep" orthography, requires phonological mediation (Frost, 1991). Along the same lines, Van Orden, Johnston, and Hale (1988) have claimed that "Word identification in reading [English] proceeds from spelling to sound to meaning." That is, they conclude that, like Serbo-Croatian, word recognition in English is necessarily phonologically mediated. This conclusion clearly is at odds with the theory that the extent to which phonological mediation is employed depends on properties of the orthography.

In the sections that follow, I address these issues in more detail. I first consider questions concerning the ODH that have arisen with regard to Serbo-Croatian, English, and Hebrew. I then try to place the issue of orthographic depth in a somewhat broader theoretical framework suggested by the Seidenberg and McClelland (1989) model. This framework suggests that a common, multi-component architecture underlies processing in different orthographies. According to this view, the critical issue concerns what I term the "division of labor" between processing mechanisms in this multi-component processing system. Orthographic depth is one of the major factors that has an impact on the division of labor. However, orthographies differ in other respects that are also relevant. It appears that the tradeoffs among these factors are such that readers converge on very similar divisions of labor despite substantive differences between orthographies.

Three challenges for the orthographic depth hypothesis

1. Serbo-Croatian

There are a prodigious number of studies of reading in Serbo-Croatian, many of which are discussed in other chapters in this book. Rather than attempting an exhaustive review with the goal of establishing definitive conclusions about processing in this orthography, I will focus on some of the paradigmatic effects that have been taken as supporting the orthographic depth hypothesis, specifically (a) phonological ambiguity effects, and (b) effects of frequency and lexical (word-nonword) status. Though these effects by no means subsume all of the relevant phenomena that have been uncovered, they provide a basis for raising some general questions concerning the interpretation of word recognition data and their relevance to the orthographic depth hypothesis.

Phonological ambiguity effects and lexical decision

Across a broad range of conditions, letter strings that are ambiguous between Roman and Cyrillic pronunciations produce longer lexical decision latencies than unambiguous strings. The effect obtains for both words and nonwords (e.g., Feldman, 1981; Lukatela et al., 1989a; Turvey et al., 1984); it is eliminated when the stimuli appear in disambiguating contexts that specify the target's alphabet (Lukatela et al., 1989b). The effect also obtains when the stimuli are named aloud (Feldman, 1981; Lukatela et al., 1989a).

The conclusion that these effects show that word recognition is phonologically mediated rests on a particular theory of the representation of lexical knowledge in memory and of the lexical decision task. As in most theories, it is assumed that words are represented in terms of entries in a mental lexicon. Word recognition involves accessing the entries corresponding to the stimulus ("lexical access"). A critical assumption is that positive lexical decisions—"yes" responses for words—are made by accessing their meanings. Hence, phonological ambiguity effects obtained with this task indicate that phonological information is used to access meaning. Much of the evidence for phonological mediation in Serbo-Croatian is moot, however, if this account of the task is incorrect. The question, then, is whether it can be sustained in light of other evidence.

Few of the Serbo-Croatian studies provide direct evidence that subjects accessed meaning in making lexical decisions. Rather, it is typically *assumed* that this must have taken place if subjects responded correctly. Unfortunately, there is other evidence that making a lexical decision is not as simple as merely deciding if a letter string has a meaning or not. Words and nonwords differ along dimensions other than meaning that can also provide bases for making lexical decisions. For example, words are more familiar orthographic patterns than nonwords; words have been seen before and nonwords have not. Similarly, words have more familiar sound patterns than nonwords; only words are heard in speech. As set out before the subject, the lexical decision task is not to determine if the target stimulus has a meaning or not; it is to reliably discriminate between the two types of stimuli. As in a signal detection task, the subject must establish decision criteria that allow fast responses with acceptable error rates. These criteria could in principle involve any of the several dimensions along which words and nonwords differ. Thus, many discussions of the lexical decision task assume that subjects make their responses by judging the familiarity of stimuli in terms of orthography, phonology, or semantics (Balota & Chumbley, 1984; Gordon, 1983; Seidenberg, Waters, Sanders, & Langer, 1984b). Perhaps the primary conclusion from extensive use of the task is that response criteria vary as a function of the properties of the stimuli in an experiment. As subjects' response criteria vary, so do the effects of variables such as frequency, orthographic-phonological regularity, and contextual congruence (e.g., Forster, 1981; Stanovich & West, 1981; Neely, 1977; Seidenberg et al., 1984b).

There is good evidence that subjects vary their decision criteria in response to properties of the stimuli. Consider, for example, the lexical decision studies reported by James (1975). The word stimuli varied in terms of concreteness. Different types of nonwords were used across experiments. There were two types of pronounceable nonwords (pseudohomophones, which sound like words, and nonpseudohomophones, which do not), as well as nonpronounceable nonwords (formed by transposing letters in the pronounceable strings). There was an effect of concreteness only when the nonwords were pronounceable. The results suggest that semantic information, indexed by the concreteness factor, only entered into the decision process under some conditions. When the nonwords are nonpronounceable, subjects can base their decisions on whether the target sounds (or looks) like a word. Hence the semantic factor has no effect. When the stimuli look and sound very wordlike, subjects are forced to use semantic information in making their decisions.

Complementary effects involving phonological information were reported by Shulman, Hornak, and Sanders (1978) and Waters and Seidenberg (1985). The Shulman et al. studies examined phonological priming effects (i.e., BRIBE-TRIBE facilitation; FREAK-BREAK inhibition). These effects were obtained with the lexical decision task when the nonword stimuli were pronounceable, replicating Meyer et al.'s (1974) earlier finding. With nonpronounceable nonwords, however, the phonological priming effect was eliminated. Hence, whether phonological information was used in making the lexical decision depended on the properties of the nonwords. Waters and Seidenberg (1985) obtained similar results, but with two significant differences in procedure. Instead of phonological priming effects, they examined phonological regularity effects. Instead of manipulating the properties of the nonwords, they manipulated the properties of the words. When the word stimuli consisted of regular and exception words and nonwords, no regularity effect obtained. When the stimuli also included the orthographically strange words, a regularity effect resulted. In the former case, the words and nonwords could be discriminated on an orthographic basis; hence the phonological factor was irrelevant. Including strange words in the latter case disabled this orthographic decision strategy; once phonological information entered into the decision process a regularity effect obtained.

One additional bit of evidence concerning the lexical decision task is provided by the model of word recognition (in English) developed by McClelland and myself (Seidenberg & McClelland, 1989). Although we presented a general framework for thinking about lexical processing, the simulation model that we implemented only addressed part of the system. The model takes a letter string as input and produces two types of output: an orthographic code and a phonological code. These codes are represented as patterns of activation over sets of units representing these types of information. There are no entries for individual words; thus, orthographic and phonological codes are computed rather than accessed (Seidenberg, 1990). We developed an account of how these computed codes are used in performing tasks such as lexical decision and naming (see Seidenberg & McClelland, 1989, for details). The observation relevant to this discussion is that the model accounts for the results of several lexical decision studies even though it is wholly *lacking* any representation of meaning. Briefly, we show that decisions can be based on the properties of the computed patterns of activation over the sets of orthographic and phonological units. Although we have not attempted to simulate all of the conditions in the many lexical decision experiments in the literature, we have provided a principled explanation for why decision criteria vary in response to orthographic and phonological properties of the stimuli and applied this analysis to some representative cases (e.g., the Waters & Seidenberg, 1985, studies; the Gordon, 1983, study). This modeling exercise suggests that under at least some conditions, lexical decisions need not involve access to meaning, calling into question whether the Serbo-Croatian studies can necessarily be taken as providing evidence concerning access of meaning.

The model also exhibits behaviors that are relevant to phonological ambiguity effects. In the model, the presentation of an orthographic input results in the computation of phonological output. The characteristics of this computation are determined by the weights on connections between units, which govern the spread of activation. In the simulations, the values of the weights were determined during a learning phase in which the model was exposed to a large subset of the English lexicon (2897 monosyllabic words). The weights were modified on the basis of feedback concerning the correct orthographic and phonological codes for input strings using the backpropagation learning algorithm (Rumelhart, Hinton & Williams, 1986). The weights come to reflect facts about the correspondences between these codes picked up during the learning phase. The training corpus included 12 heterophonic homographs—words such as WIND, BASS, and LEAD that have two pronunciations, corresponding to different words. We arbitrarily decided to train the model on both pronunciations equally often; for example, half the time it was told that the pronunciation of WIND is /wind/ and half the time that it is /wInd/. Given this inconsistent feedback, which reflects the fact that these words are genuinely phonologically ambiguous, the model performed poorly on these items, even after a large amount of training. This result is consistent with the finding that homographs yield very long naming latencies when they are presented as isolated stimuli (Seidenberg et al., 1984a).

Two points should be noted: first, words like WIND are the closest English analogues to phonologically-ambiguous Serbo-Croatian items like BETAP, and produce similar effects. Second, in the Seidenberg and McClelland model, phonological ambiguity effects derive from the properties of a computation from orthography to phonology that is independent of meaning.

In summary, it is clear that lexical decisions are sometimes based on nonsemantic information. This fact limits the kinds of inferences that can be drawn from LD results about the factors that influence access to meaning. It cannot be assumed that subjects necessarily access meaning in this task; they might simply access the codes that are necessary to perform it. Hence, the effect of a factor such as phonological ambiguity may simply indicate that subjects have accessed phonological information, not that they have used it to access meaning.

I hope it is clear that these considerations merely establish that phonological ambiguity effects are indeterminate with regard to the claim that phonological mediation is obligatory in Serbo-Croatian, not that the claim is false. We do not know whether the stimuli in the experiments afforded the non-semantic response strategies that have been observed in some studies of English. Moreover, the fact that our model can make lexical decisions without access of meaning doesn't necessarily mean that people do the same thing. My point is only that theoretical implications of these effects are ambiguous, something that has not been acknowledged in discussions of orthographic depth. This means that other types of evidence will be required in order to resolve the issues.

Lexicality effects and naming

Naming aloud is the other principal task that has been used in studies of Serbo-Croatian and the ODH. The data have been interpreted within the standard "dual-route" model of reading developed by Coltheart (1978) and others. Again, however, recent theoretical advances call into question assumptions about the task that affect the interpretation of the data. The standard model assumes that there are "lexical" and "nonlexical" pronunciation processes. "Lexical" pronunciation involves recognizing a word on a visual basis ("accessing an entry in the orthographic input lexicon") and then looking-up a stored representation of the word's pronunciation ("accessing an entry in the phonological output lexicon"), which is then used to formulate the articulatory response. The lexicons are assumed to be organized in terms of frequency; hence, pronunciation by this route is frequency-dependent. "Nonlexical" pronunciation involves applying rules governing the correspondences between spelling and pronunciation; in Coltheart's formulation, these were termed "grapheme-phoneme correspondence rules." These rules can be applied to any letter string without regard to frequency or lexical (word-nonword) status. Thus, frequency and lexicality effects were taken as evidence for involvement of the "lexical" route.

Frost et al. (1987) employed this logic in a comparative study of word recognition in English, Serbo-Croatian, and Hebrew. This is an important paper, providing a wealth of interesting data. Frost et al. assumed that orthographic depth reflects the extent to which the pronunciations of words can be derived by lexical vs. nonlexical processes. In Serbo-Croatian, pronunciations can be derived nonlexically, because the orthography is shallow; in unpointed Hebrew, pronunciations can only be determined lexically because the orthography is deep; English falls somewhere in between. Thus, depth of the orthography should be related to the degree of "lexical" involvement in naming. This generated the prediction that the size of frequency effects and word-nonword differences should be a function of the depth of the orthography, which is what was found. Hebrew produced the largest effects, English produced medium-sized effects, and Serbo-Croatian the smallest. Thus, the authors concluded that the study provided evidence for greater "lexical involvement" in naming in the deeper orthographies.

This reasoning was valid with respect to the principal theory of the day, but not the Seidenberg and McClelland model. As discussed elsewhere (Seidenberg & McClelland, 1989; Seidenberg, 1989; Seidenberg, in press), ours is a "dual-route" model in the sense that there is a "direct" computation from orthography to semantics and an "indirect" computation from orthography to phonology to semantics. Insofar as the former applies only to words and the latter to both words and nonwords, these computations can be seen as analogous to the routes in the dual-route model. Importantly, however, our computation from orthography to phonology does not involve pronunciation rules; rather it involves weighted connections between units encoding distributed representations. This change in the type of knowledge representation employed has two consequences. First, the model can generate correct output for both "rule-governed" words and exceptions, whereas the rules will, by definition, generate incorrect phonological codes for the exceptions. Second, the orthographic to phonological computation in our model is affected by frequency and lexicality, the factors previously thought to differentiate between "lexical" and "nonlexical" naming mechanisms.

These aspects of the model have important implications concerning the interpretation of data such as Frost et al.'s. Their view is that frequency and lexicality effects reflect the degree to which the "lexical" route contributes to naming, which is a function of orthographic depth. In contrast, our model suggests that these effects simply reflect properties of a single computation from orthography to phonology. This "route" exhibits a property, sensitivity to word frequency, that was previously assigned to the "lexical" route. That such effects are larger in deeper orthographies is simply a consequence of how the weights on this pathway are set during learning, not the degree to which a second route is implicated. A connectionist model trained using backpropagation or another error-correcting learning algorithm will pick up on statistical properties of the correspondences between input and output codes. The values of the weights reflect the aggregate effects of training on a corpus of words. Thus, performance on any given word is affected by exposure to all other words in the corpus. Given the properties of the English lexicon to which our model was exposed—specifically, the distributions of orthographic patterns and orthographic-phonological patterns in the training lexicon—performance on any given word was largely determined by two factors: how often the word was presented during training (that is, its frequency), and how often the model was trained on similarly spelled words (its "neighbors"). Thus, performance on GAVE is affected by exposure to GAVE; it also benefits from exposure to SAVE, PAVE, and RAVE, and is penalized by exposure to HAVE. All of these effects are realized in the same manner, changes to the weights. In fact, it is one of the theoretical claims of our model that frequency and neighborhood effects derive from the same source, the learning procedure (Seidenberg & McClelland, 1989).

Frequency and neighborhood effects interact in ways that are relevant to understanding the effects of orthographic depth on naming observed in the Frost et al. study. As the number of neighbors of a word increases, the effect of lexical frequency decreases (Jared et al., 1990). Intuitively, performance on a word such as GAVE, which has a large number of neighbors, is not very dependent on frequency of exposure to it because of support from the similarly spelled and pronounced neighbors. The neighbors shift the weights in ways that also benefit GAVE. Performance on an irregular word such as HAVE, however, is highly dependent on frequency of exposure; because its pronunciation is irregular, it does not benefit from exposure to similarly spelled and pronounced neighbors. In fact, there must be enough exposures to the word itself to overcome the negative effects of exposure to the GAVE et al. cohort. The model therefore correctly simulates the fact that frequency effects in naming are larger for irregular words than for regular words.

The larger frequency effects and word-nonword differences in English suggest that the readers' frequency of exposure to particular word-forms exerts more of an effect in the deeper orthography. The effects are smaller in Serbo-Croatian because the dominant factor is simply the very regular correspondences between graphemes and phonemes. In English the pronunciations of graphemes—especially vowels—depend on the lexical contexts in which they occur. Whether the pattern -AVE occurs in the context H- or S- determines how the vowel is pronounced. Pronunciation is less context-sensitive in Serbo-Croatian, because it is shallow; vowels, and other graphemes, are pronounced the same way in all contexts (within the Cyrillic or Roman alphabet, of course). Hence, latency to pronounce a given word in English should be more dependent on how often it has been experienced than in Serbo-Croatian. The same account applies to the word/nonword effects. Within this type of model, the pronunciations of both words and nonwords are generated in exactly the same way; in effect, nonwords simply function as very low frequency words. Since the pronunciations of graphemes and phonemes in Serbo-Croatian are not dependent on the contexts in which they occur, it should matter less whether a particular letter string happens to form a familiar word.

This account can be tested by running a version of the model that simulates what would occur if English were as shallow as Serbo-Croatian. The model can be trained exactly as in our original simulation, correcting the weights on the basis of feedback concerning the correct pronunciations of words. However, the training set can be formulated so that grapheme-phoneme correspondences within it are entirely regular. I took the 2897 word training set used in the original simulation, and "regularized" the pronunciations of all exception words. Thus, the model was trained that HAVE is pronounced /hAv/, GIVE as /gIv/, DONE as /dOn/, and so on. In order to make the corpus as regular as possible, I eliminated items containing spelling patterns that only occur in one word (e.g., AISLE, ONCE). The model was then trained exactly as in the original simulation. I examined the effects of this change in the training corpus in reference to a set of 48 high and low frequency regular words that were used in a study by Taraban and McClelland (1987), and 48 nonwords derived from these items by changing the initial consonant or consonant-cluster. For example, the list contained the high frequency word MUST and the derived nonword NUST. The model's performance on these items is measured in terms of error scores that reflect the discrepancy between the computed phonological output and the output that would be produced if the model performed without error. Larger error scores indicate poorer performance.

In the original simulation, the average error score for the 24 higher frequency words after 250 training epochs was 2.968; for lower frequency words it was 3.829, and for nonwords it was 8.376. There was a small frequency effect for words and a somewhat larger difference between words and nonwords. In the modified simulation, the mean scores were as follows: high frequency words, 3.456; low frequency words, 3.848; nonwords, 6.911. Both frequency and lexicality effects decreased in magnitude, resulting in much smaller differences between the stimulus types.[1]

Thus, given the inconsistencies in the spelling-sound correspondences of English, lexical frequency has a larger impact than exposure to other words containing some of the same graphemes and phonemes. As the consistencies at the level of graphemes and phonemes increases—in shallow orthographies—the effects of lexical frequency decrease. In studies such as Frost et al.'s (1987), frequency and lexicality effects were taken as evidence for use of the direct, orthographic recognition process. The simulations suggest that this conclusion is not necessarily valid because the effects can be derived from properties of the orthographic-phonological computation.

These observations must be weighed against other important data provided in the Frost et al. paper. It is difficult to arrive at a theory that can accommodate all of their results. I have suggested that frequency and lexicality effects do not necessarily implicate anything other than properties of the computation from orthography to phonology. This reanalysis is also consistent with their lexical decision data, which indicate both frequency and lexicality effects in all three orthographies. Thus, in contrast to naming, orthography had very little effect on lexical decision performance. This surprising result is hard to reconcile with the orthographic depth hypothesis given the assumptions that (a) lexical decisions require access to meaning, and (b) there are differences between orthographies in the extent to which phonology influences the access of meaning. These assumptions suggest that there should be *some* differences between the orthographies with the lexical

[1] There is a small residual frequency effect and somewhat larger lexicality effect in both the simulation and in the Frost et al. Serbo-Croatian condition. Besner and Hildebrandt (1987) also obtained a frequency effect in Japanese Kana. These results might be taken as problematical for the view that words in shallow orthographies are read on the basis of nonlexical rules. Frequency is a "lexical" factor to which the rules are not supposed to be sensitive. However, frequency and lexicality effects could arise in part because of differences between items in terms of the familiarity of their orthographic patterns. The initial process of recognizing letter patterns may be affected by frequency of exposure to them. In the Seidenberg and McClelland model, for example, high and low frequency words (such as MUST and BUST) and nonwords (such as NUST) produce different orthographic error scores, which are a measure of orthographic familiarity, even thought their spellings are very similar. Hence, the mere finding of a frequency or lexicality effect in a shallow orthography cannot be taken as evidence against the use of nonlexical pronunciation rules. The point of the simulations of deep and shallow English is that the relationship between orthographic depth and the relative sizes of the frequency and lexicality effects observed in the Frost et al. experiments can be explained in terms of the settings of the weights governing the orthographic-phonological computation.

decision task, but none were observed. If (a) is correct, then the results appear to indicate that orthographic depth has little influence on the access of meaning. If (a) is not correct, then the data are indeterminate with respect to the central claim that phonological mediation varies in proportion to orthographic depth.

This picture is complicated further by another of the Frost et al. findings, larger semantic priming effects in the deeper orthographies using the naming task, with no effect at all in Serbo-Croatian. This result was also obtained by Katz and Feldman (1983) in a study of English and Serbo-Croatian. Frost et al.'s interpretation is that whereas naming in Serbo-Croatian can be performed "nonlexically," naming in English or Hebrew requires lexical mediation. Hence there is an effect of semantic relatedness, a lexical factor, in English and Hebrew but not in Serbo-Croatian. Since the name code seems to be generated by first consulting a lexical entry in the two deep orthographies, this implies greater reliance on "direct" access and, by implication, more reliance on phonological mediation in Serbo-Croatian, consistent with the orthographic depth hypothesis. This interpretation fits the data and I cannot offer a better one. The problem is that it is hard to reconcile with other findings. First, there are the puzzling (but important) lexical decision results described above. Orthographic depth should also have affected performance on this task, especially given that it, rather than naming, is supposed to involve access of meaning. Second, if the Frost et al. account is correct, it should be possible to observe other effects of semantic information on naming in orthographies such as English, but this evidence is extremely difficult to come by. For example, one might expect factors such as concreteness of meaning or imageability to affect naming in English (as in the James, 1975, lexical decision studies). In a large-scale study of naming, Gloria Waters and I obtained latencies for 2900 words in English from 30 subjects (Seidenberg & Waters, 1989). For 530 of these words we have measures of abstractness/concreteness and imageability. The correlations between these factors and naming latencies are close to zero. Similar results have obtained for an unpublished set of data for 108 words collected by Paula Schwanenflugel. I have to stress that we have looked quite hard for these effects, using large samples of items and several measures related to semantics. We have also considered, and had to reject, more specific hypotheses concerning the role of semantics in naming (e.g., that the effects are specific to either very high or very low frequency words). These negative effects do not, of course, establish that naming is not mediated by semantics in English, and further research is required. At the same time, it is quite puzzling that these effects have not been observed as yet, given the Frost et al. interpretation of their priming results.

To summarize, the many studies of reading in Serbo-Croatian have not yielded a clear picture of the role of phonological recoding in this orthography. One might ask why this is so. Is this just another unhappy case in which theories in cognitive psychology are stronger and more elaborate than the methods available for testing them? Should we expect all studies employing these tasks to yield similarly ambiguous results? I think that the answer to both of these questions is no. There is no reason why further studies employing these same methods (or similar ones) cannot yield definitive answers. Rather than simply assuming that lexical decisions require access of meaning, however, it will be necessary to examine the conditions under which this assumption is valid. This will require obtaining more direct evidence concerning activation of meaning itself. By "direct" I mean evidence that a meaning-related factor has affected performance. Such evidence is provided by priming studies such as Frost et al.'s (1987) and Katz and

Feldman's (1983), in which activation of meaning is diagnosed by semantic priming effects (see also Baluch & Besner, 1991). Insofar as these experiments also employed the lexical decision and naming tasks, they suggest that the tasks themselves are not the limiting factor. Rather, it is necessary to couple them with appropriate stimulus manipulations. The pseudohomophone effects in the Van Orden (1987; Van Orden et al., 1988) studies also provide direct evidence that meaning has been activated, and therefore suggest an important direction for future cross-orthography research. Thus, although the data are at present somewhat cloudy, the methods for resolving the issues seem to be at hand.

2. English

Most discussions of the orthographic depth hypothesis assume that English is relatively "deep" because of irregularly-pronounced words such as HAVE, DEAF, and COLONEL, which do not occur in Serbo-Croatian (or in other shallow orthographies such as Spanish). In the standard dual-route framework, the meanings of exception words can only be accessed on a visual basis, because nonlexical spelling-sound correspondence rules (GPCs) will generate inappropriate, "regularized" phonological codes (e.g., HAVE rhymed with GAVE, DEAF rhymed with LEAF). Because there are no irregularly pronounced words in shallow orthographies, application of spelling-sound rules is guaranteed to yield the correct phonological codes for all words. These observations imply greater reliance on the direct, visual route in deep orthographies and greater reliance on the phonologically-mediated route in shallow orthographies, in accordance with the orthographic depth hypothesis.

This view has been challenged on both theoretical and empirical grounds. The theoretical challenge comes from recent accounts of how knowledge of spelling-sound correspondences is represented in memory and used in reading. The standard view is that this knowledge is represented in terms of rules. Such rules correctly specify the pronunciations of patterns such as GAVE, SAVE, and PAVE, and necessarily fail to correctly specify the pronunciation of exception words such as HAVE. Given this conception of the pronunciation rules, it follows that the meanings of exception words cannot be accessed by first deriving their phonological codes.

Connectionism provides a novel form of knowledge representation that is an alternative to the standard conception of rules. Knowledge is represented in terms of weights on connections between units. Thus, instead of a rule governing the pronunciation of -AVE, there are weights on connections between units in a lexical network that produce the correct input (orthographic) - output (phonological) mappings. Importantly, the same set of weighted connections can be used to generate both HAVE and GAVE. The reason this is possible is because such networks make use of distributed representations. The spelling pattern -AVE is not represented by a single unit, and therefore does not have to be assigned a single pronunciation. Thus, -AVE is pronounced /Av/ in the context of S- or P- but /av/ in the context of H-, with the same weights being used in all cases. I consider the invention of this type of knowledge representation (which is largely due to Geoff Hinton) to be a profound development. Until such systems were introduced, it was difficult—I would say impossible—to envision a system that simultaneously encoded both regular, "rule-governed" pronunciations and exceptional ones. One has to imagine what would have happened if, several years ago, someone had proposed a box-and-arrow model of word recognition (like the ones discussed in Coltheart, 1978 or Patterson, Marshall, & Coltheart, 1985) and arbitrarily assigned to a "grapheme-phoneme correspondence box" the capacity to generate correct pronunciations for both regular and irregular words. Such a proposal would have been perceived as incoherent, I am sure. However, that is exactly what models such as Seidenberg and McClelland's (1989) entail. There is a net that maps from orthography to phonology, producing correct output for all words; whether they are rule-governed or exceptions according to previous theories is irrelevant. Moreover, we do not have to conjecture whether this process will work; the net is implemented as a computational model that simulates human performance.

The development of connectionist theories of word recognition has important implications for the orthographic depth hypothesis. If the pronunciations of all words—including exceptions—can be generated by means of a "nonlexical" computation from orthography to phonology, then the meanings of all words can be reliably derived on the basis of phonological recoding, even in a putatively "deep" orthography such as English. Thus, recent connectionist models invalidate one of the primary assumptions on which the orthographic depth hypothesis is based, namely that irregularly pronounced words require a separate, "lexical" processing mechanism.

The challenge to the standard assumption that human knowledge is represented in terms of rules is, of course, an important part of the connectionist program. The extent to which different types of knowledge, previously conceptualized in terms of rules, can be explained in terms of connectionist types of knowledge representation is currently the focus of considerable attention. People are actively reassessing the assumption that different types of linguistic knowledge, in particular, are necessarily represented in terms of rules. To take one recent example, assigning the stress to syllables is invariably explained in terms of rules within standard linguistic theories, but Gupta and Touretzky (1991) describe connectionist models that perform this task without explicit rules. Whether such connectionist theories can explain all of the relevant facts ("achieve descriptive adequacy," in Chomsky's terminology) and whether they capture generalizations that standard rule-based theories miss are not yet known but under intensive investigation (for discussion of the controversial case of verb morphology, see Pinker, 1991, and Seidenberg, 1991). Certainly no upper limits on the explanatory power of the connectionist approach have been definitively established, though of course many unknowns remain (Seidenberg, 1989).

In the domain of word recognition, these issues arose with the development of the Sejnowski and Rosenberg (1986) and Seidenberg and McClelland (1989) models; Van Orden, Pennington, and Stone (1990), while not presenting an implemented model, discuss many properties of such systems and their potential relevance to word recognition. Seidenberg and McClelland showed that their model simulates a number of empirical phenomena concerning lexical decision and naming in English. Importantly, their model provides a unified account of some phenomena that are problematic for standard dual-route models, namely consistency effects, first studied by Glushko (1979), and later by Seidenberg et al. (1984), Jared and Seidenberg (1990) and Jared, McRae, and Seidenberg (1990). The standard dual-route account recognizes two types of words: regular (rule-governed) words such as MUST and exceptions such as HAVE. The studies mentioned above examined other words, such as GAVE, that are "regular" but "inconsistent" (in Glushko's terminology). GAVE is rule-governed, according to the standard Coltheartian approach, but it has an irregular neighbor, HAVE. The studies show that inconsistent words yield longer naming latencies than regular words, even though both are "rule-governed" according to the standard approach. Thus, words differ in the degree of consistency in the mapping between spelling and pronunciation; MUST is highly consistent, HAVE is inconsistent, and GAVE falls somewhere in between. Jared and Seidenberg (1990) show that these effects occur in multisyllabic words, and Jared et al. (1990) show that the results of more than a dozen studies in the literature can be explained in terms of a specific measure, the ratio between a word's friends (e.g., for GAVE, the rhyming -AVE words such as SAVE and PAVE) and enemies (for GAVE, the nonrhyming but similarly-spelled word HAVE).

The Seidenberg and McClelland model provides a simple account of these effects. Degree of consistency is encoded in terms of the weights on connections between units. This model correctly predicts effects of differing degrees of consistency observed in many studies. Thus, a single mechanism accounts for the processing of words that differ greatly in terms of degree of consistency; moreover, the model correctly predicts the intermediate cases, and the standard dual-route approach does not (for discussion of some ways in which the standard model could be modified in order to accommodate these phenomena, see Patterson & Coltheart, 1987. No modified dual-route model that actually produces the effects exists, however).

Connectionist models such as Seidenberg and McClelland's provide an alternative to the standard dual-route account, and the simulations in our 1989 paper suggest that the approach has some face validity. Though the model is not without limits (owing principally to the relatively small size of the training corpus and certain flaws in the phonological representation; Besner et al., 1990; Seidenberg & McClelland, 1990), it provides the most detailed account of the widest range of phenomena to date. The flaws in the model are ones that were built in from the start, and do not reflect any fundamental limitations of the approach.

Insofar as it affords the possibility of phonologically-mediated access of meaning even for "irregular" words, this type of model presents a challenge to the orthographic depth hypothesis. These theoretical developments have led researchers to return to the empirical question as to whether, in fact, word recognition is phonologically mediated in English, the putatively "deep" orthography. Empirical data from recent studies also present a challenge to the orthographic depth hypothesis. Van Orden (1987; Van Orden, Johnston, & Hale, 1988) addressed the issue of phonological mediation in an interesting away. Cognizant of problems that had arisen with the interpretation of lexical decision data, Van Orden used a semantic decision task in which subjects decide if a target word is a member of a pre-specified category. This semantic decision task, which demands the access of meaning, was combined with a manipulation of the phonological properties of the targets. On critical trials, the target was a homophone or pseudohomophone of a correct category exemplar. Thus, the category FLOWER might be followed by the homophone target ROWS; the category ARTICLE OF CLOTHING might be followed by the pseudohomophone SUTE. Van Orden and colleagues observed significant false positive responses on such trials, compared to appropriate nonhomophonic controls. These false positive responses would only occur if subjects had phonologically recoded the target stimulus and used the phonological code to access meaning. Because the target sounds like a category exemplar, subjects make false positive responses on a small but significant proportion of trials. The pseudohomophone results are particularly compelling, insofar as these nonwords could not be represented in lexical memory, and therefore could only activate meaning on a phonological basis. Thus, Van Orden et al. (1988) concluded that phonological mediation is obligatory in reading English words, echoing Turvey et al.'s earlier claim about Serbo-Croatian. If this claim is correct, it would appear to utterly refute the orthographic depth hypothesis.

It might be a mistake to write off the ODH on the basis of these results, however. The Van Orden et al. studies provide convincing evidence for phonologically-based activation of meaning. Considering how difficult it has been to develop methods that provide unambiguous evidence for phonological mediation, this is a considerable achievement. What the studies do not address is the range of conditions under which phonological mediation occurs. One question that must always be asked is whether the results are sensitive to aspects of the experimental design that do not carry over to normal reading. A second question is whether the effects occur for all words or only certain types of words. Many researchers have assumed that skilled readers utilize both "direct" and "phonologically mediated" recognition processes; which route dominates for a given word depends on factors such as frequency and reading skill. For example, McCusker et al. (1981), reviewing research to that point in time, concluded that recognition is direct for higher frequency words, and phonologically mediated for lower frequency words. Seidenberg et al. (1984) and Seidenberg (1985) made the same proposal. It has been widely assumed that phonological mediation occurs; the question is when. Van Orden et al. make the strong claim that phonological mediation is obligatory, but a broader range of stimuli and conditions would have to be examined in order to substantiate it.

Jared and Seidenberg (1991) provide additional evidence on this score. They describe six experiments using variants of Van Orden's methodology. They replicated Van Orden et al.'s basic finding of significant false positive effects for trials such as FLOWER-ROWS and ARTICLE OF CLOTHING-SUTE. However, several other important findings were also obtained, which indicate that the false positive rate in this paradigm depends on at least 3 factors. First, it depends on the specificity of the category. Van Orden has employed relatively specific categories such as FLOWER and PART OF A HORSE'S HARNESS, ones that allow subjects to rapidly generate a small number of common exemplars (ROSE, TULIP; REIN, BIT). Jared and Seidenberg (1991) used both narrow categories such as these and broader, more general categories such as INANIMATE OBJECT or VERB. These broader categories preserve the important constraint that subjects must access meaning in order to perform the task. However, they eliminate the possibility of generating targets in advance. The size of the false positive effect is greatly reduced when broader categories are used. Second, the effect depends on word frequency. Jared and Seidenberg manipulated the frequency of the target stimulus, using high frequency, low frequency, and pseudohomophone targets. With narrow categories such as Van Orden used, both high and low frequency targets, as well as pseudohomophones, produced larger false positive rates than nonhomophonic controls. With broader categories, only lower frequency targets and pseudohomophones showed the effect. Subjects were able to determine that a high frequency homophone foil is not a member of a category as rapidly as a nonhomophone control. Thus, the meaning of the homophone exemplar was not activated enough to cause any interference, suggesting that phonologically-mediated activation of meaning did not occur. In short, there was direct access for high frequency words. Finally, the false positive rate also depends on whether the homophone foil and exemplar are spelled similarly or not. Effects are larger for similarly-spelled pairs (e.g., PAIL-PALE) than for dissimilarly-spelled words (e.g., WAIT-WEIGHT). Van Orden (1987) also obtained this result. In summary, the false positive effects that Van Orden et al. took as evidence for obligatory phonological recoding actually vary systematically in response to properties of the experiment. The conditions that maximize the effects are narrow categories with lower frequency homophones whose spellings are not easily distinguished (and, indeed, may not be securely known by all subjects); the conditions that minimize the effects are broad categories with high frequency words and homophone pairs that are not spelled alike.

Jared and Seidenberg also examined effects of homophony on positive semantic decisions (i.e., latencies to decide that the target is an exemplar of the category—FLOWER-ROSE). These experiments eliminated the use of anomalous trials in which a target sounds like a member of the designated category but is not in fact an exemplar. Our thought was that including a large number of trials like FLOWER-ROWS might cause subjects to become aware of the homophone manipulation and therefore change their response strategies. We also factorially manipulated the frequencies of both the homophone exemplar (e.g., for FLOWER, ROSE) and homophone foil (e.g., FLOWER, ROWS). The frequency of the member of the homophone pair that is *not* presented on a trial provides important evidence as to whether subjects engage in a spelling check, as Van Orden (1987; Van Orden et al., 1988) proposed. The results of these experiments confirmed our earlier findings: with broad categories, lower frequency targets showed an effect of homophony, whereas high frequency targets did not. When the stimuli in the experiment contained the anomalous trials, subjects did engage in a spelling check for lower frequency words. However, when these trials were eliminated, so was the spelling check. Hence, the spelling check seems to be a response to demand characteristics of the experiment rather than part of the normal word recognition process.

These results narrow the scope of the Van Orden effects considerably. Rather than suggesting that phonological mediation is obligatory, they suggest that its occurrence depends on word frequency, with lower frequency words showing larger effects. This conclusion is in accord with earlier proposals by McCusker et al. (1981) and Seidenberg et al. (1984; Seidenberg, 1985).

If this picture is correct, there is still some life in the orthographic depth hypothesis. Assume, for the moment, that Turvey et al. are correct and that phonological mediation is obligatory in Serbo-Croatian. Then the Jared and Seidenberg results suggest that there is less phonological mediation in reading English, as suggested by the ODH. Two things need to be determined before these conjectures could be taken as fact. First, experiments like Jared and Seidenberg's need to be performed in Serbo-Croatian (as well as in Hebrew). There is a need for studies examining effects of phonological mediation using both words that differ in frequency and a semantic decision task. At this point, we simply do not know whether, as in English, the extent to which phonological mediation occurs in Serbo-Croatian depends on frequency. My own guess is that it does (as discussed further below). The Seidenberg and Jared studies suggest that common words in English are recognized on a visual basis. I see no reason why the shallowness of the orthography should prevent this from occurring in Serbo-Croatian as well. This is an empirical question, of course, and the methods for addressing it are available. A second issue that needs to be addressed is whether Jared and Seidenberg's conclusions are correct. It is possible that there is phonological mediation for higher frequency words and that existing methods are simply not sensitive enough to reveal them. High frequency words are recognized very rapidly. Effects due to phonological mediation are likely to be small and difficult to detect; therefore, some caution is certainly warranted before concluding that there is no phonological activation of meaning for higher frequency words. All that can be said at this point is that, using the same methods and experimental logic as in the original Van Orden (1987; Van Orden et al., 1988) studies, there was no evidence of phonological mediation for higher frequency words, whereas such evidence was obtained for lower frequency words. These results certainly suggest that it would be valuable to develop further refinements of these methods that might be capable of picking up very small effects if they occur.

To summarize, the Van Orden results present an important challenge to the orthographic depth hypothesis. However, the Jared and Seidenberg studies suggest that there is not as yet a definitive answer to the question of phonological mediation in English, and point to a need for similar studies in Serbo-Croatian. Assuming that progress continues at the rate it has over the past few years, the answers will, I believe, finally emerge.

3. Hebrew

The other major writing system that has been extensively studied in connection with the orthographic depth hypothesis is Hebrew. At the outset I should acknowledge feeling less confident about how Hebrew is read than either English or Serbo-Croatian. Nonetheless, quite a bit can be discerned from existing studies. The basic facts about written Hebrew are well known. Almost all Hebrew letters are consonants. Vowels are represented by diacritical marks that may be located above, within or below letters. Children are taught to read "pointed" script that contains these explicit vowels. The pointed Hebrew script is shallow: the mappings between pointed letter strings and pronunciations are entirely regular and consistent. However, the vowels are typically omitted in texts for skilled readers (e.g., newspapers, books). Unpointed words are often ambiguous; different words are formed depending on which vowels are (mentally) added to a given consonantal root. Thus, each root can be seen as generating a neighborhood of lexical items. Some of these lexical items will be semantically and/or morphologically related but others will not. Rough analogies in English are provided by the consonant string SNG, which generates the neighborhood SING/SANG/SUNG/SONG, all of which are morphologically and/or semantically related, and the consonant string DG, which generates two morphologically related words (DIG and DUG) and one that is semantically and morphologically unrelated (DOG).[2]

These consonantal roots differ from words in English in an important way: they are not merely ambiguous between different lexical items, they are vague in the sense of failing to completely specify *any* lexical item. Hence they are compatible with multiple words associated with multiple pronunciations. In contrast, the written forms of words in English do completely specify words; thus they can be unambiguously associated with pronunciations. The reader merely has to know which pronunciation is associated with each written word. The letter pattern BOOK contains sufficient information to permit it to be distinguished from other words in the language; the letter pattern BK does not. Therefore, if the word BOOK is in my vocabulary, I can identify the associated pronunciation. There are a small number of homographs such as WIND or BASS that are associated with two pronunciations, but the point is the same. Hebrew consonantal roots, in contrast, do not fully specify any pronunciations, and are often compatible with several different ones. I will refer to this dimension as *orthographic transparency*. English and Serbo-Croatian are orthographically transparent; Hebrew is not. These observations

---

[2]These simplified examples are for illustrative purposes only. I am ignoring, for example, the fact that DG also generates words such as ADAGE or DOGE.

strongly imply that word recognition and pronunciation in Hebrew must be more dependent upon contextual information than in English. Again, using a rough English analogue, the word DIG is pronounced the same in all contexts. Pronouncing DG, however, would depend on whether it appears in the context JOHN BEGAN TO DG compared to JOHN PETTED HIS DG. Of course, to get a full sense of the problem, one would have to omit the vowels from all of the words in the context as well (JHN PTTD HS DG) and also recognize that contextual constraints are probabilistic (the sentence could be JOHN BEGAN TO DOG HIS RIVAL).

Hebrew exhibits unique characteristics that make it difficult to relate to the orthographic depth hypothesis. One view would be that unpointed Hebrew is the extreme example of a "deep" orthography, insofar as the pronunciation of each word is highly underdetermined by its orthographic representation. Hence it might be assumed that the access of meaning will not be phonologically mediated. The orthographic depth hypothesis emphasizes the extent to which an orthography specifies phonology. However, this ignores the other important feature of written Hebrew: unpointed words are orthographically opaque. As such, recognition on a direct, visual basis would seem to be equally problematical.

There have been several proposals as to how Hebrew is read. Frost et al. emphasize the greater degree of "lexical" involvement in recognizing and pronouncing Hebrew words, in contrast to shallower orthographies that afford "nonlexical" mechanisms. As I noted previously, however, "lexical involvement," as indexed by frequency and lexicality effects, is ambiguous as to its source. It could reflect an orthographic-phonological computation that is sensitive to frequency of exposure, as in the Seidenberg and McClelland model, or it could implicate a nonphonological route mapping from orthography to semantics. Moreover, this greater degree of "lexical" involvement does not necessarily implicate a nonphonological recognition process, insofar as the mapping from orthography to semantics seems equally indeterminate. In more recent studies, Frost (1991) has suggested that word recognition in Hebrew relies on phonological mediation. His view is that subjects recognize a consonantal root on a visual basis, phonologically-recode (possibly with the help of contextual input), and use the derived phonological code to access meaning. If this view is correct, it suggests quite a different picture than the orthographic depth hypothesis in its original form. Specifically, Hebrew, the quintessential "deep" orthography, is now thought to involve phonological mediation, just as Van Orden et al. (1988) have proposed for English and Turvey et al. (1984) for Serbo-Croatian.

In summary, orthographic depth does appear to affect how Hebrew is read, though perhaps not in the manner suggested by the orthographic depth hypothesis. The effect seems to be realized in terms of the degree of reliance on contextual information in identifying words, rather than in terms of the degree of reliance on phonology.

An alternative to orthographic depth: Equitable division of labor

What is the status of the orthographic depth hypothesis in light of the research reviewed above? Research on all three of the orthographies that have played a crucial role in the development of the ODH has raised questions about its validity.[3] Clearly, there remain a

[3]I do not have space to consider research on Japanese and Chinese bearing on the ODH (see Besner, this volume for discussion of some of this work), although I do not think it changes the overall picture very much. There seems to be evidence for rapid activation of phonological information in both writing systems, even in the "deep" Chinese orthography and in Japanese Kanji (see, for example, Perfetti &

large number of empirical issues that are simply unresolved concerning, for example, the role of direct access in each orthography and the extent to which there are true differences between orthographies with respect to the degree of phonological involvement in visual word recognition. Still, it appears that the empirical phenomena are starting to point in a very different direction than the ODH, namely toward similarities between alphabetic orthographies in terms of how they are processed—all seem to involve heavy reliance on phonology—despite substantial differences in "depth." In this section I will attempt to situate the question of orthographic depth in a somewhat broader theoretical context. My main point is not that the orthographic depth hypothesis is mistaken; to the contrary, I think there is a basic sense in which it has to be correct. Orthographies do differ in terms of the extent to which they encode phonological information and this probably does have an impact on the extent to which meaning is activated by phonology. Rather, I will suggest that the problem with the orthographic depth hypothesis is that it is too narrow. It focuses on one property of orthographies—transparency of phonological encoding—to the exclusion of other properties, including ones that may have more impact on processing difficulty. Moreover, it ignores the broader range of perceptual and computational constraints that determine how words are read. In doing so it overstates how much orthographies differ in terms of processing and overlooks some deeper generalizations about how word recognition is achieved. In sum, the orthographic depth hypothesis could be true but (a) contribute little to functional differences between orthographies in terms of how they are read, and (b) miss the bigger picture, which seems to be about commonalities in the reading process despite seemingly radical differences among orthographies.

Instead of focusing on properties of orthographies, it might be useful to begin by analyzing the problem of word recognition itself. The goal is to determine the contextually-appropriate meaning of a word. The computation is one that takes a visual pattern as input and yields a meaning representation as output. The characteristics of this computation are determined by relationships among the various lexical codes: orthographic, phonological, semantic. There are differences among writing systems with respect to the correspondences between orthographic and phonological codes. There are similarities between orthographies with respect to relationships among the other codes. In alphabetic orthographies, the mapping between orthography and meaning is essentially arbitrary, at least at the level of individual morphemes.[4] The mapping from phonology to meaning is arbitrary in exactly the same sense. These relationships among codes are illustrated in Figure 1. Different tasks (reading, naming, spelling, etc.) involve different computations among the codes ("routes," "routines," or "pathways").

Figure 1 represents a framework or "architecture" hypothesized to underlie lexical processing in all languages. It is important to distinguish between this broader framework and the Seidenberg and McClelland model, which was an initial attempt to implement part of this system. The implemented model is limited in various ways that are theoretically important. One is that the model is not a real-time, dynamic system: activation is computed in a single step, rather than cascaded over time. Second, the

---

Zhang, 1991; Seidenberg, 1985a; Besner & Hildebrandt, 1987). It seems likely that this information contributes to the activation of meaning in both cases.

[4]Like everyone else, I am ignoring clusters such as CLASP, CLING, CLUTCH. Morphologically complex words can also be seen entailing a type of representation in which the relationship between orthography and meaning is not arbitrary. Thus, the relationship between the spelling pattern PRE- and its meaning is arbitrary, but PRE- contributes in a systematic way to the meanings of words such as PREVIEW and PRECOMPILE.

implemented model lacks the property of interactivity (see McClelland & Rumelhart, 1981, for a model that does not). The computation from orthography to hidden units to phonology is strictly feed-forward. There is a feedback loop from the hidden units back to orthography, but activation is not allowed to cycle multiple times, and this feedback loop is isolated from the computation of phonology. Moreover, we did not implement the isomorphic feedback loop from phonological output to hidden units. In addition, there are no interconnections between the units within a layer; hence there is no mechanism for the activation of phonological units to influence each other.Third, the model is entirely deterministic. There is no variability in the computation of output representations; given a particular input and a particular set of weights, the same output is always be computed. This contrasts with models such as Cohen, Dunbar, and McClelland (1990) in which output is stochastic because a random factor is added to the activation function. This seems more in keeping with the variability of actual human performance.

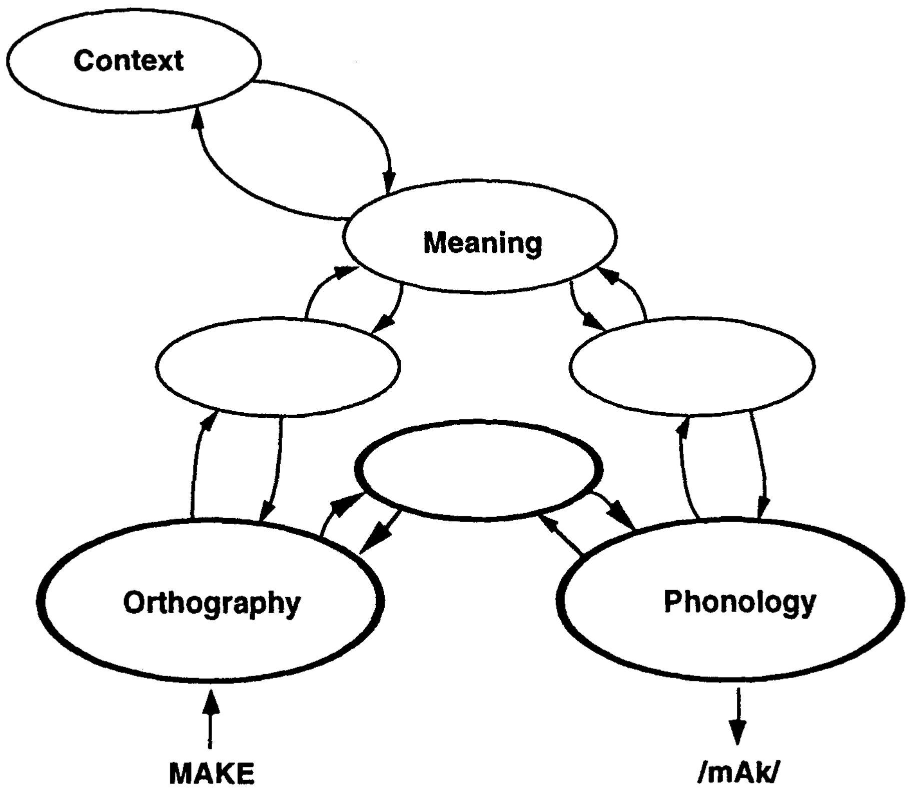

Figure 1. **A framework for thinking about lexical processing. Seidenberg and McClelland's (1989) implemented model is indicated in bold outline.**

In the discussion that follow I envision a model that incorporates all of these properties. Interactivity is especially important. Interactivity among units within a layer permits the formation of attractor basins (Hinton & Shallice, 1990; Plaut, 1991), which are to PDP models what lexical entries are to more traditional models. Interactivity between layers of units permits a network to sharpen or refine patterns through feedback (Plaut, 1991).[5]

The relationships among lexical codes illustrated in Figure 1, which seem to be characteristic of all languages for which there are conventional written codes, imply two ways to achieve the basic goal of determining the meaning of a written word: either computing the meaning directly from orthography or indirectly by means of computations from orthography to phonology and then phonology to meaning. That these two alternatives are available is the basic insight of "dual-route" models, and it seems unassailable to me.[6]

The important differences between the connectionist approach I have been advocating and earlier dual-route models derive from the use of distributed representations. I have already noted some consequences of the fact that this obviates the need for pronunciation rules. Other consequences follow from the use of distributed representations of meaning (Seidenberg, 1990). There is a set of units encoding semantic primitives; the meaning of a word corresponds to the pattern of activity over these units. Whereas in previous accounts meanings have been treated as fixed entities that can be "accessed" from memory, the alternative view suggests that meanings are computed each time a word is encountered, with the patterns of activation varying across instances. For example, a context that picks out a particular feature of a word's meaning may lead to increased activation of that feature (Tabossi, 1988).

According to this theory, codes are not accessed, they are computed; semantic activation accrues over time, and there can be *partial* activation from both orthographic and phonological sources. So, for example, whereas in the standard dual-route model, "phonological mediation" required deriving the complete phonological code for a word and using it to search lexical memory, in the present framework there can be partial activation of phonology from orthography, or of meaning from phonology. Thus, the meaning of a word is built up by means of activation from both routes (indeed, from three "routes" if contextual influence is included, and from four if there is the additional route discussed in footnote 6), rather than accessed by means of whichever route wins the race. I will continue to use the term "phonological mediation" because it is convenient, but

---

[5]One view of the Seidenberg and McClelland (1989) model is that it illustrates about how well a lexical network can do without interactivity. Specifically, a simple feed-forward net can learn to generate correct output for almost the entire training set and generalize to simple nonwords such as NUST. It is not that the model produces incorrect pronunciations for difficult nonwords such as JINJE or FAIJE (McCann & Besner, 1987); rather, it produces very unclear output that does not match *any* pronunciation very well. Interactive spread of activation between output and hidden units could be expected to shape and clarify these patterns.

[6]What *is* assailable about the standard model are its claims about how the routes work (i.e., pronunciation rules; lexical lookup, etc.). Also, it is confusing, but nonetheless true, that most standard "dual-route" models actually have three routes (Coltheart, 1987). They are (a) the nonlexical, GPC route; (b) the lexical route through semantics, and (c) a nonsemantic, lexical route that involves direct connections between orthographic and phonological word forms. Our model does not contain anything like this third route. There is some evidence for this route (e.g., Funnell, 1983), though it is by no means decisive. Funnell, for example, argues that routes (a) and (b) were impaired in her patient, who could nonetheless name some words aloud. The principal question is whether a partially impaired semantic route could nonetheless support naming of some words.

what is really meant is "partial activation of meaning on a phonological basis." Phonological mediation in this sense can co-occur with activation of meaning on a visual basis. Assuming that the encoding of a word initiates activation along all pathways in parallel, the question of direct vs. phonologically mediated activation of meaning boils down to whether, for a given word in a given orthography, there is sufficient time for activation to spread from phonological units to semantic units before the computation from orthography to semantics settles on the meaning of a word.

How are the resources that such a system makes available actually allocated in reading? The ability to read is acquired rather than innate. I suggest that the system that supports this capacity develops in such a way as to realize what I will term an equitable division of labor between the routes. The task of recognizing words and determining their meanings is allocated between the routes in such a way as to permit the task to be performed efficiently. The hypothesis to be developed here is that this division of labor develops under a variety of constraints. Orthographic depth represents one important class of constraints, but others have an impact as well. The tradeoffs among these constraints are such that the processing system tends to converge on a very similar division labor across writing systems. Exactly *what* the division of labor is for various orthographies is an empirical question, one that I have suggested is not as yet settled. The data are pointing toward smaller differences between orthographies with respect to the division of labor, however, than the orthographic depth hypothesis implied. The principal theoretical question concerns the factors that affect this division of labor, that is, *how* it is achieved.

In standard dual-route models, the division of labor in English is governed by the assumption that only the "lexical" route can handle exception words. Since the reader does not know in advance whether a word is an exception or not, the two routes are tried in parallel, with a race between them. The fastest-finishing process permits the "access" of a lexical entry. This will always be the "lexical" route for irregular words. I have suggested that this argument is invalidated by recent connectionist models. Various other arguments—I will call them "computational plausibility arguments"—have also been offered as to why one or the other route would necessarily predominate for skilled readers. For example, the direct route is often assumed to be more efficient, because it involves a single computation (from orthography to semantics) rather than the two computations (orthography to phonology; phonology to semantics) of the indirect route. Certain methods of reading instruction (e.g., "see and say") take this assumption for granted. Conversely, Van Orden et al. (1990) argue that the indirect route is necessarily more efficient. They note, for example, that whereas the mapping from orthography to semantics is essentially random, the mapping from orthography to phonology is systematic, something that can be picked up by "covariant learning" techniques (i.e., certain connectionist learning algorithms). However, it can also be noted that, whereas the mapping from orthography to phonology is systematic, the mapping from phonology to semantics, the other part of the "indirect" route, is as arbitrary as the mapping from orthography to semantics. Insofar as both routes involve learning an arbitrary mapping, but only the indirect route also requires orthographic-phonological translation, we are back to the argument that the direct route is more efficient. Van Orden et al. anticipate this rejoinder, noting that while the mapping between phonology and meaning is arbitrary, it is highly overlearned, because of its role in speech perception. The net result, they claim, is that phonological mediation will be more efficient than direct access, and therefore predominate for skilled readers. However, these arguments ignore the cost associated with a recognition system based exclusively on phonological information, namely the problem of disambiguating homophones, a serious problem given the large number of homophones in English.

As the structure of the preceding paragraph suggests, these arguments based on computational plausibility are not very decisive. Van Orden et al.'s arguments about computational efficiency, for example, are abstract rather than based on actual computational results. This kind of argument does not place any upper limit on peoples' capacities to learn a direct mapping from orthography to semantics. At the same time, it cannot be concluded that a recognition pathway with one component computation (the direct route) is necessarily more efficient than a recognition pathway with two component computations (the indirect route) without being more explicit about the computational properties of each route. One can get to the Plaza Hotel from Grand Central Station more quickly by taking a subway and a bus than by walking (on the other hand, a cab is faster than both—usually). Settling these issues will require the development of more realistic computational models.

Rather than assuming that one route will necessarily be more efficient than the other, the approach I am advocating emphasizes understanding the factors that result in a division of labor between the routes that permits word recognition to be accomplished efficiently. One such factor is orthographic depth. The more systematic the correspondences between spelling and sound, the easier it will be to compute phonological codes. Other things being equal, the easier it is to compute phonological codes, the more likelihood that meaning will be activated on this basis. The key phrase here, however, is "other things being equal," which they rarely are in the case of orthographies. Note, for example, that it is difficult to determine the extent to which orthographies differ in terms of the complexity of the mapping between orthography and phonology because of other ways in which they differ. It is notorious, for example, that English has a large degree of inconsistency at the level of grapheme-phoneme correspondences, in contrast to Serbo-Croatian. The irregular words tend to cluster among the higher frequency items in the language, however, and both the Seidenberg and McClelland model and the empirical data suggest that these words can be pronounced as easily as entirely regular words. Moreover, English has a large pool of monosyllabic words, which include most of the exceptions. Serbo-Croatian, in contrast, is more consistent at the level of grapheme-phoneme correspondences. However, it has fewer monosyllabic words. In addition, the system for assignment of syllabic stress is quite complex. If correct assignment of syllabic stress is required in order to achieve phonological activation of meaning, this might restrict the use of phonological mediation considerably. The net result is that, despite the differences at the level of graphemes and phonemes, it is by no means obvious that it is easier to generate phonological codes for words in one of the orthographies.[7]

---

[7]Linguists inform me that the stress patterns for some words in Serbo-Croatian are lexically determined. That is, they cannot be assigned without already having identified the word. This would seem to present a problem for the assumption that meanings are activated exclusively on the basis of phonology in Serbo-Croatian (and there are analogous effects in English; e.g., im´pact vs. im`pact; blockage vs. blockade), at least if one holds the standard view that stress is assigned by rule, with exceptions "listed" in the lexicon. Of course, in the kind of model I am proposing, it would be possible to encode idiosyncratic facts about the stress patterns of individual words within the orthographic to phonological route; partial activation of meaning from phonology might also be possible even if the complete stress pattern were not computed. In either case, it would be important to gain additional empirical evidence as to how much phonology is computed in reading Serbo-Croatian words, and how cases where the assignment of stress is problematical are treated.

The important results of Sebastián-Gallés (1991) raise questions about how knowledge of spelling-sound correspondences is represented even in very shallow orthographies. Her experiments were conducted in Spanish, which is said to be shallower than Serbo-Croatian because both grapheme-phoneme correspondences and stress are highly predictable. Sebastián-Gallés exploited the fact that the pronunciation of the two letters *c* and *g*, while entirely rule-governed, depends on the contexts in which they occur. For example, *g* is always pronounced /g/ when followed by *a* or *o*, and /x/ when followed by *e* or *i*. She created consistent and inconsistent nonwords in the following way. The consistent pseudoword **encogedo** was derived from the word **encogido.** According to the rule governing *g*, it should be pronounced the same in both stimuli. The inconsistent pseudoword **arrugedo** was derived from me word **arrugado.** The same rule governs the pronunciation of the *g* in **arrugedo** as in **encogedo**. However, a different rule applies to the *g* in **arrugado.** If subjects pronounce nonwords by applying the rules, **encogedo** and **arrugedo** should be equally easy to pronounce. Sebastián-Gallés actually found that inconsistent nonwords like **arrugedo** were harder to pronounce than consistent nonwords such as **encogedo.** Thus, in pronouncing **arrugedo** subjects were affected by the lexical neighbor **arrugado,** evidence that they were not simply applying grapheme-phoneme correspondence rules. This result represents a replication of the Glushko inconsistency effect in a shallow orthography with completely regular spelling-sound correspondences. This consistency effect, of course, would be easy to derive in a connectionist model like Seidenberg and McClelland's; however, it is incongruent with the idea that nonwords are pronounced by applying spelling-sound correspondence rules. In summary, this research suggests that even in orthographies where grapheme-phoneme correspondences can be described in terms of rules, knowledge of these correspondences may not be represented in this form. If readers of these orthographies are not merely applying rules, it cannot be assumed that it is easier to generate phonological codes in these cases. Rather, the extent to which pronunciation is affected by inconsistent (but rule-governed) neighbors would have to be assessed.

Other properties of orthographies need to be considered as well. One that I have identified is orthographic transparency—the extent to which an orthographic pattern specifies the identity of a lexical item. Serbo-Croatian and English are both orthographically transparent, but Hebrew is orthographically opaque. Here too there are tradeoffs that may affect the division of labor. Because it is orthographically transparent, English has a larger number of distinct orthographic patterns than does Hebrew, with its consonantal roots. Learning to recognize the large number of distinct visual patterns in the English vocabulary on a strictly visual basis may be more difficult than learning to recognize the relatively smaller number of roots in Hebrew. Word recognition in Hebrew might exploit this fact (indeed, there doesn't seem to be any other way of recognizing the consonantal roots short of letter-by-letter reading). The cost, of course, is that recognizing a consonantal root is not equivalent to recognizing a word; additional processing is required.

There may be other bases for tradeoffs between the routes. Consider the case of English again. The recognition system has to be organized so as to deal with both regular and irregular words. As the Seidenberg and McClelland model suggests, a single route can generate phonological codes for both, affording the possibility of phonological mediation across the board. However, it might be asked whether this is either a necessary outcome or the most efficient one. It might be more efficient to allocate responsibility for some words to the direct route—in particular, some of the lower frequency exception words whose orthographic-phonological correspondences are difficult to master. Removing exception words from this computation improves its efficiency, as the simulation of "shallow English" suggested. Thus, it could be that many exception words are handled by means of the "direct" route, not because the indirect route is incapable of generating correct phonological codes for these words (as the GPC notion suggested), but rather because moving responsibility for some of the exceptions out of the indirect route improves its efficiency on the items that *are* its responsibility.

The existing Seidenberg and McClelland simulations are also suggestive in this regard. With 200 hidden units, the model is able to generate correct phonological codes for about 97% of the words in its training corpus. The items that it misses are mostly low frequency words with exceptional spelling-sound correspondences (e.g., BLITHE; BREADTH; COAX). Given the architecture we employed, the model did not have the resources to encode the correct output codes for a small number of irregular words, ones that were presented very seldom during training. This architecture nonetheless allows us to simulate quite a large number of other behavioral phenomena. It appears that the orthographic to phonological computation in the model achieves approximately the same level of efficiency as in people by allocating responsibility for a small number of very difficult items to the other route.

Of course, the simulation results depend on the choice of network architecture, specifically the number of hidden units that are employed. With a larger number of hidden units, the model could encode even the lowest frequency irregular words. The cost would be a decrement in the model's (already limited) ability to generalize (i.e., process nonwords). With a smaller number of hidden units, more of the exceptions would have to be handled by the direct route, leaving the orthographic to phonological computation to handle regular words and nonword generalization.

These observations suggest that there may be considerable plasiticity in the division of labor between the routes. The redundancy between the routes affords the possibility that readers might converge on somewhat different solutions, depending on an independent factor, the resources available for component computations. I believe that it will turn out that there are important individual differences in the division of labor between the routes within a given orthography, an interesting topic for future research. In English, for example, it appears that differences among skilled readers are manifested not in their ability to read familiar words, but rather in their facility with nonwords. Both routes are capable of handling words, but only the orthographic to phonological computation is relevant to nonwords. Moving more of the exception words to the "lexical" route would lead to better performance on nonwords. Accommodating more of the exceptions within the orthographic to phonological computation would lead to worse performance on nonwords. There is some very suggestive, though highly preliminary, evidence from neuroimaging studies that some forms of dyslexia may be related to a failure to allocate sufficient neural resources to parts of the reading task (Hynd & Semrud-Clikeman, 1989). These data are consistent with the idea that there may be significant variation between readers in terms of the computational capacities of different "routes."

In summary, there may be important individual differences in the division of labor within a given orthography, owing to variation in available computational resources. Of course, the orthographic depth hypothesis is simply a generalization of this observation to the case of orthographies rather than individuals. There may be important differences between orthographies in the division of labor, owing to variation in properties such as transparency of orthographic-phonological correspondences, the number of orthographic vs. phonological word forms, and other factors. The processing style of a given reader will then depend on interactions among these factors.[8]

As in other discussions of orthographic depth, I have largely focused on factors that affect the orthographic to phonological conversion process. The same issues arise in connection with the direct, orthography to semantics computation, of course. The computational capacities of this route are not well understood. It is clear why word recognition in Hebrew cannot rely exclusively on this computation: word forms in the written language are opaque. It is not obvious, however, why word recognition in the English orthography could not be handled entirely by this process, obviating the problem of computing phonology entirely. After all, people are able to recognize a huge number of objects on a "direct," visual basis, without first generating their phonological codes; the number of familiar objects that people can recognize in this manner is surely larger than the number of words in their vocabularies. Given that word forms in English clearly determine lexical items, why isn't word recognition simply another species of object recognition? It is sometimes argued that phonological mediation is efficient because it exploits existing knowledge that is used in speech recognition. However, word recognition might also exploit existing pattern recognition capacities that are used every day in object recognition. What are the factors that limit the efficiency of this route, thereby forcing a sharing of effort with the phonological route?

I cannot offer a definitive answer to this question, but the relevant factors seem to include the following. The visual word forms of a language constitute a certain stimulus space. The complexity of the visual recognition process depends in part on the characteristics of this space (which vary across orthographies, of course). Factors such as the number of word forms, their complexity, and their frequency determine the relative discriminibility of items from each other, given human perceptual and cognitive capacities. This stimulus space appears to differ from the one for objects. In particular, there may be fewer word forms to recognize than objects, but they are more similar to one another and therefore harder to differentiate. Moreover, the forms of objects convey information about their origin, function, and relationships to other objects, which may facilitate recognition. Word forms convey little of this information. At the same time, it may be possible to recognize some word forms on the same basis as objects; this seems to be characteristic of the very early word recognition processes of "logographic" readers (Frith, 1985) and may be something that is achieved on a wider scale by skilled readers.

The efficiency of the direct route depends in part on human learning capacities, which are not well understood. Van Orden et al. (1990) argued that "covariant learning" procedures will have an easier time with the systematic correlations between orthography and phonology in English than with the largely arbitrary mapping from orthography to semantics. Taken with the fact that people are very good at deriving the meanings of words from phonology, these observations suggested to the authors why phonological mediation must be the rule. However, it is by no means obvious what the limits are on the capacity to learn a mapping like the one between orthography and semantics in English. It is certainly the case that some connectionist learning schemes can master such arbitrary mappings, though whether they have any psychological plausibility remains to be

[8]One question that must be left for future research is how the variability with respect to the division of labor that is observed among readers of English compares to the variability that is observed between orthographies.

determined. The Plaut (1991) model, for example, learns to map from orthography to semantics, and then from semantics to phonology, though on an admittedly small scale. Moreover, it is possible that the arbitrariness of the mapping between orthography and semantics has been overstated, insofar as morphological structure is taken to represent an intermediate level at which the correspondence to meaning is non-arbitrary.

Clearly there are quite a few unknowns that make it difficult to be certain as to how the division of labor is accomplished among different orthographies and individuals. I hope to have conveyed the idea that what is important is to determine the computational demands imposed by an orthography—all of them, not just "orthographic depth"—as well as the computational capacities of the reader, both of which vary. Of course, determining the actual division of labor that is achieved is an empirical question. At the risk of being wrong—because so many of the relevant factors are still poorly understood—my current guess is that something like the following will prove to be correct. Human cognitive and perceptual capacities being what they are, there is "direct" recognition of higher frequency words in all alphabetic orthographies that exhibit the property I have termed orthographic transparency (e.g., English and Serbo-Croatian; in Hebrew there is direct access for the consonantal roots). The size of this pool of words varies depending on the size and density of the space of orthographic word forms, on the efficiency of the orthographic to phonological conversion process afforded by the orthography (i.e., orthographic depth), and on the skill of the reader. In English, this pool seems to comprise relatively few types but they account for many of the tokens that are encountered in reading. The distribution of words by frequency is highly skewed, with a small number of lexical types accounting for a large proportion of the tokens. Some relevant statistics are summarized in Table 1.

Table 1

Some data concerning the most common words in English.

Entire Kučera and Francis (1967) corpus:*

Tokens: 998,052
Types: 46,369

150 Most Frequent Words:

Length in Letters	Number of Types	Σ Frequency of Tokens	Cumulative % of Entire Corpus
1	2	28,421	2.8
2	24	168,282	19.7
3	38	190,405	38.8
4	52	91,157	47.9
5	25	31,892	51.1
6	4	3,584	51.5
7	5	3,891	51.9
Total	150	517,632	51.9

*These calculations are based on an on-line version of the Kučera and Francis corpus of unknown provenance that contains 998,052 tokens.

The 150 most frequent words in the Kučera and Francis corpus account for over 50% of the approximately 1,000,000 tokens in that database (my on-line version of the Kučera and Francis corpus contains 998,052 tokens). The 150 most frequent words account for 51.9% of these tokens. The distribution of these 150 words by length is given in the table. There are 64 words from 1-3 letters in length; they account for 387,108 tokens, which is 38.8% of the entire corpus. There are 116 words from 1-4 letters in length, yielding 478,265 tokens, which is 47.9% of the corpus. Thus, almost half the words in the corpus consist of the 116 most frequent 1-4 letter words. It does not stretch the bounds of plausibility to consider these words to be likely candidates for direct access. Nothing in my knowledge of human capacities suggests that they should be unable to recognize 100-150 simple patterns by sight. I do not mean to suggest that the direct recognition process is limited to these 150 words; the range of cases in which direct access is used is simply unknown. Rather, these numbers represent plausible lower bounds.

Of course, it could also be true that these words are easy to pronounce as well, in which case that they might be rapidly recognized on a phonological, rather than orthographic, basis. The homophony problem that would result from a strictly phonological recognition process—over 25% of the 150 most frequent words are homophones such as I and RIGHT—would have to be resolved by means of some kind of spelling check (Van Orden et al., 1988, though see Jared and Seidenberg, 1991, for evidence that the spelling check only occurs in response to specific experimental conditions). The fate of these items remains to be determined conclusively. However, the Jared and Seidenberg (1991) studies certainly suggest that at least some higher frequency words are recognized on a direct basis. The point of the statistical considerations is that the capacity to recognize even a small number of lexical types on a visual basis would entail a division of labor in which quite a lot of responsibility fell to the direct route because of the type-token facts. If these considerations are valid—and it remains to be seen if they are—I see no reason why they would not apply in Serbo-Croatian as well. Nothing about the orthography would constraint the reader to a phonologically analytic strategy any more than they are so-constrained in recognizing objects. In Hebrew, it may be the consonantal roots that are recognized on a visual basis, but, as I have noted before, converging on the identity of a lexical item appears to depend on input both from phonology and context.

The size of the pool of words recognized on a visual basis may, of course, be an individual difference variable related to reading skill, which may in turn depend in part on the computational resources allocated to this computation and on properties of the orthography. Still, it appears that the pool of words that can be recognized on a visual basis is limited; that is what is suggested by the various studies of English, Serbo-Croatian and Hebrew mentioned above, which indicate at least some activation of meaning via phonology under a broad range of conditions. Aside from the factors that make this an efficient division of labor, some of which were discussed above, there are at least two other reasons why phonology may play such a prominent role. One is because of its important and widely-discussed role in the acquisition of reading skill (Adams, 1990). A second, less obvious, reason may be because phonological recoding facilitates the recognition of longer, morphologically complex words. Such words may be recognized in terms of subunits that are recovered from left to right (in languages such as English). Phonological recoding may facilitate the retention of parts of words while attention shifts to subsequent parts. It has often been suggested that word-level phonology is relevant to the use of working memory in sentence parsing (e.g., Waters, Caplan, & Hildebrandt, 1987); here I am suggesting that it may be relevant to the parsing of words as well.

The final issue that needs to be addressed concerns strategies. Assuming that a functional division of labor is achieved, can the reader flexibly modify the allocation of resources depending on task demands or instructions? A number of researchers in the dual-route tradition have assumed that skilled readers can strategically allocate attention to one or the other route. Recent studies by Baluch and Besner (1991) and Monsell et al. (in press) are consistent with this view. In a study of Persian, which contains both "deep" and "shallow" elements, Baluch and Besner (1991) obtained evidence interpreted as indicating that subjects strategically changed the attention to the direct or phonologically-mediate route in response to properties of the stimuli in the experiment (principally, presence or absence of nonwords). An alternative interpretation of such effects, however, is that, rather than changing their strategies for recognizing words, subjects change their strategies for performing the experimental task (e.g., naming or lexical decision). For example, Jared and Seidenberg (1991) varied the proportion of homophones in two of their experiments. In two experiments 80% of the trials consisted of homophones. Since the task was to decide if a target word was a member of a category, phonological recoding could only make performance on the task more difficult. Thus, it might be expected that these conditions would lead subjects to strategically avoid the use of phonological recoding if that is possible. Jared and Seidenberg found clear evidence that including a large proportion of homophones in the stimulus list did not change subjects' word recognition strategies. They showed exactly the same pattern of results as in earlier studies with a lower proportion of homophones (the false positive effects described earlier in this paper). This manipulation did have an effect on subjects' behavior, however: it resulted in much more cautious decision-making. Subjects computed word meanings in the same way, but took longer to make their decisions and made fewer errors. In summary, it is possible that readers may be able to flexibly allocate attention to one or the other recognition process, though my own guess is that it is much easier to change the strategies used in performing a task such as lexical decision than to alter the process used in computing the meanings of words. Once the division of labor is established, it would seem to take extraordinary conditions (such as brain injury; Patterson et al., 1985) to change it. This is an important issue that needs to be addressed in further research.

Conclusions

I have proposed a view in which the meaning of a word is built up on the basis of activation from multiple sources. It therefore does not have to be determined exclusively on the basis of orthographic or phonological information. This changes the terms of the debate over direct vs. phonologically-mediated access, insofar as there can be simultaneous activation from both sources. I have suggested that a key theoretical issue concerns the division of labor between component processes in a multi-route system. Developing a more detailed account of the factors that influence the division of labor, their interactions, and their relative salience, can be seen as an important direction for future research. Orthographic depth is probably one of these factors. However, the effects of orthographic depth on processing are not as simple as the original hypothesis suggested, because the depth factor is confounded with many other differences among orthographies and the languages they represent. The tradeoffs among these factors appear to yield very similar outcomes across orthographies with regard to the division of labor. In

English it appears that a relatively small pool of high frequency words can be recognized on a visual basis; the same may be true of other writing systems as well. Phonology does contribute in a central way to the acquisition of word recognition skills and to the recognition of many words even among skilled readers, even in putatively "deep" orthographies such as English and Hebrew. Thus, humans seem to converge on remarkably similar mechanisms in reading, despite apparent differences among orthographies.

Acknowledgment
I am grateful to Maryellen MacDonald for comments on earlier drafts of this article (stretching back over several years, actually). I thank Shlomo Bentin for patiently trying to explain some of the mysteries of reading Hebrew to me; he may not agree with everything I say about Hebrew but he has had a large impact on my thinking nonetheless. I also thank my colleagues Karalyn Patterson and James L. McClelland, who greatly affected my thinking about the division of labor issue.

References

Adams, M. (1990). *Beginning to read.* Cambridge, MA: MIT Press.

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

Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17,* 644-652.

Baron, J. (1973). Phonemic stage not necessary for reading. *Quarterly Journal of Experimental Psychology, 25,* 241-246.

Barron, R. W. (1978). Access to the meanings of printed words: Some implications for reading and learning to read. In F. B. Murray (Ed.), *The recognition of words: IRA series on the development of the reading process.* Newark, DE: International Reading Association.

Barron, R. W. (1981). Reading skill and reading strategies. In A.M. Lesgold & C.A. Perfetti (Eds.), *Interactive processes in reading.* Hillsdale, NJ: Erlbaum.

Besner, D., Twilley, L., McCann, R., & Seergobin, K. (1990). On the connection between connectionism and data: Are a few words necessary? *Psychological Review, 97,* 432-446.

Carr, T. H., & Pollatsek, A. (1985). Recognition of printed words: A look at current models. In D. Besner, T. G. Waller, & G. E. MacKinnon (Eds.), *Reading research, Vol. 5.* New York: Academic Press.

Chomsky, N., & Halle, M. (1968). *Sound pattern of English.* New York: Harper & Row.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* 332-361.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing.* New York: Academic Press.

Coltheart, M. (1987). Introduction. In M. Coltheart, G. Sartori, & R. Job (Eds.), *Cognitive neuropsychology of language.* London: Erlbaum.

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

Doctor, E.A., & Coltheart, M. (1980). Children's use of phonological encoding when reading for meaning. *Memory & Cognition, 8,* 195-209.

Feldman, L. B. (1981). Visual word recognition in Serbo-Croatian is necessarily phonological. *Haskins Laboratories Status Report on Speech Research, SR-66.*

Feldman, L. B., & Turvey, M. T. (1983). Word recognition in Serbo-Croatian is phonologically analytic. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 288-298.

Forster, K. I. (1981). Frequency blocking and lexical access: One mental lexicon or two? *Journal of Verbal Learning and Verbal Behavior, 20,* 190-203.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M.Coltheart (Eds.), *Surface dyslexia.* London: Erlbaum.

Frost, R. (1991). Phonetic recoding of phonologically ambiguous printed words. Paper presented at the 32nd Annual Meeting of the Psychonomic Society (November). San Francisco.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographic depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 104-115.

Funnell, E. (1983). Phonological processes in reading: New evidence from acquired dyslexia. *British Journal of Psychology, 74,* 159-180.

Gordon, B. (1983). Lexical access and lexical decision: Mechanisms of frequency sensitivity. *Journal of Verbal Learning and Verbal Behavior, 22,* 24-44.

Gupta, P., & Touretzky, D. (1991). Connectionist models and linguistic theory: Investigations of stress systems in language. Unpublished manuscript, Carnegie Mellon University, Pittsburgh.

Henderson, L. (1982). *Orthography and word recognition in reading.* London: Academic Press.

Hinton, G., & Shallice, T. (1990). Lesioning a connectionist network: Investigations of acquired dyslexia. *Psychological Review.*

Hung, D.L., & Tzeng, O.J.L. (1981). Orthographic variations and visual information processing. *Psychological Bulletin, 90,* 377-414.

Hynd, G. W., & Semrud-Clikeman, M. (1989). Dyslexia and brain morphology. *Psychological Bulletin, 106,* 447-482.

James, C.T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 104,* 130-136.

Jared, D., McRae, K., & eidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29,* 687-715.

Jared, D., & Seidenberg, M.S. (1990). Naming multisyllabic words. *Journal of Experimental Psychology: Human Perception and Performance, 16, 92-105.*

Jared, D., & Seidenberg, M. S. (1991). Does word recognition proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General.*

Katz, L., & Feldman, L. B. (1981). Linguistic coding in word recognition: Comparisons between a deep and a shallow orthography. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading.* Hillsdale, NJ: Erlbaum.

Katz, L., & Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 157-166.

Lukatela, G., & Turvey, M. (1980). Some experiments on the Roman and Cyrillic alphabets of Serbo-Croatian. In J. F. Kavanagh & R. L. Venezsky (Eds.), *Orthography, reading, and dyslexia.* Baltimore, MD: University Park Press.

Lukatela, G., Feldman, L.B., Turvey, M. T., Carello, C., & Katz, L. (1989a). Context effects in bi-alphabetical word perception. *Journal of Memory and Language, 28,* 214-236.

Lukatela, G., Popadić, D., Ognjenović, P., & Turvey, M. T. (1980). Lexical decision in a phonologically shallow orthography. *Memory & Cognition, 8,* 124-132.

Lukatela, G., Turvey, M.T., Feldman, L., Carello, C., & Katz, L. (1989b). Alphabetic priming in bi-alphabetical word perception. *Journal of Memory and Language, 28,* 237-254.

McCann, R., & Besner, D. (1987) Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 14-24.

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.

McCusker, L. X., Hillinger, M. L., & Bias, R. G. (1981). Phonological recoding and reading. *Psychological Bulletin, 89,* 217-245.

Meyer, D.E., Schvaneveldt, R.W., & Ruddy, M.G. (1974). Functions of graphemic and phonemic codes in visual word recognition. *Memory & Cognition, 2,* 309-321.

Monsell, S., Patterson, K. E., Graham, A., Hughes, C. H., & Milroy, R. (in press). Lexical and sublexical translation of spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

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.

Patterson, K. E., Marshall, J. C., & Coltheart, M. (Eds.). (1985). *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading.* London: Erlbaum.

Patterson, K. E., & Coltheart, V. (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: Reading.* London: Erlbaum.

Pinker, S. (1991). Rules of language. *Science.*

Plaut, D. C. (1991). Connectionist neuropsychology: The breakdown and recovery of behavior in lesioned attractor networks. Carnegie-Mellon University Department of Computer Science Technical Report CMU-CS-91-185.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 10,* 645-657.

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In D. Rumelhart & J. McClelland (Eds.), *Parallel distributed processing, Vol. 1.* MIT Press.

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition, 19,* 1-30.
Seidenberg, M. S. (1989). Word recognition and naming: A computational model and its implications. In W.D. Marlsen-Wilson (Ed.), *Lexical representation and process.* MIT Press.
Seidenberg, M. S. (1990). Lexical access: Another theoretical soupstone? In D. Balota, G. Florais d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading.* Hillsdale, NJ: Erlbaum.
Seidenberg, M. S. (1991). Connectionism without tears. In S. Davis (Ed.), *Connectionism: theory and practice.* Cambridge University Press.
Seidenberg, M. S. (1991). Dyslexia as a computational disorder. In P. Gough, R. Treiman, & L. Ehri (Eds.), *Reading acquisition.* Hillsdale, NJ: Erlbaum.
Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523-568.
Seidenberg, M. S., & McClelland, J. L. (1990). More words but still no lexicon: Reply to Besner et al. *Psychological Review, 97,* 447-452.
Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre- and post-lexical loci of contextual effects on word recognition. *Memory and Cognition, 12,* 315-328.
Seidenberg, M. S., & Waters, G. S. (1989). A mega study of word naming. Paper presented at the 30th annual meeting of the Psychonomic Society (November). Chicago, IL.
Sejnowski, T. J., & Rosenberg, C. R. (1986). NETtalk: A parallel network that learns to read aloud. Baltimore, MD: Johns Hopkins University EE and CS Technical Report JHU/EECS-86/01.
Shulman, H. G., Hornak, R., & Sanders, E. (1978). The effects of graphemic, phonetic, and semantic relationships on access to lexical structures. *Memory & Cognition, 6,* 115-123.
Stanovich, K. E., & West, R. F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 658-672.
Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language, 27,* 324-340.
Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word recognition. *Journal of Memory and Language, 26,* 608-631.
Turvey, M. T., Feldman, L. B., & Lukatela, G. (1984). The Serbo-Croatian orthography constrains the reader to a phonologically analytic strategy. In L. Henderson (Ed.), *Orthographies and reading.* London: Erlbaum.
Turvey, M. T., Feldman, L. B., & Lukatela, G. (1984). The Serbo-Croatian orthography constrains the reader to a phonologically analytic strategy. In L. Henderson (Ed.), *Orthographies and reading.* London: Erlbaum.
Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition, 15,* 181-198.
Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 371-386.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of a subsymbolic psycholinguistics. *Psychological Review, 97,* 488-522.

Wang, W. S.-Y. (1979). Language change: A lexical perspective. *Annual Review of Anthropology, 8,* 353-371.

Waters, G., Caplan, D., & Hildebrandt, N. (1987). Working memory and written sentence comprehension. In M. Coltheart (Ed.), *Attention and performance XII: Reading.* Hillsdale, NJ: Erlbaum.

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: time course and decision criteria. *Memory & Cognition, 13,* 557-572.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 6

Automatic Activation of Linguistic Information in Chinese Character Recognition

Daisy L. Hung
Ovid J. L. Tzeng
Angela K. Y. Tzeng
University of California, Riverside
The Salk Institute for Biological Studies, San Diego
National Chung Cheng University, Taiwan

The invention of writing systems is undoubtedly one of the most important cultural achievements of humankind. Because of the ability to transcribe spoken language into some kind of written representation, communication has become vastly expanded to overcome the limitations of space and time that are usually imposed on the spoken sound. But development of useful and efficient systems came very slowly and the evolution of writing took many different twists until it stumbled onto a single path: Any fully developed writing system has to be speech-based (DeFrancis, 1989), even though the way speech is represented in the script varies from one language to another. Indeed, it took a span of many thousands of years for our ancestors to come up with systems that work for different languages. But the relationship between script and speech is not a simple one: At every advance, the number of graphic symbols in the script decreases, and as a direct consequence, the abstractness of the relation between script and meaning increases as the link between grapheme and speech becomes clearer. Consider, however, the Chinese writing system which was constructed essentially on the basis of a syllabic principle but also with the addition of a great deal of morphological information to increase the graphic distinctiveness. As a consequence, the same syllable may be represented by different logographs with different meanings.

Due to the morphosyllabic nature of Chinese, its logographs number in the thousands and are complex spatial configurations. Such unique script/speech relationships have led some investigators to speculate that reading a Chinese text may require different visual information processing strategies from those involved in reading texts written in an alphabetic script. For example, it has been assumed that the only way to read Chinese is by direct unmediated visual access (Barron, 1978). However, such an orthographic-specific reading view is not shared by many cognitive psychologists who regard the orthographic differences to be only skin deep (Seidenberg, 1985; Tzeng & Hung, 1988). The controversy focuses on how much (if any) phonological information can be recovered, given that the logographs contain such an opaque script/speech relationship. Of course, by now, no one would hold the simple-minded view of treating Chinese reading as picture recognition. A more realistic view is the one that regards reading as a set of basic processes attached to a set of primary language processes which include phonological processes. Under such a conceptualization, the arousal of phonological processes is not optional or strategic but automatic and obligatory in any orthography.

In the past, various memory paradigms have been used to investigate the process of speech recoding in reading Chinese. Usually, experimenters presented lists of printed characters first; then, after a period of delay, subjects were instructed to recall all of the characters they previously read. Researchers hypothesized that the subjects' performance would be poorer on the trials containing stimulus characters which shared some degree of phonological similarity if the visual images had been recoded into a phonological format. Thus, a performance decrement due to the intralist phonological similarity was taken as evidence for the speech recoding process in reading Chinese (Tzeng, Hung & Wang, 1977). Another method which has also widely been used is the use of concurrent tasks in which subjects are asked to read sentences while they are engaging in a concurrent activity of articulating speech sounds. Again, performance decrement in conditions where the pronunciations of the characters in the sentence share common articulatory features with the concurrent speech sounds are taken as evidence for the occurrence of speech recoding (Baddeley, 1979). Results from studies of the last ten years have shown that speech recoding processes play an important role in memory and comprehension (Tzeng & Hung, 1988). In other words, regardless of the writing systems, speech codes are kept activated in working memory in support of comprehension (Perfetti & McCutchen, 1982).

Although the demonstration of speech recoding while reading texts written in such an opaque script is indeed intriguing, it can be argued that speech codes are activated strategically only in response to memory demands of a text, rather than routinely as part of language processing, per se. In recent years, however, studies began to examine the speech process at the word level using experimental procedures which minimize the requirement of working memory (e.g., lexical priming, character naming, etc.). The results are in general agreement with the idea of automatic speech recoding at the level of the character itself. Moreover, they also indicate that graphic units, especially those that serve as the phonetic stems, play a much more important role in the activation of the syllabic image (Fang, Horng, & Tzeng, 1986; Lien, 1985). By now, the question of interest is not whether or not there is an automatic and obligatory arousal of speech processes; rather, the question is how the various linguistic cues (e.g., graphic shapes, phonetic components, and semantic radicals) are incorporated to activate phonological codes during reading. Conventional experimental memory paradigms are hardly useful for finding the answers to this question. We intend to investigate this problem with a different experimental procedure.

In a typical Stroop paradigm, the "semantic color" and the "physical color" are the two stimulus dimensions being manipulated. That is, different colors of ink are used to print the names of different colors. When these two kinds of information are congruent with each other (e.g., the printed word "RED" in red, "BLUE" in blue), subjects have no problem either naming aloud the words they read, or calling out the name of the color used to print the words. However, interference from one dimension to another occurs when these two kinds of information conflict with each other; usually, word naming is much easier than color naming (e.g., "BLACK" in red, or "YELLOW" in blue). This is the so-called Stroop interference effect. It occurs because two different kinds of information are competing for articulation at the same time. For example, in the conflicting situation in which subjects need to answer "blue" when the target stimulus is a

"YELLOW" word in blue ink, the correct response of /blue/ is interfered by the automatically activated phonological code of /yellow/. This results in longer reaction times and a higher error rate for color naming.

The Stroop effect is robust and easy to obtain. Variations in actual experimental procedures have been created to study different aspects of the reading process with respect to the issue of automaticity in decoding (Tzeng & Wang, 1983). In our experiment, we employed a modified picture-word interference Stroop paradigm to investigate what kinds of information will be automatically activated and how they are combined in the process of character recognition. Golinkoff and her colleagues (Golinkoff & Rosinski, 1976; Rosinski, Golinkoff, & Kukish, 1975) had used such a paradigm to test the relationship between children's development of speech recoding and their reading ability in the process of learning English. In their experiments, twenty pictures of common objects were arranged in a 4 × 5 matrix on an 8 1/2 × 11-inch paper. In the center of every cell, an English word was printed together with the picture (see Figure 1). Subjects were instructed to name the picture as fast and as accurately as possible.

Based upon the relationship between the printed pictures and their embedded words, three experimental conditions were created in Golinkoff and Rosinski's (1976) study. In the 100% congruent condition, the printed word and the word representing the name of the picture were congruent with each other. For example, the word CUP was printed on the picture of a cup. In the 0% congruent condition, the same 20 pictures and 20 words as in the first condition were used again; however, the word and the name of the picture were not congruent for all 20 cells. For example, the word HAT was printed on the picture of a cup. Finally, in the CVC condition, what were printed in the pictures were consonant-vowel-consonant letter strings that were not real words. For example, YAT was printed in the center of the picture of a cup. Using the reaction time for naming all 20 pictures as the dependent variable, they found that picture naming in the 100% congruent condition was the fastest, while in the 0% congruent condition it was slowest, with the time in the CVC condition falling in between.

The investigators draw two major conclusions from these results. First, the facilitation in the congruent condition provides a strong evidence for the automatic activation of a phonological code in the reading of the distracting word. When the phonological representation of the printed word and the name of pictures are congruent with each other, the automatically activated linguistic information of the word helps subjects to generate a phonological code for the name of the picture. On the other hand, when the two pieces of information conflict with each other, the phonological code of the dominant word suppresses the generation of the name code for the object depicted in the picture. Second, based upon the comparison of the picture-naming times in the 0% congruent condition and the CVC condition, the investigators conclude that the Stroop interference is neither due to an attentional factor nor to a simple response competition disruption. Instead, they suggest that there must be a semantic component in the interference because the CVCs used in the experiment are pronounceable but nonsense. These two conclusions point out parallel activations of multiple linguistic cues (e.g., graphic, phonological, and semantic information) in word recognition. They also highlight the importance of a weighted integration at the response stage in a Stroop task.

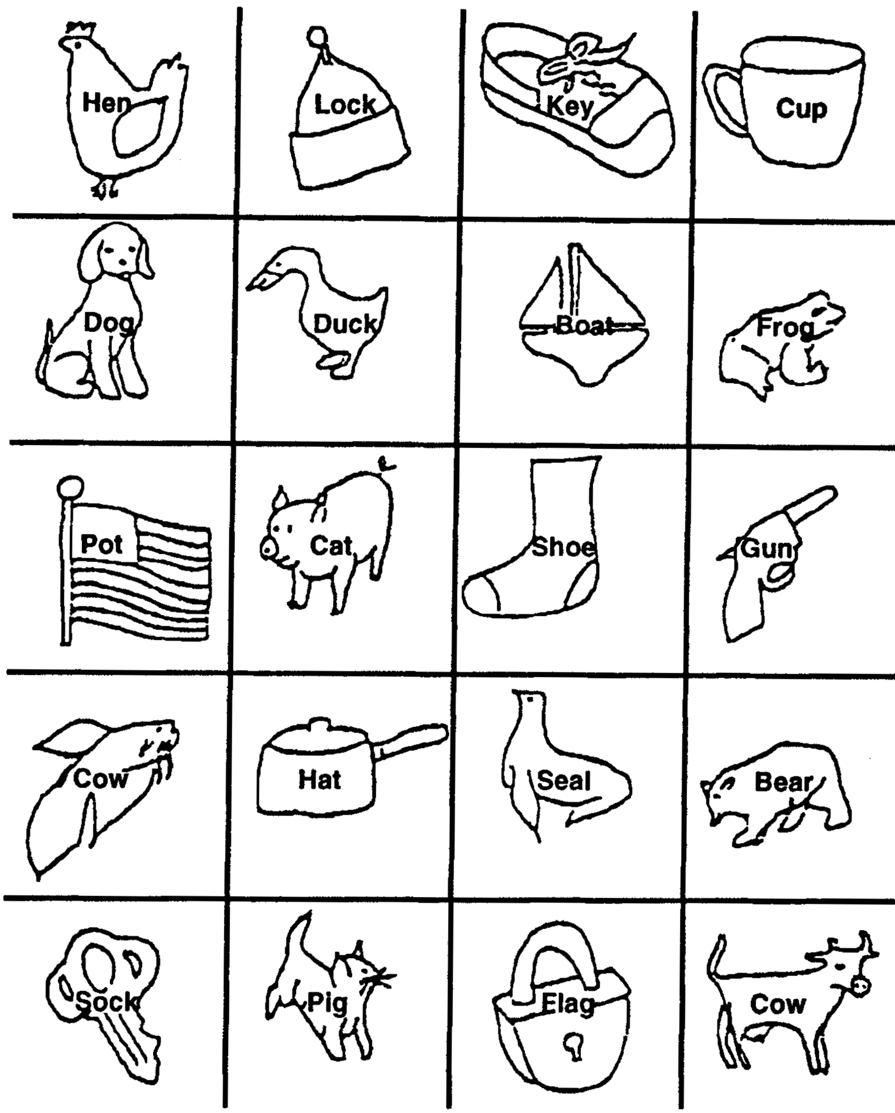

Figure 1. **Picture-Word Interference Paradigm, Golinkoff et al., 1975.**

These studies by Golinkoff and her associates establish the usefulness of the picture-word Stroop paradigm for the study of word recognition. In our study, we adopted such a modified Stroop paradigm to examine the process of character recognition in Chinese readers. Subjects were asked to call out the names of the pictures as fast and as accurately as possible, and to try not to be influenced by the characters printed with the pictures. By appropriate arrangements of the relationship between the word represented by the picture and the distracting word printed in the center of the picture, and by comparing the picture naming times across the various experimental conditions, one may be able to specify what kinds of linguistic information are available and how they are combined. Since in the Chinese writing system characters with entirely different configurations may have identical pronunciations and characters with similar graphemic components may have very different pronunciations, it is much easier for the investigators to construct experimental stimuli which vary orthogonally along the dimensions of graphic configuration and pronunciations. Thus, seven experimental conditions were created to investigate the role of graphic, phonological, and semantic information in character recognition with a modified picture-character Stroop interference paradigm.

Method

Subjects. Thirty fluent Chinese readers (15 males, 15 females) participated in this experiment. All of them were graduate students enrolled in various departments at the University of California, Riverside. They were from Taiwan and should be considered as skilled readers of Chinese. They were paid for their participation in the experiment.

Materials and Design. The experiment was divided into two sessions. In the first session, subjects received seven warm-up sheets. Each of the warm-up sheets was 8 1/2 × 11-inch paper divided into 20 cells of equal size. Six of these warm-up sheets contained all of the characters (including real characters and pseudo-characters to be explained below) and the other contained 20 pictures (10 animals and 10 objects like those in Figure 1) to be used in the actual experimental session. The purpose of the warm-up session was to familiarize the subjects with the characters or pseudo-characters to be used in the experiment.

In the experimental session, the paired experimental stimuli (i.e., character and picture) were printed in each of the 20 cells on a piece of paper. The same 20 pictures were used in all experimental conditions. However, the characters embedded in the pictures were changed according to different experimental conditions. Depending on the relationship between the object depicted in each picture and the printed character, seven experimental conditions were created. Examples of the stimulus pair in each of the seven conditions are presented in Figure 2.

(1) In the CC (Completely Congruent) condition, the printed character was exactly the name of the pictured object.

(2) In the CI (Completely Incongruent) condition, the same 20 pictures and 20 characters as in the CC condition were used; however, the pairings of the characters and pictures were rearranged such that none of them were congruent with each other.

(3) In the SGSS (Similar Graph/Same Sound) condition, the character shared a similar graph and a same sound with the character which was the name of the pictured object.

(4) In the SGDS (Similar Graph/Different Sound) condition, the character shared a similar graph but had a different sound from the character representing the pictured object.

(5) In the DGSS (Different Graph/Same Sound) condition, the character had a different graph but shared the same sound with the character representing the object.

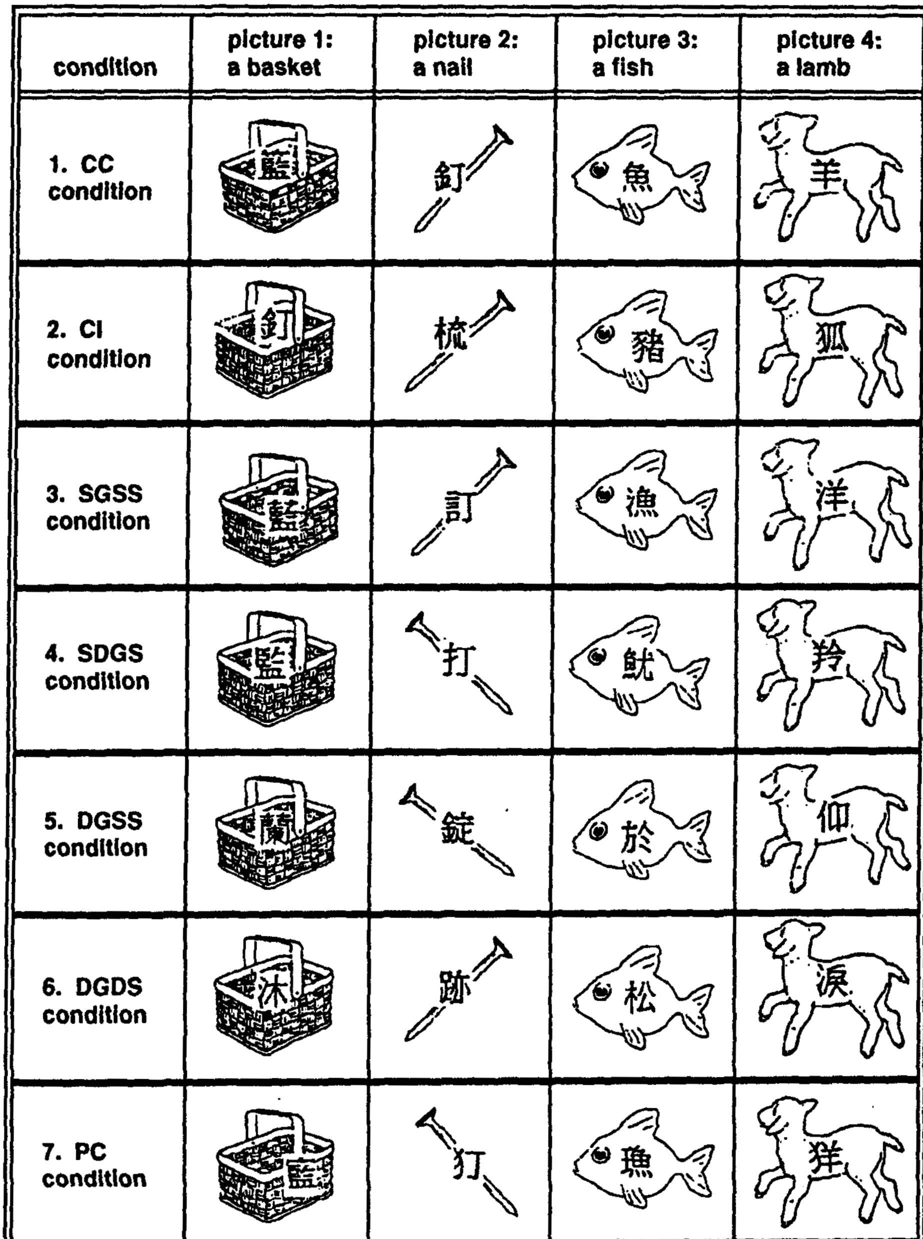

Figure 2. **Examples from seven experimental conditions in the Picture-Word Stroop Paradigm.**

(6) In the DGDS (Different Graph/Different Same) condition, the character shared neither the graph nor the sound with the character naming the object.

(7) In the PC (Pseudo-Character) condition, the printed character-like symbol was a pseudo-character in the sense that it was not a real character but was constructed according to orthographic principle and thus might be pronounceable because it contained a phonogram.

A complete within-subject design was employed. Every subject was required to participate in all seven experimental conditions. The orders of the seven conditions were counterbalanced across subjects by a Latin Square procedure.

Procedure. Subjects were run individually. In the warm-up stage, subjects were asked to call out the names of all characters, pseudo-characters, or pictures printed on a sheet of paper as fast and as accurately as possible. The purpose of this was to familiarize subjects with the experimental setup as well as the stimulus materials. At the end of the warm-up session, subjects were briefed on the tasks to be performed in the experimental stage. They were told that on each sheet of papers, 20 picture-character pairs were arranged in a 4 by 5 matrix. They were asked to call out the names of all the 20 pictures as fast and as accurately as possible and to ignore the printed characters.

The time for naming all 20 pictures was recorded as the dependent variable. All errors were recorded as well.

Results

Mean reaction times (in seconds) for naming all 20 pictures under the 7 different distracting conditions are shown in Table 1. As expected, subjects took the least time (10.15 s) to name all twenty pictures under the CC condition in which the names of the pictures and the embedded Chinese characters were completely congruent with one another.In contrast, they took the longest time(17.17 s) and made the largest number of errors (21 % error rate) under the CI condition in which the names of the pictures and the embedded characters were completely incongruent with one another. Since the pseudo-characters used in the PC condition have no meaning, the mean total time (11.73 s) required for naming all twenty pictures in this particular condition can be taken as a baseline time for general interference. Against this time, the CC condition shows an facilitation effect whereas the CI condition shows an inhibition effect; both effects are statistically significant (for both, $p<.01$).

Like Golinkoff and her colleagues, we also found the robustness of the modified picture-word interference effects from analyses of data of the above three basic conditions. Table 1 also shows results from other experimental conditions; the interference effect is apparent in each and every condition. Again, compared to the baseline time in the PC condition, mean naming times in conditions 3, 4, 5 and 6 (i.e., SGSS, SGDS, DGSS, and DGDS, respectively) are all significantly slower (all $p<.01$). In other words, the manipulations of the graphic similarity and the phonological similarity between the character denoting the pictured object and character embedded in the picture produce interference effect. Since we are more interested in the magnitude of the interference produced by each of the two variables and in their interaction, if any, and since these four conditions represent the orthogonal arrangement of these two variables, an ANOVA for a two-by-two factorial design (within-subject) was performed on the naming time data. Results of the statistical analyses are presented in the following section.

Table 1

The Average Reaction Tine of Picture-Word Stroop Interference Experiment in Chinese.

Conditions	Chinese Character †	Pronunciation	Average RT (s)	Error Rates %
1. Completely Congruent (CC)	籃 (a basket)	/lan/	10.15	0
2. Completely Incongruent (CI)	釘 (a nail)	/ding/	17.17	21
3. Similar Graph Same Sound (SGSS)	藍 (blue)	/lan/	12.30	1
4. Similar Graph Different Sound (SGDS)	監 (jail)	/jian/	13.23	2
5. Different Graph Same Sound (DGSS)	蘭 (orchid)	/lan/	14.00	4
6. Different Graph Different Sound (DGDS)	沐 (bath)	/muh/	15.45	7
7. Pseudo-Character (PC)	 (n/a)	/jian/ (?) /lan/ (?) *	11.73	2

†: Chinese characters embedded on the picture of "a basket" in different conditions. The words in the parentheses are the meaning of these Chinese characters.
*: These are two alternatives of possible pronunciation according to the phonogram contained in this pseudo-character.

As expected, there is a significant main effect of the graphic similarity, with a pronounced reduction of interference from the different-graph conditions (i.e., 14.34 s by averaging across the DGSS and DGDS conditions) to the similar-graph conditions (i.e., 13.15 s by averaging across the SGSS and SGDS conditions), $F(1,29)=21.56$, $p<.01$. Similarly, there is a significant effect of the phonological similarity, with a reduction of interference from the different sound conditions (i.e., 14.34 s by averaging across the SGDS and DGDS conditions) to the similar-sound conditions (i.e., 13.15 s by averaging across the SGSS and DGSS conditions), $F(1, 29)=6.97$, $p<.05$. More importantly, there is no interaction between these two factors, $F(1,29)=.60$, suggesting that these two effects are independent of each other.

Discussion

Word recognition is central to the theoretical understanding of reading behaviors. In recent years, reading researchers have provided evidence for a speech recoding process in the early stage of reading printed words. The importance of transforming the visible symbols into a speech format can easily be appreciated by readers of an alphabetic script in which every letter or letter group has a corresponding phonemic representation, even if the mapping relation may be rather crude. Based upon this observation, investigators of the reading process can build theories about lexical access by proposing different "routes" (or pathways) from print to word recognition. For beginning readers, the essential task is to build enough rule-based knowledge for the activation of these different routes, whereas for becoming a skilled reader, the major task is to automatize the rule-activation process given the orthographic information. Such an analytical approach makes a great deal of sense for an alphabetical writing because the orthography does provide important clues for the recovery of the phonological information. Indeed, experimental as well as neuropsychological (e.g., acquired dyslexia) evidence for the psychological reality of such a phonological route is abundant, albeit there is strong disagreement on how to activate the phonological representation (i.e., rule-based vs. assembled phonology).

In contrast, for some investigators the analytical approach to word recognition is not readily applicable to the character recognition process in Chinese. The reason for the objection is simply that decomposition of the character would not allow the reader to generate the phonological information necessary for the access of lexical information. They claim that as a logographic writing system, the Chinese character can only be read in its entirety. Since no systematic mapping between script and sound can be found, it would be a futile practice to attempt a phonological route (Zhou, 1978).

The above conceptualization about the characters is, of course, not correct. In a seminal work on the classification of the writing systems, DeFrancis (1989) makes a detailed analysis of the Chinese writing system from the perspective of its historical developments. He convincingly shows that there is much phonological information in the characters; 85% or more characters are phonograms in which a part of the character carries clues about its pronunciation. DeFrancis' observation is backed up by recent experimental demonstrations of a consistency effect in which the degree of consistency of a phonetic component as a pronunciation clue is found to be negatively correlated with its naming latency (Fang et al., 1984; Lien, 1985; Seidenberg, 1985; Tzeng, Hung, & Lee, 1991). The effect is analogous to that found in English in which the regularity of spelling or pronunciation is found to influence word recognition (Seidenberg & McClelland, 1989). Moreover, in a study with a group of bilingual children who were learning to read both English and Chinese at the same time, the ability to analyze the internal structure of the speech sound, as assessed by a so-called phonemic awareness task, was found to have a significant correlation not only with learning to read English, but with learning to read Chinese as well (Lee, Chang, Tzeng, & Hung, 1991).

In fact, knowledge of the involvement of a speech recoding process during the reading of Chinese text has been around since its demonstration by Tzeng et al. (1977). However, the observed phonological effect has been interpreted by others as memory based. In other words, the phonological representation is generated postlexically and associatively in order to meet the memory demand in sentence comprehension. But, again, such a memory based phonological view has been disputed by the demonstration of a consistent phonological priming effect in LDT (lexical decision task) paradigms which involve SOAs (stimulus onset asynchrony) as short as 50 ms (Perfetti & Zhang, 1991). Interestingly, the effect of tone-sandhi rules can readily be observed in the speech recoding process during the reading of multi-syllablic words written in Chinese characters

(Xu, 1991). The conclusion seems clear: The phonological effect in the reading of the Chinese characters is real and its nature seems to be similar to that generated in an alphabetic script.

Results from the present experiment strengthen the above conclusion by showing the automatic aspect of the speech recoding process in a modified picture-word Stroop experiment. The experimental task has a minimal demand on memory. Nevertheless, a significant effect of the phonological factor has been established beyond doubt. In this respect, it is important to note that the phonological information generated by the distracting character in fact speeds up, rather than slows down picture naming times in the similar-sound conditions (i.e., SGSS and DGSS) as compared to the naming time required for the condition which contains no such information (i.e., the CI condition). In other words, we observe a reduction of interference due to the availability of the phonological information in the distracting stimulus. In a sense, we obtain a phonological facilitating effect and the magnitude of such a positive priming effect is as great as that of the graphic priming effect observed under the similar-graph conditions (i.e., SGSS and the SGDS conditions). But how do we explain the latter effect when there is no graphic information on the pictured object itself?

In fact, under the SGDS condition, subjects see a distracting character which actually generates a phonological code different from the name code of the target picture. One would expect an interference. Instead, compared with the CI condition (0% congruence condition), naming time is much shorter in the SGDS condition, again showing a reduction in interference. In order to explain this puzzling facilitating effect of the graphic similarity, one has to assume that a mental image of the character representing the name of the object is activated by the pictured object itself and the graphic information provided by the distracting character has a priming effect on the activation of this particular mental character. The fact that the factors of graphic similarity and phonological similarity do not interact supports the idea of two separate and independent routes for character recognition in reading Chinese.

Summary

This experiment examined fluent Chinese readers' character recognition performance with a modified Stroop interference paradigm (i.e., picture-character interference) in order to see what cues were effective and how they were integrated during reading. Depending on the relationship between the object depicted in each picture and the printed character, seven experimental conditions were created. Several interesting results were observed. First, the subjects' picture-naming performance was disrupted more when the printed character and the character representing the name of the picture shared neither graphic nor phonological information. Second, little disruption was observed when the distracting symbol was a pseudo-character, suggesting that its non-lexical status prevented the activation of other linguistic information (e.g., phonological and graphic clues) embedded in the stimulus. Third, when the disrupting symbol was a real character, the magnitude of interference was the same for lack of graphic similarity as for lack of phonological similarity, supporting the notion that both graphic and phonological information were automatically activated in Chinese character recognition. Moreover, the integration of these two types of information was found to be additive in nature.

Acknowledgment
The research reported here was supported in part by a search grant from the CCk Cultural Foundation to Daisy Hung and in part by a research grant from the National Science Council of the Republic of China (NSC80-0301-H194-05) to Ovid Tzeng.

References

Baddeley, A. D. (1979). Working memory and reading. In P. A. Kolers, M. E. Wrostad, & H. Bouma (Eds.), *Processing of visible language*. Vol. 1 (pp. 355-370). New York: Plenum Press.

Barran, R. W. (1978). Access to the meaning s of printed words: Some implications for reading and learning to read. In F. B. Murray (Ed.), *The recognition of words: IRA series on the development of the reading process* (pp. 34-56). Newark, DE: International Reading Association.

DeFrancis, J. (1989). *Visible language: The diverse oneness of writing systems*. Honolulu: University of Hawaii Press.

Fang, S. P., Horng, R. Y., & Tzeng, O. J. L. (1986). Consistency effects in the Chinese character and pseudo-character naming tasks. In H. Kao & R. Hoosain (Eds.), *Linguistics, psychology, and the Chinese language*. Hong Kong: University of Hong Kong Press.

Golinkoff, R. M., & Rosinski, R. R. (1976). Decoding, semantic processing, and reading comprehension skill. *Child Development, 47*, 252-258.

Lee, W. L., Chang, J. M., Tzeng, O. J. L., Wee, G. C., & Hung, D. L. (1991, November). Phonemic awareness in reading logographic and alphabetic scripts in children bilingual in Chinese and English. Paper presented at the annual meeting of the Orton Society of Dyslexia, Portland, Oregon.

Lien, Y. W. (1985). *Consistence of the phonetic clues in the Chinese phonograms and their naming latencies*. A Master's thesis submitted to the National Taiwan University.

Perfetti, C. A., & McCutchen, D. (1982). Speech processes in reading. In N. Lass (Ed.), *Speech and language: Advances in basic research and practice* (pp. 237-269). New York: Academic Press.

Perfetti, C. A., & Zhang, S. (1991). Phonological processing reading Chinese words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 633-643.

Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Development, 46*, 247-253.

Seidenberg, M. (1985). The time course of phonological code activation in two writing systems. *Cognition, 19*, 1-30.

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

Tzeng, O. J. L., & Hung, D. L. (1988). Orthography, reading, and cerebral functions. In D. de Kerckhove & C. J. Lumsden (Eds.), *The alphabet and the brain: The lateralization of writing* (pp. 273-290). Berlin: SpringerVerlag.

Tzeng, O. J. L., Hung, D. L., & Lee, W. L. (1991, March). A conspiracy effect in the recognition of Chinese characters. Paper presented at the Conference on From Spoken to Written Languages, Bellagio, Italy.

Tzeng, O. J. L., Hung, D. L., & Wang W. S.-Y. (1977). Speech recoding in reading Chinese characters. *Journal of Experimental Psychology, 6,* 621-630.

Tzeng, O. J. L., & Wang W. S.-Y. (1983). The first two R's. *American Scientist, 71,* 238-243.

Xu, Y. (1991). Depth of phonological recoding in short-term memory. *Memory & Cognition, 19,* 263-273.

Zhou, Y. G. (1978). To what degree are the "phonetics" of present-day Chinese characters still phonetic? *Zhongguo Yuwen, 146,* 172-177.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 7

Orthographic Neighborhoods and Visual Word Recognition

Jonathan Grainger
Laboratory of Experimental Psychology C.N.R.S.
René Descartes University, Paris

Languages with alphabetic orthographies have the great advantage of providing the experimental psychologist with a simple metric for measuring formal similarities between words. In this chapter I will review some recent experimental work on visual word recognition in such languages (Dutch, English and French), where researchers have attempted to evaluate the role of similarity neighborhoods in the word recognition process.

Written words are similar to other words in many different ways: visually (e.g., try-fog); orthographically (e.g., foe-fog); phonologically (e.g., fought-fog); and semantically (e.g., mist-fog). As can be seen from the above examples, these different similarity relationships are often confounded and difficult to disentangle. Some writing systems do allow specific solutions to unconfounding these similarity relationships, one of the most studied examples being Serbo-Croatian (see, for example, Lukatela & Turvey, 1990). Nevertheless, the research to be presented here is centered on orthographic neighborhoods between words in languages where this particular similarity relationship is often partially confounded with both visual and phonological similarity.

At the theoretical level, the research to be presented here supports a view of visual word recognition as a process in which incoming sensory information defines a set of lexical candidates that then compete with each other for identification. The recognition process can, according to this point of view, be (artificially) divided into two stages: 1) sensory information is *mapped* onto the stored representations of words in memory; and 2) one lexical representation is *selected* as the best candidate for identification. The principal goal of any model of visual word recognition should be to provide a formal description of the processes underlying these two basic operations.

The present chapter will attempt to show how recent research on neighborhood effects in visual word recognition provides strong support for the activation metaphor (McClelland & Rumelhart, 1981; Morton, 1969; Seidenberg & McClelland, 1989) as opposed to the "access code/guided search" metaphor (Forster, 1976) as an appropriate means of describing the mapping process by which sensory information makes contact with lexical representations. Moreover, the "frequency ordered search/verification" metaphor (Becker, 1976; Forster, 1976; Paap, Newsome, McDonald, & Schvaneveldt, 1982) is shown to be an inadequate description of later selection processes, at least in its present formulation within the above cited models. The type of model that appears capable of accomodating the data in this field is characterized by 1) a hierarchical structure of sublexical and lexical representations; 2) a feedforward facilitatory flow of activation from the sublexical to the lexical level; and 3) inhibitory connections between representations at the lexical level. These three characteristics are an integral part of the interactive activation model of visual word recognition (McClelland & Rumelhart, 1981).

The results that support this type of cascaded, hierarchical, activation model of visual word recognition have been generated from research manipulating both the orthographic similarities between words, and the relative frequencies of these different words. These relationships have been manipulated in isolated word recognition paradigms, and in a masked priming paradigm where target words are preceded by briefly presented prime stimuli.

Word frequency and neighborhood frequency effects

Word frequency is operationally defined as the number of times a given word occurs in a given corpus of written text (typically expressed as number of occurrences per million). This measure is expected to reflect the frequency with which the average reader will encounter a particular word in his or her reading experience. The validity of this measure will depend on the quality of the corpus in terms of size and representativity. With the development of computer facilities in scientific research, languages such as English, French and Dutch now have excellent lexical data bases providing a high quality tool for reading research. Using the frequency counts generated from such sources, experimental psychologists have systematically observed that more frequently occurring words are easier to recognize (in a variety of tasks) than less frequently occurring words (see Monsell, Doyle, & Haggard, 1989, for a recent review).

In a number of recent articles (Grainger, 1990; Grainger, O'Regan, Jacobs, & Segui, 1989; Grainger & Segui, 1990) it has been suggested that one major predictor of ease of word recognition, other than the frequency of occurrence of the stimulus word itself, is the frequency of words that are orthographically similar to the stimulus (neighborhood frequency). Words that are orthographically similar to a more frequently occurring word (e.g., BLUR similar to BLUE) are harder to recognize than words that have no such higher frequency neighbors. A more extreme version of this hypothesis had previously been proposed by Havens and Foote (1963), who concluded on the basis of their perceptual identification results that visual duration thresholds for isolated printed words "...are not primarily a function of the frequency of prior usage of stimuli but of the ability or inability of the stimuli to evoke high frequency competitive responses." In other words, according to Havens and Foote (1963), and contrary to the conclusions drawn by other researchers at that time (e.g., Howes & Solomon, 1951), the ease with which we identify written words is determined essentially by neighborhood frequency and not stimulus word frequency.

Exactly the same hypothesis is embodied in two of the most cited versions of serial search models of visual word recognition (Forster, 1976; Paap et al., 1982). In these models there is a frequency ordered search among a set of lexical candidates determined by initial bottom-up processing. The search process works through this candidate set until a satisfactory fit is obtained between stored information about the word and information generated on-line from the stimulus. According to these models, word recognition time will depend essentially on the position the word occupies in the search set, high frequency words will be at the beginning and low frequency words at the end. This therefore allows such models to accommodate the pervasive effects of stimulus word frequency reported in the literature. What these models really predict, however, is not that stimulus word frequency per se affects word recognition performance, but rather the frequency of the stimulus word relative to the frequencies of the other words that figure in the candidate set. Words that are similar to more frequent words should take longer to recognize than words that are not similar to any words that are more frequent than themselves. On the other hand, when the frequencies of these orthographically similar words are controlled, then stimulus word frequency alone should not affect word recognition performance.

Recent experimental results suggest that this hypothesis can be rejected. In a number of word recognition experiments manipulating both stimulus word frequency and neighborhood frequency (existence or not of a word that is both orthographically similar to, and more frequent than the stimulus) it has been shown that effects of stimulus word frequency are obtained with words that are not similar to a more frequent word (Grainger, 1990; Grainger & Segui, 1990). The results of Grainger and Segui (1990) are summarized in Figure 1.

In Grainger and Segui's (1990) study, word recognition performance was measured using the lexical decision task and the progressive demasking task (derived from Feustel, Shiffrin, & Salasoo, 1983). In the latter paradigm, the stimulus word and a pattern mask are presented in successive cycles with the duration of the mask decreasing and the duration of the stimulus word increasing. Subjects simply have to press a response button when they have recognized a word. The screen is then cleared and subjects are asked to type in the word using the computer keyboard. This technique provides a very sensitive measure of word recognition latencies and both quantitative and qualitative error data. The results show that stimulus word frequency affects word recognition in both categories of words, with or without a higher frequency neighbor, but the effects are larger in the latter category. This interaction between stimulus word frequency and neighborhood frequency, although not observed in the lexical decision latencies, did appear in the lexical decision error data. The interaction reflects the greater neighborhood interference observed with low frequency words compared to medium frequency words.

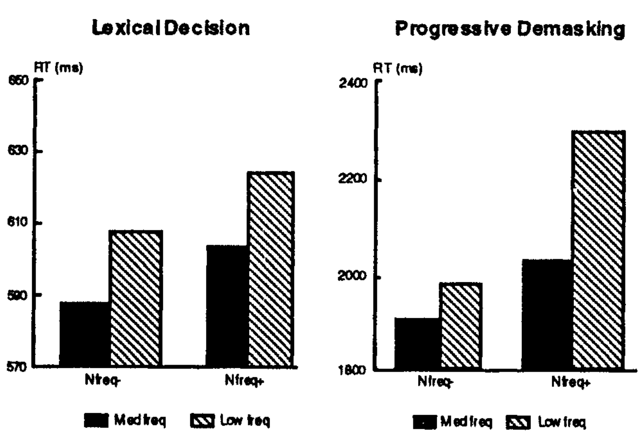

Figure 1. **Results of Grainger and Segui (1990) showing the effects of stimulus word frequency (medium frequency versus low frequency) and neighborhood frequency (Nfreq- : words with no higher frequency neighbors; Nfreq+ : words with higher frequency neighbors) observed in the lexical decision and progressive demasking paradigms.**

The vast amount of research in the past on word frequency effects was of course carried out without controlling for effects of neighborhood frequency. This is problematical in that low frequency words tend to have higher frequency neighbors whereas high frequency words do not. This therefore suggests that previous reports of word frequency effects in visual word recognition were generally combining both a "pure" stimulus word frequency component with the effects of neighborhood frequency. This would therefore lead to an overestimation of the amplitude of stimulus word frequency effects in tasks such as lexical decision where neighborhood frequency has a negative effect on performance. In tasks such as word naming, however, more frequent orthographic neighbors can inhibit or facilitate performance depending on whether or not these neighbors are pronounced similarly to the stimulus word (Jared, McRae, & Seidenberg, 1990). In other words, neighborhood frequency will not systematically exaggerate stimulus word frequency effects in word naming and may even diminish such effects when the neighbors are all pronounced similarly to the target (e.g., REEL: FEEL, HEEL, REAL, REED, REEF). This would therefore explain why the reported effects of stimulus word frequency are much smaller and less robust in word naming than in lexical decision.

In support of this analysis, it was observed (Grainger, 1990) that the size of stimulus word frequency effects in lexical decision and naming can be rendered comparable or even reversed (i.e., larger effects in the latter task) when neighborhood frequency is controlled. Figure 2 provides the main results of these experiments. The lexical decision results corroborate those presented in Figure 1. Stimulus word frequency and neighborhood frequency have independent effects on lexical decision RTs. The naming data, on the other hand, show an interaction between stimulus word frequency and neighborhood frequency. Larger effects of stimulus word frequency were obtained to words with no higher frequency neighbors. This interaction is exactly the opposite to that observed in the progressive demasking paradigm and in lexical decision error rates.

These opposite interaction effects between stimulus word frequency and neighborhood frequency observed with the progressive demasking paradigm (Figure 1) and the naming task (Figure 2) reflect the fact that more frequent orthographic neighbors, provide inhibitory effects in word recognition and facilitatory effects in word naming. When the stimulus word has no higher frequency neighbors then stimulus word frequency has similar effects in both tasks. Measured in terms of percent average response time, the amplitude of the word frequency effect is 3.9% in the progressive demasking task and 4.5% in word naming for this category of words. On the other hand, when the stimulus word is orthographically similar to more frequent words, the effects of stimulus word frequency increase in progressive demasking (12.4%) and decrease in word naming (1.9%). This implies that low frequency words are more susceptible to influences (either inhibitory or facilitatory) from orthographic neighbors. The same conclusion was drawn by Jared et al. (1990) concerning the positive and negative influences of friends (orthographically similar words pronounced similarly) and enemies (orthographically similar words pronounced dissimilarly) on word naming performance.

The results of experiments where stimulus word frequency and neighborhood frequency are independently manipulated therefore provide strong evidence against serial search models of visual word recognition that involve a frequency-ordered selection stage (Forster, 1976; Paap et al., 1982). These models clearly predicted that words with no higher frequency neighbors should not show effects of stimulus word frequency.

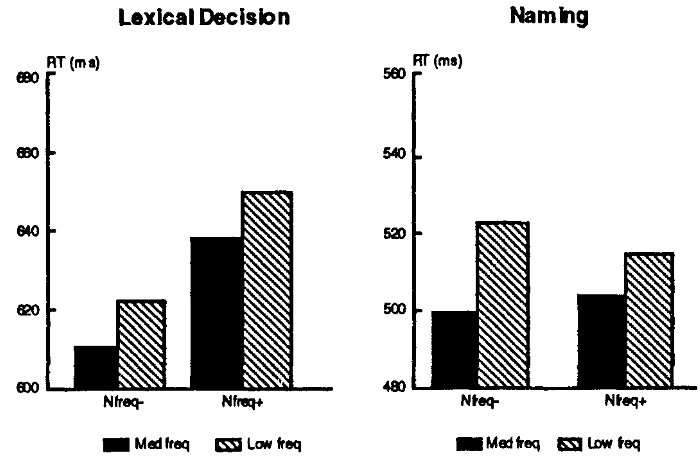

Figure 2. **Results of Grainger (1990) showing the effects of stimulus word frequency (medium frequency versus low frequency) and neighborhood frequency (Nfreq- : words with no higher frequency neighbors; Nfreq+ : words with higher frequency neighbors) observed in the lexical decision and word naming tasks.**

The results are more readily interpreted within the framework of activation based models where word frequency is reflected in the variations in resting level activations of lexical representations (McClelland & Rumelhart, 1981). In this model the effects of stimulus word frequency are a constant baseline effect which the effects of neighborhood frequency add on to. Simulations run on a French version of the model using Grainger and Segui's (1990) stimuli show the same interaction between word frequency and neighborhood frequency as observed in the progressive demasking and lexical decision error data (Figure 3). The model correctly predicts that low frequency words are more subject to neighborhood interference than are medium frequency words.

A modification of the activation verification model has recently been suggested (Grainger & Segui, 1990; Segui & Grainger, 1990) in order to accommodate the word frequency and neighborhood frequency results presented above. This modification involves abandoning a frequency ordered verification stage for a verification process triggered and ordered by lexical representations reaching a criterion activation value or verification threshold. In order to account for stimulus word frequency effects the model uses the same mechanism as the interactive activation model, that is, the resting level activations of lexical representations vary as a function of word frequency. These variations in resting level activations will affect the time it takes a given word node to reach the verification threshold and will therefore affect word recognition time. Neighborhood frequency effects are reflected in this model in terms of the probability that a word other than the stimulus word itself will reach the verification threshold first. This probability is a function of a) the number of letters this word shares with the stimulus, and b) the difference in frequency between the two words. The greater the orthographic overlap between the stimulus and its competitor and the greater the frequency difference (i.e., competitor frequency minus stimulus word frequency) the greater the probability that the competitor will be checked first. The model therefore correctly predicts that neighborhood frequency effects are stronger in low frequency words than in medium frequency words.

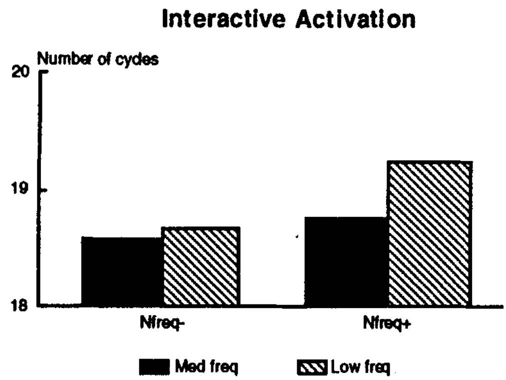

Figure 3. **Simulation of the results of Grainger and Segui (1990) run on a French version of the interactive activation model. The original parameter settings remained unchanged and number of processing cycles to reach an activation threshold of 0.7 was measured. These simulation results can be compared to the experimental results presented in Figure 1.**

Neighborhood density effects

The search metaphor has come under attack from a related series of experiments studying the role of neighborhood density in visual word recognition. Ignoring the frequencies of the stimulus word's orthographic neighbors and manipulating the number of these neighbors (Coltheart et al.'s (1977) N metric) has, however, produced conflicting results in the lexical decision task (Andrews, 1989; Coltheart et al., 1977; Grainger et al., 1989). Both Andrews and Coltheart et al. observed inhibitory effects of neighborhood density on nonword decision latencies, nonwords with many word neighbors being harder to reject than nonwords with few neighbors. However, Andrews (1989) observed facilitatory effects of neighborhood density on lexical decision latencies to words whereas Coltheart et al. (1977) and Grainger et al. (1989) found no effect of N on word responses. In the latter study, increasing the number of low frequency neighbors had no effect on lexical decision latencies or gaze durations on the stimulus word, and increasing the number of higher frequency neighbors produced a non-signficant decrease in gaze durations.

There are several possible reasons for the discrepancy between Andrews' data and the results of Coltheart et al. and Grainger et al. One is that the facilitatory effect observed by Andrews (1989) was only observed for low frequency words and was statistically robust only when the nonwords are relatively unwordlike. Coltheart et al. (1977) did not manipulate word frequency and used only orthographically regular pronounceable nonwords as in the Grainger et al. (1989) study. Moreover, the absence of an effect on N in the latter research may be quite simply due to the much smaller N values of the words compared to Andrews' stimuli.

Another possibility is that the facilitatory effect observed by Andrews (1989) can be attributed to another variable highly correlated with neighborhood density. One good candidate here is bigram frequency, since words with many neighbors typically have higher bigram frequencies than words with few neighbors. Nevertheless, the fact that bigram frequency effects are themselves notoriously difficult to obtain (Gernsbacher, 1984) is a point against this first candidate. A second candidate would be word/nonword discriminability. Low frequency words with few neighbors tend to contain unusual letter clusters (coax, fizz, flax are examples from Andrews stimulus set) and might therefore be more confusable with the nonwords used by Andrews. A final possible candidate, and one that has been examined more closely in my own research, is that it is not neighborhood density per se that is the relevant factor here but rather the number of higher frequency neighbors. This would explain why only low frequency words were affected by neighborhood density in Andrews' experiments. Neighborhood density will be strongly correlated with number of higher frequency neighbors when the stimuli are low frequency words but not when they are high frequency words.

On purely intuitive grounds neighborhood density would, in any case, seem an inappropriate variable to manipulate. Many of the words that contribute to such a count will have quite low frequencies and some may be even unknown to some subjects. One would not therefore expect such words to have much influence on word recognition performance. Moreover, at a theoretical level, frequency-ordered search models (Forster, 1976; Paap et al., 1982) do not predict an effect of neighborhood density, but an effect of number of higher frequency neighbors. Increasing number of higher frequency neighbors should lead to longer recognition latencies, according to these models. At first glance, the interactive activation model would also appear to predict that increasing the number of higher frequency neighbors should increase inhibition on the stimulus word. Simulations indicate, however, that an increase in the number of strongly activated competitors does not necessarily lead to increased inhibition on the stimulus word. These competing units all mutually inhibit (strangle) themselves. This means that when the stimulus has only one high frequency competitor this node will reach a much higher activation level during the processing of the stimulus word than when there are many high frequency competitors. Total inhibition on the stimulus word can therefore be even greater when there is only one higher frequency neighbor.

Number of higher frequency neighbors has been manipulated in two series of experiments (Grainger, 1990; Grainger et al., 1989) and the results showed no effect of this variable on lexical decision RT (although there were non-significant facilitatory trends in the lexical decision errors and the gaze duration results). These manipulations of number of higher frequency neighbors were, however, relatively weak compared to the number of higher frequency neighbors of Andrews' (1989) stimuli. The average number of higher frequency neighbors of the high density group was about four in Grainger's study whereas in Andrews' study my calculations make it closer to ten.

In recent unpublished experimental work (Grainger & Segui, 1992) we have used a stronger manipulation of number of higher frequency neighbors in the lexical decision and progressive demasking tasks. The results show a significant inhibitory effect of having one higher frequency neighbor compared to no higher frequency neighbors in both experimental tasks and a significant reduction in this inhibitory effect when the stimulus has several higher frequency neighbors. This reduction in inhibition was observed in the progressive demasking latencies and the lexical decision errors. The progressive demasking results are presented in Figure 4 along with the corresponding simulation results from the interactive activation model.

Thus, the general picture that emerges from these experiments is that the recognition of low frequency words will be adversely affected by the presence of a small number of high frequency competitors (e.g., BLUR inhibited by BLUE) but this neighborhood interference will be reduced by the presence of a large number of such high frequency competitors (e.g., HEAL: HEAR, HEAD, HEAP, REAL, MEAL, DEAL).

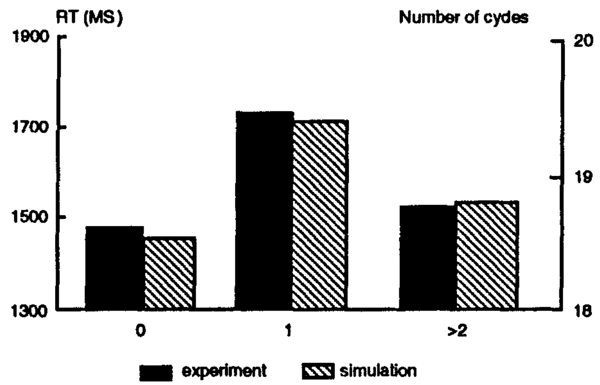

Figure 4. **Experimental results of Grainger and Segui (1992) and the corresponding simulation results from a French version of the interactive activation model. The stimuli had no higher frequency neighbors (0), a single higher frequency neighbor (1) and at least three higher frequency neighbors (>2).**

This result contradicts frequency ordered serial search models (Forster, 1976; Paap et al., 1982) which clearly predict an increase in inhibition accompanying an increase in number of higher frequency neighbors. Moreover, the modification of activation verification proposed by Grainger and Segui (1990) cannot accommodate this facilitatory effect of number of higher frequency neighbors. This modification predicts equal interference effects in words with one or several higher frequency neighbors. In this model, the magnitude of neighborhood interference is determined essentially by the frequency difference between the stimulus word and its most frequent neighbor. Number of higher frequency neighbors will not affect performance since there will be maximally one verification cycle performed before recognition. Thus, the modified activation verification model accommodates the lexical decision latency data but cannot, in its present form, explain the reduced inhibition observed with many higher frequency neighbors in the progressive demasking paradigm. It remains to be seen whether further developments of this type of model will enable it to accommodate this particular result.

From a purely empirical point of view, this result suggests that the facilitatory effect of neighborhood density on lexical decision latencies to low frequency words observed by Andrews (1989) may in fact reflect a reduced inhibition resulting from the greater number of higher frequency neighbors of the high density stimuli. Research is currently underway in an attempt to separate out the effects of neighborhood density and number of higher frequency neighbors in visual word recognition tasks.

Enhancing neighborhood effects

The importance of relative frequency (low target word frequency versus high competitor frequency) and orthographic overlap in determining neighborhood interference is supported by data obtained using the masked priming paradigm. This paradigm provides a simple methodological tool for manipulating the hypothetical competitiveness of a given target word's neighbors. Using very brief prime presentation durations and massive forward masking the prime stimuli are barely visible to subjects. Nevertheless, even in such extreme conditions, it is hypothesized that processing has been initiated on the prime stimulus causing a rise in activation of any representations involved in such processing. Thus, when the target stimulus is presented immediately after prime offset, a certain number of representations will be in a heightened state of activation when processing begins on the target word. Now, if this target word shares properties (orthographic or other) with the prime then the activation levels of representations that were raised during prime processing will continue to be supported by information from the target word. It is these representations that remain in a heightened state of activation during target processing that will influence target recognition.

Using this paradigm it has been observed that medium frequency targets are inhibited by the prior presentation of a high frequency orthographically related prime, whereas the same medium frequency targets are uninfluenced by a low frequency orthographically related prime (Segui & Grainger, 1990). This important result suggests that, due to their lower resting level activation, low frequency words do not attain a high enough activation level in these extreme presentation conditions to provoke noticeable interference on target processing.

Both the interactive activation model (McClelland & Rumelhart, 1981) and the modification of the activation verification model presented above can accommodate this basic result. Simulations run on a French version of the interactive activation model provide a reasonable reflection of the experimental data obtained with the same set of word stimuli. In Figure 5 are presented the experimental results from Segui and Grainger (1990, Experiment 2) and the corresponding simulation study with the interactive activation model.

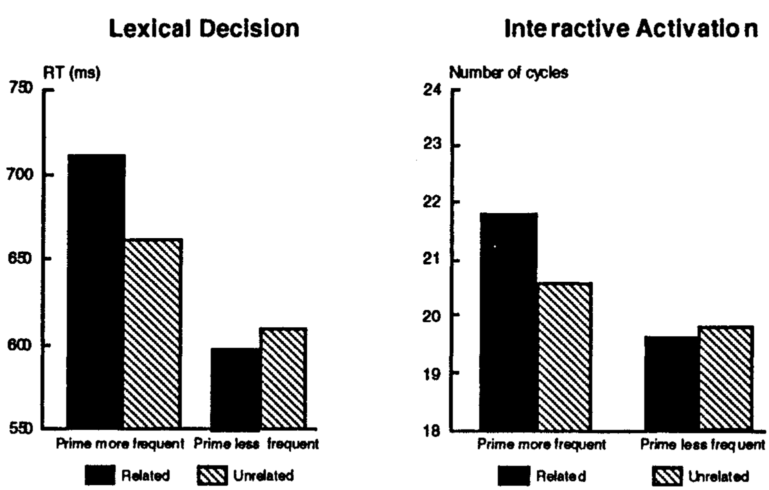

Figure 5. **Experimental results of Segui and Grainger (1990, Experiment 2) and the corresponding simulation study run on a French version of interactive activation. Word targets were preceded by orthographically related or unrelated briefly presented masked word primes. These primes could be more frequent than the target (prime high frequency, target low frequency) or less frequent than the target (prime low frequency, target high frequency).**

In the modified activation verification model, increasing the activation level of a lexical representation other than the target word will increase the probability that this representation reaches the verification threshold before the target. This increase in probability will then be reflected in larger average recognition times due to the increase in the number of occasions the wrong candidate is checked first. The model can therefore, in principle, capture these inhibitory effects of masked orthographic priming. However, only when an implemented version is available will a thorough test of this model be possible.

In more recent experimental work we have begun to isolate the various factors that determine the amount and the direction (facilitation or inhibition) of form priming effects in the lexical decision task. The general rule that emerges, from this and related research, is that inhibition is greatest when a lexical representation is maximally activated compared to the target representation during prime processing and when this activation is then maximally supported by the target stimulus. Two different observations support this analysis.

1) Primes that are word neighbors of the target produce inhibition which is strongest when the prime is more frequent than the target (Parpaillon, 1991; Segui & Grainger, 1990), whereas nonword neighbor primes tend to facilitate target processing (Ferrand & Grainger, in press; Forster, Davis, Shoknecht, & Carter, 1987).

2) Increasing prime duration produces a rise in inhibition (up to a critical duration of about 100ms) whereas reducing prime duration tends to produce facilitation (Humphreys, Evett, Quinlan, & Besner, 1987; Parpaillon, 1991; Ferrand & Grainger, in press). The results of a variable SOA study run by Parpaillon (1991) and the corresponding simulation study run on interactive activation are shown in Figure 6.

These results add further support to a general view of the word recognition process as a cascaded hierarchical activation- based process in which activation at sublexical levels (letters, letter clusters, phonological units) feeds forward to the lexical level (with or without subsequent top-down feedback). With very brief prime presentation durations (10-30ms) only sublexical representations will be significantly activated at target onset, particularly when the prime is a nonword or a low frequency word. In this situation there is not enough activity generated at the lexical level to provoke within-level competition, thus allowing sublexical facilitatory effects to emerge. As primes activate lexical representations more and more compared to target activation (with increased prime frequency relative to target frequency and/or increased prime duration) then lexical level competition increases, thus cancelling sublexical facilitation.

Visual factors in neighborhood interference

In hierarchical activation-based models of visual word recognition such as interactive activation, the degree of activation of the stimulus word's component letters influences the activation level of lexical representations containing these letters. Now, there is recent evidence (Jacobs & Grainger, 1991) that isolated letters prime the identification of physically similar letters (e.g., E-F) in the masked priming paradigm combined with an alphabetic decision task (letter/non-letter discrimination). According to hierarchical models in which letters are fundamental intermediate units in visual word recognition, this letter-letter priming should affect activation levels at the word level. In other words, the stimulus PEAT should increase the activation of the lexical representation for BEAT more than that of the lexical representation for SEAT (these two words have approximately equivalent printed frequencies) since P and B are more similar than P and S.

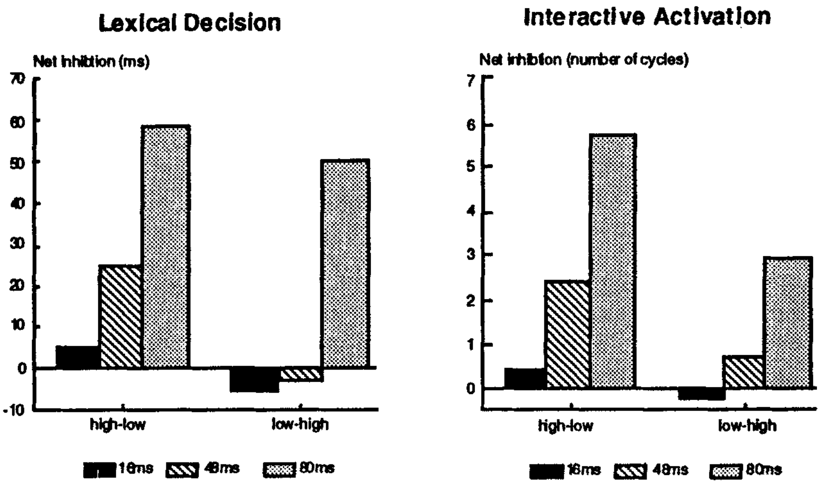

Figure 6. **Net inhibition (RT in the orthographically related condition minus RT in the unrelated condition) observed in the variable (16ms, 48ms, and 80ms) SOA orthographic priming study of Parpaillon (1991). Primes were high frequency words and targets low frequency words (high-low) or the opposite (low-high).**

In the study by Grainger (1990) there was no detectable influence of the similarity of the two critical letters in words with a single higher frequency neighbor. For example, in English the neighboring words BLUE and BLUR have the critical letter pair e/r distinguishing them. Critical letter pairs that were maximally dissimilar for lower case letters (e.g., k/y) did not appear to reduce interference in this study. However, these critical letters were always external (initial or final) letters of 4-letter words and were therefore highly visible when fixating the center of the word. Due to lateral interactions between juxtaposed letters in character arrays, the internal letters of a word are less visible than the external letters that suffer from less lateral inhibition. One might therefore expect that the similarity of the two critical letters in words with a single higher frequency neighbor would play a larger role with internal letters.

Rather than manipulating the similarity of these critical letter pairs (since this imposes extreme limitations on stimulus selection), in some recent experiments the visibility of the critical disambiguating letter in a word with a single higher frequency neighbor was varied. Grainger, O'Regan, Jacobs, and Segui (1992) observed that manipulating the visibility of this critical letter, by varying the position of initial eye fixation on the word, influences the inhibitory effects of the higher frequency neighbor. Thus, for example, the French word CHOPE (tankard) is a low frequency word that has a single higher frequency neighbor CHOSE (thing). The letter P is therefore the critical disambiguating letter in the target word CHOPE. We observed that fixating the critical disambiguating letter in a word with a single higher frequency neighbor caused a significant reduction in neighborhood interference measured relative to control words with no higher frequency neighbors. This particular result is presented in Figure 7.

In order to explain this result one must assume that the sublexical units underlying the recognition of a particular word are not all equally activated during the processing of that word. It is a well established fact that the quality of visual information sampled by the retina (visual acuity) rapidly reduces as soon as one moves away from the fixation point on the fovea (see O'Regan, 1990, for a discussion on this point). This means that when fixating a string of letters, those letters on or next to the fixation point will be maximally visible and the other letters gradually less and less visible, with the exception of the external letters, which suffer less lateral inhibition. If individual letter representations are activated during the processing of a word, then the letter on fixation will be maximally activated, with the activation levels of the other letters varying as a function of their position in the string.

As a simple example to illustrate this point, imagine that the maximum letter activation is value 0.9 on fixation and that this value drops by 0.2 for each letter position away from fixation with 0.3 being added to compensate for reduced lateral inhibition on external letters (the values chosen here are completely arbitrary). Now, take the letter string CHOPE with its higher frequency neighbor CHOSE. When the eye fixates the second letter of CHOPE the hypothetical activation values for letters one to five will be 1.0, 0.9, 0.7, 0.5, and 0.6. If one accepts that the activation values of word units are a simple sum of the activation values of the component letters when ignoring word frequency (McConckie, Kerr, Reddix, Zola, & Jacobs, 1989; Nazir, O'Regan, & Jacobs, 1991), then the word CHOPE will have an activation value of 3.7 and its higher frequency neighbor CHOSE will have an activation value of 3.2 (0.86% of the value of CHOPE). When fixation is on the fourth letter, then the activation values of letters one to five will be 0.6, 0.5, 0.7, 0.9, 1.0. CHOPE still has an activation value of 3.7 but its higher frequency neighbor CHOSE now has an activation value of 2.8 (0.76% of the value of CHOPE).

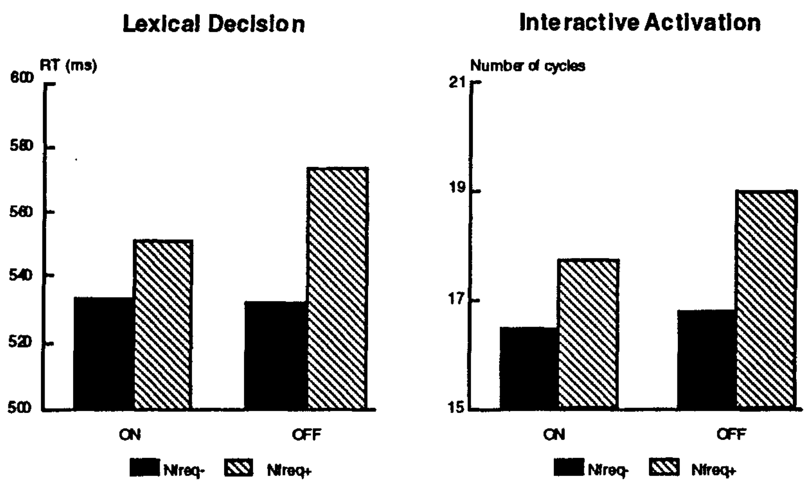

Figure 7. **Results of Grainger et al. (1992) and the corresponding simulation run on the interactive activation model, showing response latencies to words with (Nfreq+) or without (Nfreq-) a single higher frequency neighbor. Initial fixation position on the stimulus was either ON or OFF the critical letter distinguishing a word from its higher frequency neighbor. In the simulation the values of the ESTR parameter were altered to simulate the different fixation positions in the stimulus word.**

If neighborhood interference is a function of the relative activation levels of the stimulus word and its higher frequency neighbor, then this simple model correctly predicts that interference is greater when fixation is off the critical disambiguating letter in a word with a single higher frequency neighbor. These variations in letter visibility can actually be simulated in the interactive activation model using the ESTR parameter that determines the strength of feature- level activation at each letter position. The simulation results using the set of parameter values given above (i.e., 1.0, 0.9, 0.7, 0.5, 0.6 for fixation on the second letter and 0.6, 0.5, 0.7, 0.9, 1.0 for fixation on the fourth letter) are given in Figure 7.

In the same experiments (Grainger et al., 1992), it was also observed that neighborhood frequency effects are much stronger for words that differ from their higher frequency neighbor by their fourth letter (e.g., CHOPE similar to CHOSE) compared to words that differ from their higher frequency neighbor by the second letter (e.g., ASTRE similar to AUTRE). These results were interpreted within the framework of a hierarchical activation model as implying that initial letters provide relatively more activation input to the word level than do the end letters of a string. In this way, the activation of the competitor CHOSE would reach a higher level during the processing of the stimulus CHOSE than would the competitor AUTRE during the processing of the stimulus ASTRE. Nevertheless, it should be pointed out that simulations run on the interactive activation model without any such word initial bias introduced, show the same pattern of effects as the experimental data. Stronger neighborhood frequency effects were observed to CHOPE type words (2.2 cycles) than to ASTRE type words (1.3 cycles). This therefore suggests that the observed difference between CHOPE and ASTRE type words may be partly due to mechanisms other than variations in letter visibility or the strength of letter-word connections.

Once again a simple hierarchical activation model, in which competition is determined by the relative activation levels of lexical representations, provides a coherent explanation of an otherwise complicated pattern of results. Both the interactive activation model (McClelland & Rumelhart, 1981) and a modified activation verification model (Grainger & Segui, 1990; Segui & Grainger, 1990) are of this type.

Conclusions: Neighborhood effects and models of visual word recognition

The present chapter has shown that much of the research manipulating word frequency and orthographic neighborhoods in visual word recognition can be captured by a cascaded hierarchical activation model of the type proposed by McClelland and Rumelhart (1981). The key features of this type of model that allow it to simulate the experimental results presented here are: 1) its hierarchical structure with feedforward facilitatory activation flowing from sublexical levels to the lexical level, and 2) the mutual inhibition between simultaneously active word nodes. Other fundamental aspects of the interactive activation model, such as the existence of top-down facilitatory feedback from words to letters, have not been dealt with here. However, it should be pointed out that simulations run without any word-letter feedback indicate that the model is just as capable of accomodating the above results in a non-interactive form (Jacobs & Grainger, in press). It should also be pointed out that, apart from data obtained with the two-alternative forced choice paradigm (Reicher, 1969), the empirical evidence available for the existence of automatic top-down facilitation is very weak (see Grainger & Jacobs, 1991).

The results presented in this chapter were also compared with a modified activation verification model which, in general, fared extremely well in accomodating the data. This type of model adopts many of the essential charateristics of the interactive activation model while replacing the mechanism of mutual inhibition with a verification procedure. The verification mechanism failed to capture one critical aspect of the results, the fact that a reduction in inhibition is observed when the number of higher frequency neighbors is increased (Grainger & Segui, 1992). Clearly, further experimental work is required on this very critical point. Since the number of higher frequency neighbors is very strongly correlated with measures of sublexical letter cluster frequency (e.g., bigram and trigram frequency), it needs to be made quite clear that the effects of this variable can be unequivocally attributed to processes operating at the lexical rather than the sublexical level.

A further alternative to the mechanism of mutual inhibition as embodied within the interactive activation model was proposed more recently by Seidenberg and McClelland (1989). Their distributed developmental model is another example of a cascaded hierarchical activation model. Although no lexical representations exist in this model, there are a set of lower-level (input) units and higher-level (hidden) units that serve to encode the same information that is represented in the sublexical and lexical levels of local, non-distributed, interactive activation type models. The essential characteristic distinguishing this type of model from the interactive activation family is that there is no within-level inhibition (this is actually impossible in the model since the same nodes are used to represent different words). In this respect, the model (in its stable state) can be thought of as a distributional variety of the logogen model (Morton, 1969) with between-level connections but no within-level connections. In the Seidenberg and McClelland model, competitive processes are not operational during word recognition but rather have occurred during the learning phase and are encoded in the variations in weights of the between-level connections.

Seidenberg and McClelland (1989) show, rather impressively, how their model can handle most of the major empirical results on word and nonword naming (although see Besner, Twilley, McCann, and Seergobin (1990) for a critique). It remains to be seen, however, if the model can give an equally good account of word recognition performance in general and lexical decision performance in particular. In this model, lexical decisions are made on the basis of one or more computed codes (orthographic, phonological, and semantic). When reliable word/nonword discriminations can be made on the basis of orthographic information alone, then it is the orthographic error score computed by the model that is used to make lexical decisions; the smaller the error score the faster the lexical decision latencies and the lower the error rate. The orthographic error score is computed by comparing the pattern of activation generated across the input units by the orthographic input with the pattern generated by feedback from the hidden units.

Since only further simulation work will directly answer this question, it is difficult to evaluate whether the model can account for the effects of neighborhood interference discussed above (Grainger, 1990; Grainger & Segui, 1990; Grainger et al., 1989). It is conceivable (within the limits of my own understanding of the model) that larger orthographic error scores will be computed for words with a higher frequency neighbor. The stimulus BLUR, for example, may be generating a pattern of activation across the hidden units that reflects the model's training on the more frequent word BLUE. The feedback from these hidden units, which is then compared to the pattern of activation across the orthographic input units, may therefore generate a higher error score than for a word with no such higher frequency neighbor. However, it is difficult to imagine how this could be compatible with the fact that in the model the prior training on the high frequency word BLUE will lead to the computation of a lower phonological error score for its lower frequency friend FLUE. Clearly more simulation work is necessary here in order to provide a more complete evaluation of this particular model as a model of visual word recognition.

Further research on the role of orthographic neighborhoods in visual word recognition should provide the data necessary to constrain these different hierarchical activation models. In particular, the present chapter has shown that two very different mechanisms (mutual inhibition and verification) integrated within this general framework can accommodate the majority of the results at present. Future research should help decide which, if any, of these two mechanisms provides the most satisfactory explanation.

References

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

Becker, C. A. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception and Perfromance, 2,* 802-814.

Besner, D., Twilley, L., McCann, R. S., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review, 97,* 432-446.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI.* New York: Academic Press.

Ferrand, L., & Grainger, J. (in press). Phonology and orthography in visual word recognition: Evidence from masked nonword priming. *Quarterly Journal of Experimental Psychology.*

Feustel, T. C., Shiffrin, R. M., & Salasoo, A. (1983). Episodic and lexical contributions to the repetition effect in word identification. *Journal of Experimental Psychology: General, 112,* 309-346.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J.Wales & E. W. Walker (Eds.), *New approaches to language mechanisms.* Amsterdam: North-Holland.

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.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interaction between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113,* 256-281.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language, 29,* 228-244.

Grainger, J., & Jacobs, A. M. (1991). Masked constituent letter priming in an alphabetic decision task. *European Journal of Cognitive Psychology, 3,* 413-434.

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., O'Regan, J. K., Jacobs, A. M, & Segui, J. (1992). Neighborhood frequency effects and letter visibility in visual word recognition. *Perception and Psychophysics, 51,* 49-56.

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

Grainger, J., & Segui, J. (1992). Neighborhood frequency and neighborhood density effects in visual word recognition. Manuscript in preparation.

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.

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

Humphreys, G. W., Evett, L. J., Quinlan, P. T., & Besner, D. (1987). Orthographic priming: Qualitative differences between priming from identified and unidentified primes. In M. Coltheart (Ed.), *Attention and performance XII.* London: Lawrence Erlbaum.

Jacobs, A. M., & Grainger, J. (1991). Automatic letter priming in an alphabetic decision task. Perception and Psychophysics, 49, 43-52.

Jacobs, A. M. & Grainger, J. (in press). Testing a semi-stochastic variant of the interactive activation model in different word recognition experiments. *Journal of Experimental Psychology: Human Perception and Performance.*

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29,* 687-715.

Lukatela, G., & Turvey, M.T. (1990). Automatic and pre-lexical computation of phonology in visual word identification. *European Journal of Cognitive Psychology, 2,* 325-343.

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 and Psychophysics, 46,* 245-253.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General, 118,* 43-71.

Morton, J. (1969) Interaction of information in word recognition. *Psychological Review, 76,* 165-178.

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,* 375-405.

Nazir, T., O'Regan, J. K., & Jacobs, A. M. (1991). On words and their letters. *Bulletin of the Psychonomic Society, 29,* 171-174.

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). Elsevier, Amsterdam.

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: The word superiority effect. *Psychological Review, 89,* 573-594.

Parpaillon, F. (1991). Modélisation connexionniste de l'effet de voisinage orthographique. Unpublished dissertation (DEA), Université de Paris Sud, Orsay.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81,* 275-280.

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,* 1, 65-76.

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

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 8

On the Role of Cohorts or Neighbors in Visual Word Recognition

Neal F. Johnson
The Ohio State University

From an intuitive perspective, the process of visual word recognition seems quite simple. We see some letters; we put them together in our mind (which makes the word); that tells us what was written; and then we know what was meant. Unfortunately, each of these rather simplistic points conceals what appears to be some very complex processing.

For example, it would follow from such a view that perceivers should be able to make meaning-based judgments only after they know the name of the word they are seeing, but the work on "unconscious perception" raises serious doubts as to whether that is always the case (Marcel, 1983a, 1983b). In addition, that view would suggest that we should know the letters within a word before we know which word we are viewing, but the data also seem inconsistent with that effect (e.g., Johnson, 1975, 1991, and also see Huey's, 1908/1968, description of the Erdmann & Dodge experiments).

On the other hand, confirming evidence for our simplistic intuitive model also can be readily obtained. For example, the assumption that initial orthographic encoding (i.e., "seeing" the letters) must be part of word recognition is clearly supported by demonstrations that perceivers have no lexical or semantic information when there are no letters displayed. While this point seems obvious, its reality in conjunction with the fact that perceivers can identify words before they can identify letters within words, even when the two tasks require exactly the same visual information (Johnson, Turner-Lyga, & Pettegrew, 1986; Sloboda, 1976, 1977), points out the complexity of the visual processing that precedes lexical access. The problem, then, is to formulate a model of visual word recognition that reconciles our naive intuitions with the kinds of counterintuitive data we frequently obtain in experiments.

The model

In this context, the cohort model of visual word recognition (Johnson, 1992; Johnson & Pugh, 1992) was based on an earlier model of spoken word recognition (Marslen-Wilson & Welch, 1978; Marslen-Wilson, 1987). In addition to the obvious fact that letters need to be displayed, the model was formulated within the constraints of three critical parameters. The first is the fact that words conceal their letters (Drewnowski & Healy, 1977; Healy & Drewnowski, 1983; Johnson, 1991); the second is the fact that lexical processing is facilitated by increasing word frequency; and the third is the increasing body of data indicating that lexical processing is influenced (albeit, in very complicated ways) by the number of words that are visually similar to the target (Andrews, 1989; Coltheart, Davelaar, Jonasson, & Besner, 1977; Grainger, 1990; Grainger, O'Regan, Jacobs, & Segui, 1989; Johnson & Pugh, 1992; Stadtlander & Krueger, 1992). It is the manner in which this particular view of visual word recognition handles this latter issue that makes it a cohort model, and what follows is a rather abbreviated summary of the critical components of the model (Johnson, 1992; Johnson & Pugh, 1992), as well as a review of some of the relevant data.

General considerations

In terms of its global structure, the cohort model (Johnson, 1992) can be divided into two major constituents, with the first involving the visual encoding of the orthographic information, and the second being a characterization of the manner in which the orthographic encodings are used to access the lexicon and the lexical entry's semantic representation. These are viewed as being independent but overlapping stages of processing to the extent that the moment any letter information reaches its final level of orthographic encoding it is assumed to provide activation support for any lexical entries with which it is consistent, and activation is withdrawn from any activated entries with which it is inconsistent.

Orthographic concepts

Within the orthographic component of the model the relevant units of representation are: 1) feature codes; 2) letter codes; and 3) abstract representations that conform to letter triples, which are referred to as *wickelgraphs.* These latter units (wickelgraphs) are intended as representations of letter identities in their orthographic context, and, as an example, the wickelgraphs for the word *THE* would be #Th, tHe, and hE#, where # represents a word boundary.

Although this strategy for capturing orthographic context has been used by others (e.g., Brown, 1987; Seidenberg & McClelland, 1989), in this case it is assumed that the triple is not encoded as a simple trigram. That is, the wickelgraph is an encoding of the identity of the critical letter (the H in tHe), plus an opaque representation of the context letters, with the latter portion of the encoding being used as a means of defining the orthographic context in which the wickelgraph's identity can appear.

The context information within a wickelgraph is assumed to specify the manner in which that wickelgraph can be integrated with other wickelgraphs to form the patterned orthographic encoding that is used to access the lexical entry. This latter representation of the display (i.e., the pattern) is viewed as being integrated and unitary, as well as abstract, and in that sense, it is assumed that letters do not maintain any separate integrity or identity with the pattern.

Codes and knowledge states

Lexical entries, as well as letter codes, are viewed as open-ended knowledge states to which new information can be added as it is acquired. For example, a letter code is viewed as an abstract representation, which includes both information regarding the letter's identity and algorithms for instantiating the letter in a variety of graphic forms, with new algorithms being added as they are learned.

Similarly, a lexical entry for a word would be established once its auditory-acoustic (phonetic) form can be recognized. However, only later would that knowledge state be elaborated to include an articulatory routine and an orthographic representation. In addition, any syntactic information needed for the appropriate use of the word in a grammatical context also would be added at later points in time as it is acquired.

The concept of activation

Finally, in terms of general considerations, it is important to note that the model does not assume that information represented within these knowledge states is in any way involved in the recognition of words. Word recognition is assumed to be the activation of the lexical entry, as well as any subsequent semantic encoding required by the task, and the content of the lexical entry would become available only after the recognition process (activation) had been completed.

Terms like *pattern-matching* and *self-addressing* have been used to label retrieval events like those involved in lexical access, and such terms imply that the overlapping content between the antecedent event (i.e., the retrieval cue) and the retrieved or activated encoding (e.g., a lexical entry) is in some way critical to the activation or retrieval. Within this model retrieval is assumed to be the activation of a target encoding, and it is based on a simple contentless association between the retrieval or activating cue and the lexical entry. In addition, the strength of the association is assumed to be a simple function of the frequency of the pairing, and in that sense this is clearly a strength model of memory.

The initial orthographic encoding

Within the model, the level of orthographic processing is divided into three stages. These stages conform to the encoding of features, letters, and wickelgraphs, with the wickelgraphic encoding being the process that creates the abstract patterned representation of the orthographic information in the display (i.e., the abstract orthographic pattern). In addition, these three stages seem to sort themselves into just two self-contained processing modules. The first is involved with the encoding of simple features, while the second module appears to be recruited whenever there is a need to encode structural relationships, which would include both the encoding of features into structured letter units, and the higher-order encoding of the letter units into structured patterns of letters.

For example, Compton, Grossenbacher, Posner, and Tucker (1991) found that perceivers were very fast at detecting whether a letter within a letter array had one feature that was thicker than the other features. In addition, although latencies were found to be a function of the physical distance between the point of fixation and the location of the thick feature, they were not influenced by either the physical (linear) or cognitive (orthographic regularity) structure of the display.

In a second task subjects were to determine whether the display contained a lower-case letter mixed in with the otherwise all upper-case letters (physical size was held constant). Clearly, the case of a letter can be detected only after its features have been structurally encoded, and the data indicated that not only were the latencies longer than for simple feature detection, but they were a function of both the physical and the cognitive structure of the display. Feature detection was not influenced by the display's structural characteristics, while letter detection was delayed until after the structure of both the letter and the pattern as a whole had been encoded, and that suggests two modules, with the second being involved in the encoding of structural information.

The structure of the orthographic component

The functional architecture of this component of the model is identified with that of the perceptual representation system (PRS) (Tulving & Schacter, 1990), particularly with regard to the manner in which the PRS relates to the types of structural distinctions illustrated by recent experiments employing the PET technique (Peterson, Fox, Posner, Mintun, & Raichle, 1988; Peterson, Fox, Snyder, & Raichle, 1990; Posner, Peterson, Fox, & Raichle, 1988). In addition, the data on form priming (e.g., Schacter, 1990; Schacter, Cooper, & Delaney, 1990; Schacter, Cooper, Delaney, Peterson, & Tharan, 1991) offer very specific documentation for the types of processing assumed to occur within the structural-encoding phase of the model.

PET seems to detect very localized metabolic activity in the brain, and one of the effects that has been demonstrated is nonoverlapping distributions of activity for visual word recognition (occipital cortex) and spoken word recognition (parietal and temporoparietal cortex). In addition, both areas show clear differences between words and nonlanguage stimuli (e.g., random sounds or meaningless visual figures), suggesting that within both the auditory and the visual processing systems there are language specific encoding mechanisms.

However, in addition to each modality's language-specific effects, within the visual system there is also evidence for mechanisms that are specifically sensitive to structured versus unstructured letter arrays. For example, the PET data indicates that both orthographically regular and irregular displays show activity within the striate cortex, but only orthographically regular displays (both words and nonwords) also show activity in the extrastriate cortex.

With regard to visual word recognition, those data suggest that within the module that the Compton et al. (1991) data indicate is involved with structural encoding, there are really two identifiable subcomponents. One of these subcomponents is the encoding of the structural relationships among the features that form an individual letter, as documented by Compton et al., and the other is the encoding of the higher-order relationships among the letters, with those relationships then forming the structural basis for the pattern as a whole. If the displays are orthographically irregular, then the letter pattern has no structure that could be encoded, and there also is no evidence of extrastriate activity, but if the displays are structured (i.e., orthographically regular) then there is such extrastriate activity.

Feature and Letter Encoding. The model assumes that simple letter features (straight lines, diagonals, curves, etc.) are encoded, and they are assembled into higher-order feature assemblages conforming to letters when juncture features are encoded. Juncture features are the relationships between simple features, such as the three points of intersection of the simple features that make up an upper-case A, but they also included more complex relationships such as the gap in the curve that differentiates an O from a C.

Feature assemblages provide activation to letter codes with which they are consistent, and when a letter is detected its identity information is contributed to a linear array of encoded letter identities which represents the initial orthographic encoding for the display. In addition, these latter encodings are abstract in that they represent the letter's identity, but not its graphic form.

The Encoding of Wickelgraphs. Wickelgraphs are higher-order encodings of the letter identities (see Wickelgren, 1969). A wickelgraph for a letter position has the target letter's identity as its transparent core, but it also contains the contents of its two immediately adjacent letter positions as an opaque context. It is this context that defines the manner in which the target letter's identity can be combined with the identities of other wickelgraphs to form the abstract orthographic pattern that is used to activate lexical entries. In addition, wickelgraphs are assumed to be preexisting knowledge states that are based on previous encounters with the letter patterns. Given that is the case, they define the specific letter-identity contexts in which a particular letter identity can appear, and in that sense they are assumed to instantiate the rules of orthography (i.e., the rules of orthography are simply formalizations that describe the set of possible wickelgraphs).

It is also assumed that there are two routes, one direct and one computational, whereby a letter triple can be encoded into a wickelgraph. The direct route is assumed to be a case in which the triple directly activates its appropriate wickelgraph, whereas the computational route might include something like initial activation of encodings for the two bigrams followed by their activating the wickelgraph. In any event, the direct route is assumed to be faster than the computational route, and the likelihood of there being a direct route for a wickelgraph would be an increasing function of the cultural frequency of the triple. For these reasons, then, the rate of orthographic encoding is assumed to be a function of the cultural frequency of the word's wickelgraphs, and in general, high-frequency words will have high-frequency wickelgraphs.

The role of wickelgraphs in the encoding of the orthographic pattern

The abstract orthographic pattern is established the moment the first wickelgraph is encoded, and the status of the pattern at that point acts as a retrieval cue for activating all lexical entries with which it is consistent (i.e., the word's initial cohort). The pattern gradually evolves as increasing numbers of wickelgraphs are subsequently encoded into it, and during that process activation is withdrawn from lexical entries as they become inconsistent with the current status of the pattern. The cohort is resolved when all nontarget members of the cohort have lost activation.

Within the orthographic encoding component of the model, the critical concepts are the ideas of a wickelgraph and the abstract orthographic pattern, and the function of wickelgraphs is to provide a means for explaining the reader's knowledge of orthography in a form that is both plausible and consistent with available data. For example, in letter migration studies (McClelland & Mozer, 1986; Treisman & Souther, 1986) subjects are shown brief displays that consist of several words, and they are to detect whether a predesignated target word is present. If the predesignated target is not in the display set, but all its letters do appear in words that are in the set, the false positive rate is very high in comparison to a case in which the target contains a letter that is not in the display set.

However, Carnot (1988) demonstrated that not only must the target's letters appear in the displayed words, but they also must be in the same intraword letter position. If the letter position was changed, the false positive rate was even lower that if a different letter was included. Furthermore, that same low false positive rate occurred even when the letter position was not changed, but the "migration" would put the critical letter adjacent to a letter that is different from the one to which it was adjacent in the original displayed word.

For example, in a very brief display consisting of DAM, FAN, SIT, and HIP, perceivers might "see" DAN or SIP, but the false positive rate on LIP or CAN would be quite low, and it would be even lower for MAN and FIT. It appears, then, that at a point in time when the false positive rate on DAN and SIP indicates perceivers do not have the letters appropriately sorted into their correct word groupings, the unusually low false positive rate on MAN and FIT indicates that the letters are encoded in terms of their immediate local context (i.e., a wickelgraph).

Similarly, Travers (1973, 1974) presented a word for 48 ms, or the letters appeared one at a time for 48 ms each across the screen, with a masking stimulus appearing immediately to the left of each letter to prevent the co-occurrence of the letter and an iconic image of its predecessor. The results indicated that even though the display time for each letter was the same, identification rate was far superior when all the letters appeared together.

In addition, in a third condition the words appeared as a moving window of two letters each (e.g. for STOP the displays would be S, ST, TO, OP, and P), with each of the displays lasting for 24 ms (again, a total of 48 ms per letter). There also was a fourth condition which employed a 16-msec moving window of three-letter displays (e.g., S, ST, STO, TOP, OP, P) (again, a total of 48 ms per letter). As for the single letters, each display had a masking stimulus to the immediate left of the displayed letters to prevent a visual carry over from the preceding display.

The results from these two conditions indicated that performance improved with increasing window size, and performance with a window of three letters was about the same as for a single 48-msec display of the whole word. Even more interesting, however, with regard to the concept of a wickelgraph, not only did these facilitating context effects seem to have their maximum effect in wickelgraph-size units, but they occurred only for letter arrays for which the model would expect letter triples to match preexisting wickelgraphic encodings (i.e., orthographically regular displays). In general, then, these context effects indicate that very early in processing target letters are encoded in terms of their immediate orthographic context, given that context is orthographically regular, and those empirical effects are a very close match to the foregoing theoretical definition of a wickelgraph.

The abstract orthographic pattern

Within the model, the main function of the abstract orthographic pattern is to provide a mechanism that will account for the fact that words appear to be perceived holistically. For example, Healy and Drewnowski's unitization model (Drewnowski & Healy, 1977; Healy & Drewnowski, 1983) assumes that words conceal their letters, and their data from proofreading types of tasks indicates that to be the case. The pattern-unit model (Johnson, 1977, 1981) also makes such an assumption, and the word-priority effect is consistent with that expectation. That is, subjects can identify a displayed word faster than they can identify a letter within the word, even when the word identification requires the same visual information as does the letter identification (Johnson, Turner-Lyga, & Pettegrew, 1986; Sloboda, 1976, 1977). In addition, when perceivers are asked to search a displayed word for a target letter they appear to scan their memory rather than the display, although if the letter array is an unstructured string of consonants they seem to scan the display itself (Johnson, 1986; Johnson, Pugh, & Blum, 1989).

The role of attention in the word-priority effect. One account for this concealment or word-priority effect (Johnston & McClelland, 1980) is that it does not reflect any influence of the encoding process, but rather it is a function of the way in which perceivers allocate attention to the display once it is encoded. Consistent with that account, Johnson and Blum (1988) demonstrated that the letter concealment stemming from the presence of other letters in a consonant array can be eliminated by allowing subjects to prefocus their attention. However, when the displays were words, Marmurek (1987) demonstrated that although controlling the subjects' attention did modulate the word-priority effect, it was not eliminated, and that suggests that the letter concealment evident in the word-priority effect may stem from some factor in addition to the manner in which attention is allocated.

The manner in which this model handles the word-priority effect is to assume that attention is allocated to the encoding of the display in a top-down manner, just as does the Johnston and McClelland (1980) model. However, in addition to that, it is assumed that letter identities are doubly concealed by first being encoded into wickelgraphs, and then the wickelgraphs are encoded into the abstract orthographic pattern. Within the abstract orthographic pattern, it is assumed that neither individual letters nor wickelgraphs maintain any separate identity, and that the pattern as a whole is the unit of encoding. In addition, for that same reason, it is assumed that at any point in its evolution during the process of orthographic encoding, it is the then-current status of the pattern as a whole that functions as the retrieval cue or source of activation for consistent lexical entries.

The abstract orthographic pattern and the PRS. Clearly, this view of orthographic encoding assumes that the final product is an abstract structural description of the presented display, and for that reason, the encoding that is assumed to occur within this component of the model fits very closely with the types of encoding that are assumed to occur within the perceptual representation system (PRS) (Tulving & Schacter, 1990). The major experimental paradigm employed to examine the PRS involves priming within an implicit memory task. Subjects see a set of display stimuli in the context of some type of cover task, and at a later point they are to react in some way to those same stimuli in a second task. The priming effect is their ability to respond to these already-seen stimuli more quickly and/or more accurately than to a comparable stimulus set that had not been recently experienced.

For example, Schacter, Cooper, & Delaney (1990) and Schacter et al. (1991) showed subjects line drawings of three-dimensional unfamiliar objects, but half of them had a structural defect that would make them physically impossible to realize. The subjects' priming task in one condition was to judge whether the displayed object was oriented to the right or the left, and in another condition it was to judge whether the line drawing contained more vertical or horizontal lines. The assumption was that orientation judgments would require the subjects to integrate the structural information from the visual display of the object, whereas the line judgments would discourage such integration. To the extent that the encoding of the display within the PRS is in the form of an abstract structural description, these two priming-task activities should be differentially effective in establishing that encoding. In addition, in some conditions subjects also were asked to elaborate the display semantically during the priming task, which should enhance their explicit or recollective memory for the object.

The same objects then appeared in the target task, along with an equal number of previously unseen possible and impossible objects, and subjects made possible/impossible judgments (i.e., would it be possible for this object to exist in the real world). In this second task the displays were very brief and the measure was accuracy. In addition, however, there were conditions in which the second task involved the explicit recognition of whether the displayed object had appeared in the first list.

Their data indicated, first, that variables that enhanced explicit memory (e.g., semantic elaboration) did not influence the implicit-memory effects within the PRS, as reflected in the accuracy scores for the possible/impossible judgment task. Similarly, in comparison to the line-judgment task, the prior judgments of orientation facilitated performance in the implicit-memory task, but it had no effect on explicit memory. Clearly, the visual encoding of the display within the PRS is independent of the episodic memorial representation of the earlier event.

With regard to the idea that the representation within the PRS is a structural description of the display, the foregoing data indicate greater priming when the priming task encouraged the encoding of the structure than when it discouraged such structural encoding. Furthermore, for the structurally impossible displays, for which there could be no integrated structural description under any circumstances, there also was no evidence of priming, even under conditions that did facilitate the explicit recognition of those same impossible objects.

In addition, Scarborough, Cortese, & Scarborough (1977) demonstrated a similar priming effect using words as the stimuli, and again, their data, as well as those of Schacter, Rapscak, Rubens, Tharan, & Laguna, (1990), indicate that when the orthographic display cannot be assigned a structural description (e.g., a consonant array) there also is no priming effect. Furthermore, the structural description is very abstract, because a priming effect can be obtained even when the prime and target displays are in different cases (Scarborough, et al.). In fact, Clarke and Morton (1983) obtained a priming effect when the prime was crudely handwritten in a cursive form and the later target was in a standard print.

Finally, these studies employing words are not simply illustrations of lexical priming, because if the priming stimulus is spoken (Clarke & Morton, 1983), or it is a picture of the object named by the word (Scarborough, Gerard, & Cortese, 1979; Winnick & Daniel, 1970), there are no priming effects. In addition, while semantic priming effects are very short lived (e.g., a second or less), these priming effects within the PRS can last over several days (Scarborough et al., 1977).

In general, then, this precognitive visual representation appears to be a structural description that is abstract in the sense of not preserving the graphic characteristics of the display, and it is prelexical in that it is not affected by a nonvisual lexical prime. That empirical description fits the theoretical definition of the abstract orthographic pattern very closely.

Initial semantic encoding

As the orthographic encoding begins, the model assumes that there is also an implicit activation of the word's semantic category. This activation is based on a preattentional encoding of the visual information, and it is assumed to predate any orthographically based lexical access. However, once this broad-domain semantic activation occurs, it is assumed that: 1) It provides activation to all lexical entries with which it is consistent; 2) any subsequent orthographic encoding that is consistent with those lexical entries adds to their level of activation; and 3) the activation of lexical entries not supported by the orthographic encoding quickly returns to zero.

The entire point of this assumption is to provide an account for the types of unconscious perception effects demonstrated by Marcel (1983a, 1983b). In addition, however, within the model, it provides a mechanism for explaining why the lexical entry for the target word is so much more immediately available than are the entries for orthographically similar words. That is, under conditions in which it seems improbable that the cohort would be resolved, perceivers seem to have little trouble in identifying the target, and the semantically-based extra activation for the target word seems a reasonable explanation for its enhanced availability.

The nature of lexical access

As noted earlier, the moment the first wickelgraph is encoded into the abstract orthographic pattern it is assumed to provide activation to all lexical entries with which it is consistent. This activation is added to any already existing activation stemming from the initial semantic encoding, and the activation for any semantically-activated lexical entry that is not consistent with the first-encoded wickelgraph quickly returns to zero. The set of activated lexical entries at this point is referred to as the initial cohort, even though there was a prior cohort based on the initial semantic encoding.

As subsequently encoded wickelgraphs are incorporated into the abstract orthographic pattern, there is no additional activation provided to cohort members with which they are consistent, but the pattern as a whole becomes inconsistent with an increasing number of the members of the initial cohort set. When that inconsistency with a lexical entry appears, activation is withdrawn from the entry, and its activation level quickly returns to zero. The cohort is resolved when activation has been withdrawn from all the nontarget members of the cohort.

Activation and word frequency effects

At any point in time the level of activation for a lexical entry within an active cohort is a function of the frequency with which it has been paired with the then-current status of the abstract orthographic pattern. As the pattern evolves the frequency with which each of its stages would have been previously paired with consistent lexical entries would not change, and therefore, their level of activation would not change. However, the various lexical entries within the cohort would differ in terms of the frequency with which they had been paired with the pattern, and those differential frequencies, indexed by the words' cultural frequencies, would result in a variation in the level of activation for the various members of the cohort, (e.g., high-frequency words should have a higher level of activation than low-frequency words).

In addition, because high-frequency words would tend to have high-frequency wickelgraphs, their wickelgraphs would have a greater likelihood of being encoded quickly, and in a direct manner, than would those for low-frequency words. That fast orthographic encoding would mean that not only would the initial cohort be established quite quickly, but the cohort also would be resolved more rapidly than for low-frequency items. In general, then, high-frequency words should be processed more rapidly, and they should enjoy a higher-level of activation within the cohort, than low-frequency words.

The use and resolution of the cohort

It is also assumed that the cohort information is immediately available to the perceiver all during the resolution process, and they are free to respond at any time during resolution by selecting a lexical entry, or by making a lexical decision, if the then-current status of the cohort provides them with the needed information. For example, if the displays were always words, and the perceiver's task was to simply name the word that appeared, the most active member of the initial cohort would almost always be the lexical entry for the target. That would be true, because that entry would enjoy activation from both the initial semantic encoding and the orthographic encoding, and the perceiver could then select that entry without resolving the cohort. On the other hand, if perceivers had to make a lexical decision, and the nonwords were all pronounceable and had cohorts that were the same size as the words, it is unlikely that they would have the needed information until the cohort had been completely resolved. Clearly, then, both the nature of the task, and the types of displays, would determine when perceivers would sample the information from the cohort for making their selection or decision.

Since the initial work on the role of a target word's cohort on its visual recognition (Coltheart, Davelaar, Jonasson, & Besner, 1977), the measure than has been used as an index of a word's cohort size is the number of other words that can be formed by changing a single letter in the target item. For example, some of the cohorts for the word MAN would be TAN, FAN, MEN, MAT, etc., and for the nonword FON some cohorts would be SON, TON, FIN, FOG, FOX, etc.

In general, the expectation from the model is that when cohort resolution is required, the time needed to respond to the target item should be an increasing function of the size of the target word's cohort. However, it is also the case, almost by definition, that items (both words and nonwords) with large cohorts should have high-frequency letter patterns, and that would mean that they have more high-frequency wickelgraphs than do small-cohort items. Therefore, although it should take longer to resolve the cohort for large-cohort items, the initial cohort for those items should be established much more quickly. In addition, for the same reason, the rate at which cohort resolution occurs should be faster for large-cohort than for small-cohort items, even though the total amount of time needed for their resolution might be greater (i.e., despite the faster rate of eliminating nontarget cohort members, large cohorts have more such items that need to eliminated).

Finally, for the nonwords in a lexical decision task, the model assumes that the initial cohort would be resolved to the point where there are no remaining word candidates. However, even at that point, subjects cannot be certain that the display is a nonword, because there will be occasions on which the perceiver does not know the correct spelling for a word for which they can identify the spoken form (i.e., there is a lexical entry). That would be particularly true if the displays were low-frequency words, and the model assumes that under those circumstances perceivers will use grapheme-to-phoneme conversion rules to provide a phonological encoding for the display. The first phonetic element converted would activate a phonologically-based cohort, which subsequently would be resolved as the rest of the phonetic elements were encoded. The interesting prediction is that not only would responding to nonwords be delayed, but the fact that two cohorts would need to be resolved implies a larger effect of cohort size on nonwords than words.

Documentation of the resolution process

A lexical decision task seems to be the most appropriate experimental paradigm within which to examine the resolution process as defined in the model (Johnson, 1992). The reason for that, as noted above, is that depending upon the nature of the nonwords, perceivers can be allowed to respond on the basis of the initial cohort, if it provides the needed information, or they can be forced to delay responding until cohort resolution is complete, if the needed information is not available until that point.

Cohort effects when resolution is required. For example, Johnson and Pugh (1992) employed a lexical-decision task and presented the displays in blocks of trials within which the words and nonwords were homogeneous with respect to both length and cohort size. Across conditions, word frequency was controlled at a relatively low level of about 50 or less (Kučera & Francis, 1967). Given that all the nonwords had cohorts whose characteristics were indistinguishable from those for the words, perceivers would have no basis for their word-nonword decision until the cohort was completely resolved. The results indicated that, as expected, decision times were longer for large-cohort items than for small-cohort items, and although the effect was reliable for the words, it was significantly larger for the nonwords. That latter effect is consistent with the idea that nonwords require the resolution of two cohorts.

In addition, however, the pattern of errors in this experiment indicated that subjects seemed to use cohort size as the basis for their decision when, for whatever reason, they had an opportunity to make an error. Specifically, increasing cohort size reduced errors for the words, but it increased errors for the nonwords, and that data pattern suggests that in this task there is a bias to respond "word" when there is a large cohort and "nonword" when there is a small cohort. In terms of the latency data, that bias should tend to shrink the difference between large and small cohort words, but increase the difference for nonwords, and across a large number of experiments there has been a near perfect correlation between the error-based estimate of the bias and the magnitude of its expected effect on latency.

In addition, this bias effect also provides a second explanation for why the delaying effect of cohort size would be greater for nonwords. That is, overcoming the bias to respond YES or WORD to a large-cohort nonword would delay those responses, while responding NO or NONWORD to small-cohort nonword displays would be facilitated, and that would magnify the effect of cohort size. On the other hand, the bias would facilitate responding to large-cohort words, partially overcoming the delaying effect of the large cohort, but the bias would delay responding to small-cohort words, and that combined effect would shrink the overall influence of cohort-size on words. The magnifying effect of the bias on the nonwords, and its shrinking effect for the words, also would yield the response-type (YES versus NO) by cohort-size interaction that was obtained consistently across these experiments.

When delaying effects of cohort size do not appear. Although the foregoing study did demonstrate a delaying effect of a large cohort on the lexical decision times for both words and nonwords, the earlier Coltheart et al. (1977) study found that effect only for nonwords, and Andrews (1989) and Pugh, Rexer, and Katz (1992) actually found facilitation for large-cohort words. The major difference between these studies and Johnson and Pugh (1992), is that in the latter study the display types were organized into blocks of trials within which the displays were homogeneous with regard to cohort size, whereas the other studies all had them intermixed.

Johnson and Pugh (1992) then did a study that was a replication of their earlier experiments, with the single exception that items with large and small cohorts were intermixed within blocks of trials, and under those conditions they obtained the Coltheart et al. (1977) result of no effect of cohort size on words. However, the error-based index of response bias almost doubled in that experiment in comparison to their earlier experiments, and the index of that bias on the latencies increased accordingly. In that the response bias tends to shrink or reverse the delaying effect of increasing cohort size on words, but inflates it for nonwords, the increase in bias can account for the differential results between these experiments. It would appear that intermixing cohort sizes within trial blocks made cohort size a more salient dimension.

Pugh, Rexer, and Katz (1992) offer further support for this response-bias explanation of the Coltheart et al. (1977) cohort effect. They also used trial blocks in which cohort size was intermixed, but the only displays were words, and the subjects' task was to press a response button when they felt they knew the meaning of the displayed word. Clearly, given that the meaning of the target would be the only issue, the size of the initial cohort and the attendant response bias would be irrelevant, and the needed information would be available only after cohort resolution was complete. Their data indicate that with the facilitating effect of response bias eliminated, the cohort effects were in the form of interference, even with cohort sizes intermixed within trial blocks.

Cohort size versus number of position yielding cohorts. Finally, Johnson and Pugh (1992) noted that if, during resolution, all cohort members inconsistent with a newly encoded wickelgraph are eliminated from the cohort, then the critical issue should not be the size of an item's cohort, but rather the number of letter positions that yield at least one cohort. To explore this issue they had two pairs of conditions. In one pair the two conditions were equated for the number of letter positions that yielded cohorts, but the items in the two conditions differed in terms of the mean number of cohorts. The items in the other pair of conditions were equated in terms of mean number of cohorts, but they differed in the number of letter positions that yielded those cohorts.

The data were quite clear. When the number of letter positions yielding cohorts was controlled, there was no overall effect of cohort size on the decision times, but when number of cohorts was controlled, increasing the number of letter positions that yield cohorts resulted in longer lexical decision latencies (also see Pugh et al., 1992). In addition, although there was no evidence in the error data of any response bias attributable to number of positions yielding cohorts, there was a reliable (although small) bias effect attributable to cohort size, and that resulted in a facilitation effect from increased cohort size in the latency data for words.

From these data, it appears that the characteristic of the cohort that is a delaying factor during cohort resolution is the number of wickelgraphs that have to be encoded in order to achieve resolution, and not the total number of cohorts that need to be eliminated. This conclusion is further supported by an experiment described by Pugh (personal communication) in which one letter from within a low-frequency word is slightly delayed relative to the other letters when the word display appears on the computer screen. If the delayed letter yields cohorts, then lexical decision latencies are longer than if the delayed letter yields no cohorts.

These data would suggest, then, that within a lexical-decision task, with the number of letter positions yielding cohorts held constant, the effect of increasing the cohort size of words is to facilitate responding, rather than delay responding. The facilitation is assumed to occur both because the increased familiarity of the wickelgraphs should increase the rate of orthographic encoding, and because a large cohort seems to induce a bias to respond YES or WORD. Given that is the case, then if there is no other characteristic of the large-cohort words that would interfere with their processing, perceivers should respond faster to those words than to small-cohort words. The issue for nonwords is a little more complex, because while there would be the same facilitating effect of cohort size on orthographic encoding, the response bias would be an interference effect for nonwords.

Lexical decisions based on characteristics of the initial cohort

One critical assumption within the model is that perceivers can use the status of the cohort at any point in time as a basis for their decision, with the availability of the needed information being the only determiner of when that occurs. In the experiments just described, the task and materials dictated that the needed information would not be available until cohort resolution was complete. However, Johnson and Pugh (1992) also conducted experiments in which that was not true. In that situation the word displays appeared in trial blocks that were homogeneous with respect to cohort size, just as in the other experiments, but the nonwords for all the conditions were unpronounceable and they had very few cohorts.

For the trial blocks that had large-cohort words, the size of the initial cohort alone could be used for differentiating words and nonwords, but that would not be true for the small-cohort words. The fact that cohort resolution would be needed for the small-cohort items, but not for those with large-cohorts, suggests there should be a reversal of the original effect of cohort size. That is, increasing cohort size, as well as the number of positions yielding cohort, should decrease latencies, and that was the result they obtained.

Similarly, studies that have employed a naming task for words, consistently show a facilitating effect of increasing cohort size when number of positions is not controlled (Andrews, 1989; Laxon, Coltheart, & Keating, 1988), even though the same words yield an interference effect in a lexical-decision task (Johnson & Pugh, 1992). The argument from the model, however, is that the combined activation stemming from the initial semantic encoding and that coming from the initial orthographic encoding should make the total activation of the lexical entry for the target word much greater than the activation for any other member of the cohort (i.e., those items would only have orthographically-based activation). Furthermore, that would generally be true even if one of the cohorts had a greater cultural frequency, and thereby had a somewhat greater orthographically-based activation within the cohort. That being the case, for a naming task, subjects can simply name the most active cohort member and be confident that they will be correct.

On the other hand, even nonwords have cohorts, and, within those cohorts, one cohort member would have a higher level of activation than the others. To that extent the initial cohorts for nonwords would look like those for words, with the exception that it would usually be the case that the difference in activation between the most active and the next most active item would be greater for the words (i.e., for words the target would have the extra semantic activation). However, for words there would be occasions when one or more nontarget cohort members would have a higher cultural frequency than the target. In those cases the difference in activation level between that of the target and that of the most active nontarget might not be any greater than the difference between the two most active cohort members for a nonword. That would be particularly true if the words had relatively low frequencies, which generally has been the case in these experiments. For that reason, then, if nonwords were included in the series of displays, initial cohort characteristics could not be used as a basis for distinguishing words from nonwords. If the perceiver's task was to name only the items for which there were lexical entries (i.e., the words), it would be necessary for them to resolve the cohort before responding.

In a recent experiment we explored this issue using the large-cohort and small-cohort words from the Johnson and Pugh (1992) experiments. As before, the words were blocked by cohort size, and in one condition subjects were to name each word as it appeared on the computer screen. In the other condition each block of trials had an equal number of pronounceable nonwords of the same cohort size as the words, but again they were to pronounce only the words (i.e., the task required an initial lexical decision).

The expectation was that with no need to resolve a cohort, the only effect on naming latencies should be the facilitating influence of increasing cohort size stemming from the increased orthographic familiarity. On the other hand, when an initial lexical decision must be made, then cohort resolution would be needed, and increasing cohort size should delay the naming response rather than facilitate it.

The results fit the expectation very closely. When the task involved only naming, the latencies were shorter for the large-cohort than the small-cohort words, but the reverse was true when a lexical decision preceded the naming response. These data, along with those obtained using unpronounceable nonwords, suggest that subjects do let the nature of both the task and the materials determine how and when they sample and use the cohort information in the decision process.

The role of word frequency in visual word recognition

One of the primary motivating forces for search models of word recognition is that they provide a very simple way of accounting for word frequency effects, whereas it is somewhat more cumbersome to account for those effects within the context of activation models. Even some models that employ the concept neighborhood or cohort (e.g., Becker, 1976; Paap, Newsome, McDonald, & Schvaneveldt, 1982) also include a search component to handle frequency effects. This model (Johnson, 1992), however, does not include a search mechanism, but instead it has two mechanisms that in combination can account for the traditional frequency effects in visual word recognition.

Specifically, high-frequency words would tend to have high-frequency wickelgraphs, which would mean that they should have a higher rate of orthographic encoding than low-frequency words. Under the constraints of a degraded (tachistoscopic) viewing condition that would mean that they could be encoded more fully than low-frequency words, and thereby yield higher accuracy levels. In addition, however, the more familiar wickelgraphs also would mean that the initial cohort would be established and resolved more quickly. Furthermore, the lexical entry for a high-frequency target would tend to stand out in the initial cohort, as a result of both the activation from the orthographic encoding and the initial semantic encoding, and the fact that one item stands out in that way would reduce the likelihood of there being a need to resolve the cohort. All of these effects would tend to shorten the processing time needed for high-frequency words, and that would result in reduced latencies for nondegraded displays. In general, then, without including any special mechanisms, the model does seem able to handle the usual accuracy and latency effects attributable to word frequency.

In addition, however, there are some special influences of word frequency on cohort effects in word recognition. One of the effects is the fact that the influence of increasing cohort size is reduced as word frequency increases, and that seems to occur regardless of whether the basic effect of increasing cohort size is to facilitate processing and the response decision (Andrews, 1989; Laxon, Coltheart, & Keating, 1988), or the cohort-size effect is obtained under conditions that yields interference which delays responding (Fox & Koenigsknecht, 1990; Johnson & Pugh, 1992).

With regard to the interference effect, it is assumed that the delay in the decision associated with increasing cohort size stems from the increasing time needed for resolution, but if increasing word frequency reduces the need for cohort resolution, then the cohort-size effect should be attenuated, which the data indicates to be the case (Johnson & Pugh, 1992). When increasing cohort size facilitates responding it is assumed that no cohort resolution occurred, and that the facilitation arises from the enhanced rate of orthographic encoding. That is, on average, large-cohort items will have wickelgraphs that are more familiar than will small-cohort items. However, as word frequency increases that also will increase the familiarity of the wickelgraphs, but the enhancement in familiarity should be greater for the small-cohort items, thereby reducing the overall difference in wickelgraphic familiarity between the two cohort sizes. In general, then, the model is able to handle the reduced effect of cohort size as word frequency increases, both for conditions under which large cohorts delay responding and for conditions under which they facilitate responding.

Finally, again in the context of the effect of word frequency on cohort effects, as the foregoing discussion would suggest, there are marked effects of whether there are members of the cohort whose word frequency exceeds that of the target (Grainger, 1990; Grainger, O'Regan, Jacobs, & Segui, 1989; Stadtlander & Krueger, 1991). In general, as the cultural frequency of the critical nontarget cohort member increases, there is an increase in the decision latency. Within the model (Johnson, 1992), for word displays, if the cultural frequencies of all nontarget members of the cohorts were lower than those of the targets, then the level of activation of the target, in comparison to the next most active cohort, would be so great that it could be readily identified within the initial cohort, (no resolution would be needed). Not only would that be true for a naming task, but for a lexical-decision task the difference in activation between the two most active cohort members would be much greater for words than nonwords, and that also could be used as a basis for decision.

However, if the cultural frequency of a nontarget cohort member, and therefore its orthographically-based activation, exceeded that of the target member, the difference in activation between the target and the next most active cohort member would be less discriminable. That would reduce the likelihood that such information could be used in either a lexical-decision or a naming task, and the increased response latency would come from the need for cohort resolution (albeit, the effect should be larger for a lexical-decision task).

Summary

The primary goal of this model (Johnson, 1992) was to account for the obvious fact that letters are used in the process of reading words, while at the same time that very letter information has no immediate availability to the reader. In addition, a similar comparison that needs to be handled is the fact that word recognition is influenced by the number of other words that share letters with the target word, while on the other hand, there is an even larger effect of the cultural frequency of the word as a whole. In general, then, the critical issue has been to account for the fact that although for word displays, word-level characteristics have an almost overwhelming impact on recognition processes, it is also the case that attributes of letter-level encodings are absolutely critical.

The cohort model of visual word recognition (Johnson, 1992; Johnson & Pugh, 1992) draws a basic distinction between the encoding of orthographic information and the use of the product of that encoding process for lexical and semantic access. It is assumed that the critical event during orthographic processing is the encoding of letter identities in terms of their immediate orthographic context (i.e., their two adjacent letters). These context-sensitive encodings are termed wickelgraphs, and the possible set of such wickelgraphs instantiates the rules of orthography. In addition, the context information within a wickelgraph is assumed to dictate the manner in which it can be integrated into the unitary and abstract orthographic pattern that is used for lexical access.

Lexical access is assumed to begin when the first wickelgraph is encoded into the abstract orthographic pattern. However, it is also assumed that prior to that orthographically-based access, there is a preliminary semantically-based activation of all lexical entries consistent with the semantic category of the target item. Any semantically activated items not consistent with the first-encoded wickelgraph loose their activation at that time, while the orthographically-based activation is added to all lexical entries with which it is consistent, including those that already have some activation stemming from the initial semantic encoding. The set of activated lexical entries at that point is termed the initial cohort, and, in general, within that cohort, the most active item would be the target, because it would have received activation from both the initial semantic and the initial orthographic encoding.

The abstract orthographic pattern subsequently evolves as increasing numbers of wickelgraphs are encoded, and activation is withdrawn from lexical entries as they become inconsistent with the evolving pattern. The cohort of lexical entries is resolved when all but the target has been deactivated, but perceivers do not have to wait until that point before they respond. That is, they can make their decision or lexical selection at any point in time, given the characteristics of the cohort provides them with the needed information.

With regard to the orthographic component of the model, the data on the role of context information in orthographic encoding is strikingly consistent with the concept of a wickelgraph, and the data supporting the concept of an abstract orthographic pattern is equally clear. Similarly, the data on the influence of the target item's cohort also fits the model very well. In addition, the model is able to provide an account for both the conditions under which a large cohort facilitates lexical selections and decisions, and the conditions under which a large cohort results in interference. Finally, it also can handle the relatively complex frequency effects that have been apparent in these tasks. Overall, then, the model does seem to provide a reasonable account of lexical access, and there is independent evidence supporting the individual constructs within the model.

References

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

Becker, C. A. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 556-566.

Brown, G. D. A. (1987). Resolving inconsistency: A computational model of word naming. *Journal of Memory and Language, 26,* 1-23.

Carnot, M. J. (1988). *The encoding of positional information in visual displays.* Unpublished masters thesis, The Ohio State University, Columbus, OH.

Clarke, R., & Morton, J. (1983). Cross modality facilitation in tachistoscopic word recognition. *Quarterly Journal of Experimental Psychology, 35A,* 79-96.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535-556). New York: Academic Press.

Compton, P., Grossenbacher, P., Posner, M. I., & Tucker, D. (1991). A cognitive-anatomical approach to attention in lexical access. *Journal of Cognitive Neuroscience, 3,* 304-312.

Drewnowski, A., & Healy, A. F. (1977). Detection errors on *the* and *and*: Evidence for reading units larger than the word. *Memory & Cognition, 5,* 636-647.

Fox, R. A., & Koenigsknecht, R. (1990). *The effect of cohort size on gated word recognition.* Paper presented at the meeting of the American Speech-Language-Hearing Association, St. Louis, MO.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language, 29,* 228-244.

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.
Healy, A. F. & Drewnowski, A. (1983). Investigating the boundaries of reading units: Letter detection in misspelled words. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 413-426.
Huey, E. B. (1908/1968). *The psychology and pedagogy of reading.* Cambridge, MA: MIT Press.
Johnson, N. F. (1975). On the function of letters in word identification: Some data and a preliminary model. *Journal of Verbal Learning and Verbal Behavior, 14,* 17-29.
Johnson, N. F. (1977). A pattern-unit model of word identification. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension* (pp. 91-126). Hillsdale, NJ: Erlbaum.
Johnson, N. F. (1981). Integration processes in word recognition. In O. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology* (pp. 29-64). Hillsdale, NJ: Erlbaum.
Johnson, N. F. (1986a). On looking at letters within words: Do we "see" them in memory? *Journal of Memory and Language, 25,* 558-570.
Johnson, N. F. (1991). Holistic models of word recognition. In R. Hoffman & D. Palermo (Eds.), *Cognitive psychology: The state of the art* (pp. 79-94). Hillsdale NJ: Lawrence Erlbaum Associates.
Johnson, N. (1992). *A cohort model of visual word recognition.* Manuscript submitted for publication.
Johnson, N. F., & Blum, A. J. (1988). When redundancy hurts letter detection: An attempt to define one condition. *Perception & Psychophysics, 43,* 147-155.
Johnson, N. F., & Pugh, K. R. (1992). *The nature of cohort effects in visual word recognition: Some preliminary experiments.* Manuscript submitted for publication.
Johnson, N. F., Pugh, K. R., & Blum, A. J. (1989). More on the way we "see" letters from words within memory. *Journal of Memory and Language, 28,* 155-163.
Johnson, N. F., Turner-Lyga, M., & Pettegrew, B. S. (1986). Part-whole relationships in the processing of small visual patterns. *Memory and Cognition, 14,* 5-16.
Johnston, J. C., & McClelland, J. L. (1980). Experimental tests of a model of word identification. *Journal of Verbal Learning and Verbal Behavior, 19,* 503-524.
Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.
Laxon, V. J., Coltheart, V., & Keating, C. (1988). Children find friendly words friendly too: Words with many orthographic neighbours are easier to read and spell. *British Journal of Educational Psychology, 58,* 103-119.
Marcel, A. J. (1983a). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15,* 197-237.
Marcel, A. J. (1983b). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology, 15,* 238-300.
Marmurek, H. H. C. (1987). Attentional holism in visual word processing. *Psychological Research, 49,* 45-52.
Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition, 25,* 71-102.
Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10,* 29-63.
McClelland, J. L., & Mozer, M. C. (1986). Perceptual interactions in two-word displays: Familiarity and similarity effects. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 18-35.
Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: The word superiority effect. *Psychological Review, 89,* 573-594.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature, 331,* 585-589.

Petersen, S. E., Fox, S. E., Snyder, A. Z., & Raichle, M. E. (1990). Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science, 249,* 1041-1049.

Posner, M. I., Petersen, S. E., Fox, P. T., & Raichle, M. E. (1988). Localization of cognitive operations in the human brain. *Science, 240,* 1627-1631.

Pugh, K. R., Rexer, K., & Katz, L. (1992). *Effects of orthographic neighborhood distribution in visual word recognition.* Manuscript submitted for publication.

Scarborough, D. L., Cortese, C., & Scarborough, H. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 3-12.

Scarborough, D. L., Gerard, L., & Cortese, C. (1979). Accessing lexical memory: The transfer of word repetition effects across task and modality. *Memory & Cognition, 7,* 3-12.

Schacter, D. L. (1990). Perceptual representation systems and implicit memory: Toward a resolution of the multiple memory systems debate. In A. Diamond (Ed.), *Development and neural bases of higher cognition. Annals of the New York Academy of Sciences, 608* (pp. 543-571).

Schacter, D. L., Cooper, L. A., & Delaney, S. M. (1990). Implicit memory for unfamiliar objects depends on access to structural descriptions. *Journal of Experimental Psychology: General, 119,* 5-24.

Schacter, D. L., Cooper, L. A., Delaney, S. M., Peterson, M. A., & Tharan, M. (1991). Implicit memory for possible and impossible objects: Constraints on the construction of structural descriptions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 3-19.

Schacter, D. L., Rapcsak, S. Z., Rubens, A. B., Tharan, M., & Laguna, J. M. (1990). Priming effects in a letter-by-letter reader depend upon access to the word form system. *Neuropsychologia, 28,* 1079-1094.

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

Sloboda, J. A. (1976). Decision times for word and letter search: A wholistic word identification model examined. *Journal of Verbal Learning and Verbal Behavior, 15,* 93-101.

Sloboda, J. A. (1977). The locus of the word-priority effect in a letter detection task. *Memory & Cognition, 5,* 371-376.

Stadtlander, L. M., & Krueger, L. E. (1991). *The effect of word frequency and size of neighborhood: Evidence for a frequency-verification and heuristic model of visual word recognition.* Manuscript submitted for publication.

Travers, J. R. (1973). The effects of forced serial processing on identification of words and random letter strings. *Cognitive Psychology, 5,* 109-137.

Travers, J. R. (1974). Word recognition with forced serial processing: Effects of segment size and temporal order variation. *Perception & Psychophysics, 16,* 35-42.

Treisman, A., & Souther, J. (1986). Illusory words: The role of attention and top-down constraints in conjoining letters to form words. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 3-17.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247,* 301-306.

Wickelgren, W. A. (1969). Context sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review, 76,* 1-15.

Winnick, W. A., & Daniel, S. A. (1970). Two kinds of response priming in tachistoscopic recognition. *Journal of Experimental Psychology, 84,* 74-81.

PART 2

Orthography and Phonology

This Page Intentionally Left Blank

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 9

The Relation of Speech to Reading and Writing

Alvin M. Liberman
Haskins Laboratories

Theories of reading/writing and theories of speech typically have in common that neither takes proper account of an obvious fact about language that must, in any reckoning, be critically relevant to both: there is a vast difference in naturalness (hence ease of use) between its spoken and written forms. In my view, a theory of reading should begin with this fact, but only after a theory of speech has explained it.

My aim, then, is to say how well the difference in naturalness is illuminated by each of two theories of speech—one conventional, the other less so—and then, in that light, to weigh the contribution that each of these can make to an understanding of reading and writing and the difficulties that attend them. More broadly, I aim to promote the notion that a theory of speech and a theory of reading/writing are inseparable, and that the validity of the one is measured, in no small part, by its fit to the other.

What does it mean to say that speech is more natural?

The difference in naturalness between the spoken and written forms of language is patent, so I run the risk of being tedious if I elaborate it here. Still, it is important for the argument I mean to make that we have explicitly in mind how variously the difference manifests itself. Let me, therefore, count the ways.

(1) Speech is universal. Every community of human beings has a fully developed spoken language. Reading and writing, on the other hand, are relatively rare. Many, perhaps most, languages do not even have a written form, and when, as in modern times, a writing system is devised—usually by missionaries—it does not readily come into common use.

(2) Speech is older in the history of our species. Indeed, it is presumably as old as we are, having emerged with us as perhaps the most important of our species-typical characteristics. Writing systems, on the other hand, are developments of the last few thousand years.

(3) Speech is earlier in the history of the individual; reading/writing come later, if at all.

(4) Speech must, of course, be learned, but it need not be taught. For learning to speak, the necessary and sufficient conditions are but two: membership in the human race and exposure to a mother tongue. Indeed, given that these two conditions are met, there is scarcely any way that the development of speech can be prevented. Thus, learning to speak is a precognitive process, much like learning to perceive visual depth and distance or the location of sound. In contrast, reading and writing require to be taught, though, given the right ability, motivation, and opportunity, some will infer the relation of script

to language and thus teach themselves. But, however learned, reading/writing is an intellectual achievement in a way that learning to speak is not.

(5) There are brain mechanisms that evolved with language and that are, accordingly, largely dedicated to its processes. Reading and writing presumably engage at least some of these mechanisms, but they must also exploit others that evolved to serve nonlinguistic functions. There is no specialization for reading/writing as such.

(6) Spoken language has the critically important property of 'openness': unlike nonhuman systems of communication, speech is capable of expressing and conveying an indefinitely numerous variety of messages. A script can share this property, but only to the extent that it somehow transcribes its spoken-language base. Having no independent existence, a proper (open) script is narrowly constrained by the nature of its spoken-language roots and by the mental resources on which they draw. Still, within these constraints, scripts are more variable than speech.

One dimension of variation is the level at which the message is represented, though the range of that variation is, in fact, much narrower than the variety of possible written forms would suggest. Thus, as DeFrancis (1989) convincingly argues, any script that communicates meanings or ideas directly, as in ideograms, for example, is doomed to arrive at a dead end. Ideographic scripts cannot be open—that is, they cannot generate novel messages—and the number of messages they can convey is never more than the inventory of one-to-one associations between (holistically different) signals and distinctly different meanings that human beings can master. Indeed, it is a distinguishing characteristic of language, and a necessary condition of its openness, that it communicates meanings indirectly, via specifically linguistic structures and processes, including, nontrivially, those of the phonological component. Not surprisingly, scripts must follow suit; in the matter of language, as with so many other natural processes, it is hard to improve on nature.

Constraints of a different kind apply at the lower levels. Thus, the acoustic signal, as represented visually by a spectrogram, for example, cannot serve as a basis for a script; while spectrograms can be puzzled out by experts, they, along with other visual representations, cannot be read fluently. The reason is not primarily that the relevant parts of the signal are insufficiently visible; it is, rather, that, owing to the nature of speech, and especially to the coarticulation that is central to it, the relation between acoustic signal and message is complex in ways that defeat whatever cognitive processes the 'reader' brings to bear. Narrow phonetic transcriptions are easier to read, but there is still more context-, rate-, and speaker-conditioned variation than the eye is comfortable with. In any case, no extant script offers language at a narrow phonetic level. To be usable, scripts must, apparently, be pitched at the more abstract phonological and morphophonological levels. That being so, and given that reading-writing require conscious awareness of the units represented by the script, we can infer that people can become conscious of phonemes and morphophonemes. We can also infer about these units that, standing above so much of the acoustic and phonetic variability, they correspond approximately to the invariant forms in which words are presumably stored in the speaker's lexicon. A script that captures this invariance is surely off to a good start. At all events, some scripts (e.g., Finnish, Serbo-Croatian) do approximate to purely phonological renditions of the language, while others depart from a phonological base in the direction of morphology. Thus, English script is rather highly morphophonological, Chinese even more so. But, as DeFrancis (1989; see also Wang, 1981) makes abundantly clear, all these scripts, including even the Chinese, are significantly phonological, and, in his view, they would fail if they were not; the variation is simply in the degree to which some of the morphology is also represented.

Scripts also vary somewhat, as speech does not, in the size of the linguistic segments they take as their elements, but here, too, the choice is quite constrained. Surely, it would not do to make a unit of the script equal to a phoneme and a half, a third of a syllable, or some arbitrary stretch—say 100 milliseconds—of the speech stream. Still, scripts can and do take as their irreducible units either phonemes or syllables, so in this respect, too, they are more diverse than speech.

(7) All of the foregoing differences are, of course, merely reflections of one underlying circumstance—namely, that speech is a product of biological evolution, while writing systems are artifacts. Indeed, an alphabet—the writing system that is of most immediate concern to us—is a triumph of applied biology, part discovery, part invention. The discovery—surely one of the most momemtous of all time—was that words do not differ from each other holistically, but rather by the particular arrangement of a small inventory of the meaningless units they comprise. The invention was simply the notion that if each of these units were to be represented by a distinctive optical shape, then everyone could read and write, provided he knew the language and was conscious of the internal phonological structure of its words.

How is the difference in naturalness to be understood?

Having seen in how far speech is more natural than reading/writing, we should look first for a simple explanation, one that is to be seen in the surface appearance of the two processes. But when we search there, we are led to conclude, in defiance of the most obvious facts, that the advantage must lie with reading/writing, not with speech. Thus, it is the eye, not the ear, that is the better receptor; the hand, not the tongue, that is the more versatile effector; the print, not the sound, that offers the better signal-to-noise ratio; and the discrete alphabetic characters, not the nearly continuous and elaborately context-conditioned acoustic signal, that offers the more straightforward relation to the language. To resolve this seeming paradox and understand the issue more clearly, we shall have to look more deeply into the biology of speech. To that end, I turn to two views of speech to see what each has to offer.

The conventional view of speech as a basis for understanding the difference in naturalness. The first assumption of the conventional view is so much taken for granted that it is rarely made explicit. It is, very simply, that the phonetic elements are defined as sounds. This is not merely to say the obvious, which is that speech is conveyed by an acoustic medium, but rather to suppose, in a phrase made famous by Marshall McLuhan, that the medium *is* the message.

The second assumption, which concerns the production of these sounds, is also usually unspoken, not just because it is taken for granted, though it surely is, but also because it is apparently not thought by conventional theorists to be even relevant. But, whatever the reason, one finds among the conventional claims none which implies the existence of a phonetic mode of action—that is, a mode adapted to phonetic purposes and no other. One therefore infers that the conventional view must hold (by default, as it were) that no such mode exists. Put affirmatively, the conventional assumption is that speech is produced by motor processes and movements that are independent of language.

The third assumption concerns the perception of speech sounds, and, unlike the first two, is made explicitly and at great length (Cole & Scott, 1974; Crowder & Morton, 1969;

Diehl & Kluender, 1989; Fujisaki & Kawashima, 1970; Kuhl, 1981; Miller, 1977; Oden & Massaro, 1978; Stevens, 1975). In its simplest form, it is that perception of speech is not different from perception of other sounds; all are governed by the same general processes of the auditory system. Thus, language simply accepts representations made available to it by perceptual processes that are generally auditory, not specifically linguistic. So, just as language presumably recruits ordinary motor processes for its own purposes, so, too, does it recruit the ordinary processes of auditory perception; at the level of perception, as well as action, there is, on the conventional view, no specialization for language.

The fourth assumption is required by the second and third. For if the acts and percepts of speech are not, by their nature, specifically phonetic, they must necessarily be made so, and that can be done only by a process of cognitive translation. Presumably, that is why conventional theorists say about speech perception that after the listener has apprehended the auditory representation he must elevate it to linguistic status by attaching a phonetic label (Crowder & Morton, 1969; Fujisaki & Kawashima, 1970; Pisoni, 1973), fitting it to a phonetic prototype (Massaro, 1987; Oden & Massaro, 1978), or associating it with some other linguistically significant entity, such as a 'distinctive feature' (Stevens, 1975).

I note, parenthetically, that this conventional way of thinking about speech is heir to two related traditions in the psychology of perception. One, which traces its origins to Aristotle's enumeration of the five senses, requires of a perceptual mode that it have an end organ specifically devoted to its interests. Thus, ears yield an auditory mode; eyes, a visual mode; the nose, an olfactory mode; and so on. Lacking an end organ of its very own, speech cannot, therefore, be a mode. In that case, phonetic percepts cannot be the immediate objects of perception; they can only be perceived secondarily, as the result of a cognitive association between a primary auditory representation appropriate to the acoustic stimulus that excites the ear (and hence the auditory mode) and, on the other hand, some cognitive form of a linguistic unit. Such an assumption is, of course, perfectly consistent with another tradition in psychology, one that goes back at least to the beginning of the 18th century, where it is claimed in Berkeley's "New Theory of Vision" (1709) that depth (which cannot be projected directly onto a two-dimensional retina) is perceived by associating sensations of muscular strain (caused by the convergence of the eyes as they fixate objects at various distances) with the experience of distance. In the conventional view of speech, as in Berkeley's assumption about visual depth, apprehending the event or property is a matter of perceiving one thing and calling it something else.

Some of my colleagues and I have long argued that the conventional assumptions fail to account for the important facts about speech. Here, however, my concern is only with the extent to which they enlighten us about the relation of spoken language to its written derivative. That the conventional view enlightens us not at all becomes apparent when one sees that, in contradiction of all the differences I earlier enumerated, it leads to the conclusion that speech and reading/writing must be equally natural. To see how comfortably the conventional view sits with an (erroneous) assumption that speech and reading/writing are psychologically equivalent, one need only reconsider the four assumptions of that view, substituting, where appropriate, 'optical' for 'acoustic' or 'visual' for 'auditory.'

One sees then, that, just as the phonetic elements of speech are, by the first of the conventional assumptions, defined as sounds, the elements of a writing system can only be defined as optical shapes. As for the second assumption—viz., that speech production is managed by motor processes of the most general sort—we must suppose that this is exactly true for writing; by no stretch of the imagination can it be supposed that the writer's movements are the output of an action mode that is specifically linguistic. The third assumption of the conventional view of speech also finds its parallel in reading/writing, for, surely, the percepts evoked by the optical characters are ordinarily visual in the same way that the percepts evoked by the sounds of speech are supposed to be ordinarily auditory. Thus, at the level of action and perception, there is in reading/writing, as there is assumed to be in speech, no specifically linguistic mode. For speech, that is only an assumption—and, as I think, a very wrong one—but for reading/writing it is an incontrovertible fact; the acts and percepts of reading/writing did not evolve as part of the specialization for language, hence they cannot belong to a natural linguistic mode.

The consequence of all this is that the fourth of the conventional assumptions about speech is, in fact, necessary for reading/writing and applies perfectly to it: like the ordinary, nonlinguistic auditory and motor representations according to conventional view of speech, the correspondingly ordinary visual and motor representations of reading/writing must somehow be made relevant to language, and that can only be done by a cognitive process; the reader/writer simply has to learn that certain shapes refer to units of the language and that others do not.

It is this last assumption that most clearly reveals the flaw that makes the conventional view useless as a basis for understanding the most important difference between speech and reading/writing—namely, that the evolution of the one is biological, the other cultural. To appreciate the nature of this shortcoming, we must first consider how either mode of language transmission meets a requirement that is imposed on every communication system, whatever its nature and the course of its development. This requirement, which is commonly ignored in arguments about the nature of speech, is that the parties to the message exchange must be bound by a common understanding about which signals, or which aspects of which signals, have communicative significance; only then can communication succeed. Mattingly and I have called this the requirement for 'parity' (Liberman & Mattingly, 1985; Liberman & Mattingly, 1989; Mattingly & Liberman, 1988). One asks, then, what is entailed by parity as the system develops in the species and as it is realized in the normal communicative act.

In the development of writing systems, the answer is simple and beyond dispute: parity was established by agreement. Thus, all who use an alphabet are parties to a compact that prescribes just which optical shapes are to be taken as symbols for which phonological units, the association of the one with the other having been determined arbitrarily. Indeed, this is what it means to say that writing systems are artifacts, and that the child's learning the linguistic significance of the characters of the script is a cognitive activity.

Unfortunately for the validity of the conventional assumptions, they require that the same story be told about the development of parity in speech. For if the acts and percepts of speech are, as the conventional assumption would have it, ordinarily motor and ordinarily auditory, one must ask how, why, when, and by whom they were invested with lingistic significance. Where is it written that the gesture and percept we know as [b] should count for language, but that a clapping of the hands should not? Is there somewhere a commandment that says, Thou shalt not commit [b] except when it is thy clear intention to communicate? Or are we to assume, just as absurdly, that [b] was incorporated into the language by agreement? It is hard to see how the conventional view of speech can be made to provide a basis for understanding the all-important difference in evolutionary status between speech and reading/writing.

The problem is the worse confounded when we take account of both sides of the normal communicative act. For, on the conventional view the speaker deals in representations of a generally motor sort and the listener in representations of a generally auditory sort. What is it, then, that these two representations have in common, except that neither has anything to do with language? One must thus suppose for speech, as for writing and reading, that there is something like a phonetic idea—a cognitive representation of some kind—to connect these representations to each other and to language, and so to make communication possible.

Thus it is that at every biological or psychological turn the conventional view of speech make reading and writing the equivalents of speech perception and production. Since these processes are plainly not equivalent, the conventional view of speech can hardly be the starting point for an account of reading and writing.

The unconventional view of speech as a basis for understanding the difference in naturalness. The first assumption of the unconventional view is that the units of speech are defined as gestures, not as the sounds that those gestures produce. (For recent accounts of the unconventional view, see: Liberman & Mattingly, 1985; Liberman & Mattingly, 1989; Mattingly & Liberman, 1988; Mattingly & Liberman, 1990). The rationale for this assumption is to be understood by taking account of the function of the phonological component of the grammar and of the requirements it imposes. As for the function of phonology, it is, of course, to form words by combining and permuting a few dozen meaningless segments, and so to make possible a lexicon tens of thousands of times larger than could ever have been achieved if, as in all natural but nonhuman communication systems, each 'word' were conveyed by a signal that was holistically different from all others. But phonology can serve this critically important function only if its elements are commutable; and if they are to be commutable, they must be discrete and invariant.

A related requirement has to do with rate, for if all utterances are to be formed by variously stringing together an exiguous set of signal elements, then, inevitably, the strings must run to great lengths. It is essential, therefore, if these strings are to be organized into words and sentences, that they be produced and perceived at reasonable speed. But if the auditory percepts of the conventional view are to be discrete and invariant, the sounds and gestures must be discrete and invariant, too. Such sounds and gestures are possible, of course, but only at the expense of rate. Thus one could not, on the conventional view, say 'bag,' but only [b] [a] [g], and to say [b] [a] [g] is not to speak but to spell. Of course, if speech were like that, then everyone who could speak or perceive a word would know exactly how to write and read it, provided only that he had managed the trivial task of memorizing the letter-to-sound correspondences. The problem is that there would be no language worth writing or reading.

There seems, indeed, no way to solve the rate problem and still somehow preserve the acoustic-auditory strategy of the conventional view. It would not have helped, for example, if Nature had abandoned the vocal tract and equipped her human creatures with acoustic devices adapted to producing a rapid sequence of sounds—a drumfire or tattoo—for that strategy would have defeated the ear. The point is that speech proceeds at rates that transmit up to 15 or even 20 phonemes per second, but if each phoneme were represented by a discrete sound, then rates that high would seriously strain and sometimes overreach the ability of the ear to resolve the individual sounds and to divine their order.

According to the unconventional view, Nature solved the problem by avoiding the acoustic-auditory strategy that would have created it. The alternative she chose was to define the phonetic elements as gestures, as the first assumption of the unconventional view proposes. Thus, [b] is a closing at the lips, [h] an opening at the glottis, [p] a combination of lip closing and glottis opening, and so forth. In fact, the gestures are far more complex than this, for a gesture usually comprises movements of several articulators, and these movements are exquisitely context-conditioned. Given such complications, I must wait on others to discover how best to characterize these gestures and how to derive the articulatory movements from them. But while I'm waiting, I can be reasonably sure that the unconventional view heads the theoretical enterprise in the right direction, for it permits coarticulation. That is, it permits the speaker to overlap gestures that are produced by different organs—for example, the lips and the tongue in [ba]—and to merge gestures that are produced by different parts of the same organ—for example, the tip and body of the tongue, as in [da]—and so to achieve the high rates that are common.

But the gestures that are coarticulated, and the means for controlling them, were not lying conveniently to hand, just waiting to be appropriated by language, which brings us to the second assumption of the unconventional view: the gestures of speech and their controls are specifically phonetic, having been adapted for language and for nothing else. As for the gestures themselves, they are distinct as a class from those movements of the same organs that are used for such nonlinguistic purposes as swallowing, moving food around in the mouth, licking the lips, and so on. Presumably, they were selected in the evolution of speech in large part because of the ease with which they lent themselves to being coarticulated. But the control and coordination of these gestures is specific to speech, too. For coarticulation must walk a fine line, being constrained on either side by the special demands of phonological communication. Thus, coarticulation must produce enough overlap and merging to permit the high rates of phonetic segment production that do, in fact, occur, while yet preserving the details of phonetic structure.

The third assumption of the unconventional view is that, just as there is a specialization for the production of phonetic structures, so, too, is there a specialization for their perception. Indeed, the two are but complementary aspects of the same specialization, one for deriving the articulatory movements from the (abstract) specification of the gestures, the other for processing the acoustic signals so as to recover the coarticulated gestures that are its distal cause. The rationale for this assumption about perception arises out of the consequences of the fact that coarticulation folds information about several gestures into a single piece of sound, thereby conveying the information in parallel. This is of critical importance for language because it relaxes by a large factor the constraint on rate of phonetic-segment perception that is set by the temporal resolving power of the ear. But this gain has a price, for coarticulation produces a complex and singularly linguistic relation between acoustic signal and the phonetic message it conveys. As is well known, the signal for each particular phonetic element is vastly different in different contexts, and there is no direct correspondence in segmentation between signal and phonetic structure. It is to manage this language-specific relation between signal and appropriate percept that the specialization for speech perception is adapted. Support for the hypothesis that there is such a specialized speech mode of perception is to be found elsewhere. (See references given at the beginning of this section.) What is important for our present purposes is only that, according to this hypothesis, the percepts evoked by the sounds of speech are immediately and specifically phonetic. There is no need, as there is on the conventional view, for a cognitive translation from an initial auditory representation, simply because there is no initial auditory representation.

Now one can see plainly the difference between speech and reading/writing. In reading, to take the one case, the primary perceptual representations are, as we have seen, inherently visual, not linguistic. Thus, these representations are, at best, arbitrary symbols for the natural units of language, hence unsuited to any natural language process until and unless they have been translated into linguistic form. On the other hand, the representations that are evoked by the sounds of speech are immediately linguistic in kind, having been made so by the automatic processes of the phonetic module. Accordingly, they are, by their very nature, perfectly suited for the further automatic and natural processing that the larger specialization for language provides.

As for parity and its development in evolution and in the child, it is, on the unconventional view, built into the very bones of the system. For what evolved, on this view, was a specifically phonetic process, together with representations that were thus categorically set apart from all others and reserved for language. The unconventional view also allows us to see, as the link between sender and receiver, the specifically phonetic gestures that serve as the common coin for the conduct of their linguistic business. There is no need to establish parity by means of (innate) phonetic ideas—e.g., labels, prototypes, distinctive features—to which the several nonlinguistic representations must be cognitively associated.

How can reading/writing be made to exploit the more natural processes of speech?

The conventional view of speech provides no basis for asking this question, since there exists, on this view, no difference in naturalness. It is perhaps for this reason that the (probably) most widely held theory of reading in the United States explicitly takes as its premise that reading and writing are, or at least can be, as natural and easy as speech (Goodman & Goodman, 1979). According to this theory, called 'whole language,' reading and writing prove to be difficult only because teachers burden children with what the theorists call "bite-size abstract chunks of language such as words, syllables, and phonemes" (Goodman, 1986). If teachers were to teach children to read and write the way they were (presumably) taught to speak, then there would be no problem. Other theorists simply ignore the primacy of speech as they describe a reading process in which purely visual representations are sufficient to take the reader from print to meaning, thus implying a 'visual' language that is somehow parallel to a language best described as 'auditory' (see, for example, Massaro & Schmuller, 1975; F. Smith, 1971).

On the unconventional view, however, language is neither auditory nor visual. If it seems to be auditory, that is only because the appropriate stimulus is commonly acoustic (*pace* Aristotle). But optical stimuli will, under some conditions, evoke equally convincing phonetic percepts, provided (and this is a critical proviso) they specify the same articulatory movements (hence, phonetic gestures) that the sounds of speech evoke. This so-called 'McGurk effect' works powerfully when the stimuli are the natural movements of the articulatory apparatus, but not when they are the arbitrary letters of the alphabet. Thus, language is a mode, largely independent of end organs, that comprises structures and processes specifically adapted to language, hence easy to use for linguistic purposes. Therefore, the seemingly sensible strategy for the reader is to get into that mode, for once there, he is home free; everything else that needs to be done by way of linguistic processing is done for him automatically by virtue of his natural language capacity. As for where the reader should enter the language mode, one supposes that earlier is better, and that the phonological component of the mode is early enough. Certainly, making contact with the phonology has several important advantages: it makes available to the reader a generative scheme that comprehends all the words of the language, those that died yesterday, those that live today, and those that will be born tomorrow; it also establishes clear and stable representations in a semantic world full of vague and labile meanings; and, not least, it provides the natural grist for the syntactic mill—that is, the phonological representations that are used by the working memory as it organizes words into sentences.

The thoroughly visual way to read, described earlier, is the obvious alternative, doing everything that natural language does without ever touching its structures and processes. But surely that must be a hard way to read, if, indeed, it is even possible, since it requires the reader to invent new and cognitively taxing processes just in order to deal with representations that are not specialized for language and for which he has no natural bent.

What obstacle blocks the natural path?

As we have seen, the conventional view allows two equivalent representations of language—one auditory, the other visual—hence two equally natural paths that language processes might follow. In that case, such obstacles as there might be could be no greater for the visual mode; indeed, accepting the considerations I mentioned earlier, we should have to suppose that visual representations would offer the easier route.

The unconventional view, on the other hand, permits one to see just what it is that the would-be reader and writer (but not the speaker/listener) must learn, and why the learning might be at least a little difficult. The point is that, given the specialization for speech, anyone who wants to speak a word is not required to know how it is spelled; indeed, he does not even have to know that it has a spelling. He has only to think of the word; the speech specialization spells it for him, automatically selecting and coordinating the appropriate gestures. In an analogous way, the listener need not consciously parse the sound so as to identify its constituent phonological elements. Again, he relies on the phonetic specialization to do all the hard work; he has only to listen. Because the speech specialization is a module, its processes are automatic and insulated from consciousness. There are, therefore, no cognitively formed associations that would make one aware of the units being associated. Of course, the phonological representations, as distinguished from the processes, are not so insulated; they are available to consciousness—indeed, if they were not, alphabetic scripts would not work—but there is nothing in the ordinary use of language that requires the speaker/listener to put his attention on them. The consequence is that experience with speech is normally not sufficient to make one consciously aware of the phonological structure of its words, yet it is exactly this awareness that is required of all who would enjoy the advantages of an alphabetic scheme for reading and writing.

Developing an awareness of phonological structure, and hence an understanding of the alphabetic principle, is made the more difficult by the coarticulation that is central to the function of the phonetic specialization. Though such coarticulation has the crucial advantage of allowing speech production and perception to proceed at reasonable rates, it has the disadvantage from the would-be reader/writer's point of view that it destroys any simple correspondence between the acoustic segments and the phonological segments they convey. Thus, in a word like 'bag,' coarticulation folds three phonological segments into one seamless stretch of sound in which information about the several phonological segments is thoroughly overlapped. Accordingly, it avails the reader little to be able to identify the letters, or even to know their sounds. What he must know, if the script is to make sense, is that a word like 'bag' has three pieces of phonology even though it has only one piece of sound. There is now much evidence (1) that preliterate and illiterate people (large and small) lack such phonological awareness; (2) that the amount of awareness they do have predicts their success in learning to read, and (3) that teaching phonological awareness makes success in reading more likely. (For a summary, see, for example, I. Y. Liberman & A. M. Liberman, 1990).

Why should the obstacle loom especially large for some?

Taking the conventional view of speech seriously makes it hard to avoid the assumption that the trouble with the dyslexic must be in the visual system. It is, therefore, not in the least surprising to find that by far the largest number of theories about dyslexia do, in fact, put the problem there. Thus, some believe that the trouble with dyslexics is that they cannot control their eye movements (Pavlides, 1981), or that they have problems with vergence (Stein, Riddell, & Fowler, 1989) or that they see letters upside down or wrong side to (Orton, 1937), or that their peripheral vision is better than it should be (Geiger & Lettvin, 1989), and so on.

The unconventional view of speech directs one's attention, not to the visual system and the various problems that might afflict it, but rather to the specialization for language and the reasons why the alphabetic principle is not self-evident. As we have seen, this view suggests that phonological awareness, which is necessary for application of the alphabetic principle, does not come for free with mastery of the language. As for dyslexics—that is, those who find it particularly hard to achieve that awareness—the unconventional view of speech suggests that the problem might well arise out of a malfunction of the phonological specialization, a malfunction sufficient to cause the phonological representations to be less robust than normal. Such representations would presumably be just that much harder to become aware of. While it is difficult to test that hypothesis directly, it is possible to look for support in the other consequences that a weak phonological faculty should have. Thus, one would expect that dyslexics would show such other symptoms as greater-than-normal difficulty in holding and manipulating verbal (but not nonverbal) materials in working memory, in naming objects (that is, in finding the proper phonological representation), in perceiving speech (but not nonspeech) in noise, and in managing difficult articulations. There is some evidence that dyslexics do show such symptoms. (For a summary, see: I. Liberman, Shankweiler, & A. Liberman, 1985).

What are the implications for a theory of speech?

Those who investigate the perception and production of speech have been little concerned to explain how these processes differ so fundamentally in naturalness from those of reading and writing. Perhaps this is because the difference is so obvious as to be taken for granted and so to escape scientific examination. Or perhaps the speech researchers believe that explaining the difference is the business of those who study reading and writing. In any case, neglect of the difference might be justifiable if it were possible for a theory of speech to have no relevant implications. But a theory of speech does inevitably have such implications, and, as has been shown, the implications of the conventional theory run counter to the obvious facts. My concern in this paper has been to show that, as a consequence, the conventional theory is of little help to those who would understand reading and writing. Now I would suggest that, for exactly the same reason, the theory offers little help to those who would understand speech, for if the theory fails to offer a reasonable account of a most fundamental fact about language, then we should conclude that there is something profoundly wrong with it.

The unconventional theory of speech described in this paper was developed to account for speech, not for the difference between its processes and those of reading and writing. That it nevertheless shows promise of also serving the latter purpose may well be taken as one more reason for believing it.

Summary

The difference in naturalness between speech and reading/writing is an important fact for the psychology of language and the obvious point of departure for understanding the processes of literacy, yet it cannot be accounted for by the conventional theory of speech. Because this theory allows no linguistic specialization at the level of perception and action, it necessarily implies that the primary representations of speech are just like those of reading/writing: neither is specifically linguistic, hence both must first be translated into linguistic form if they are to serve a linguistic function. Thus, the effect of the conventional theory is to put speech and reading/writing at the same cognitive remove from language and so make them equally unnatural.

A less conventional view shows the primary motor and perceptual representations of speech to be specifically phonetic, the automatic results of a precognitive specialization for phonological communication. Accordingly, these representations are naturally appropriate for language, requiring no cognitive translation to make them so; in this important respect they differ from the representations of reading/writing. Understanding the source of this difference helps us to see what must be done if readers and writers are to exploit their natural language faculty; why reading and writing should be at least a little difficult for all; and why they might be very difficult for some.

Acknowledgment

This work was supported by NIH grant HD-01994 to Haskins Laboratories. This paper was presented at a conference on "Language and Literacy: Comparative Approaches" in Bellagio, Como, Italy in March, 1991. The conference was organized by Beatrice de Gelder and José Morais.

References

Berkeley, G. (1709). *An essay towards a new theory of vision.* Dublin: Printed by Aaron Rhames for Jeremy Pepyal.

Cole, R. A., & Scott, B. (1974). Toward a theory of speech perception. *Psychological Review, 81*, 348-374.

Crowder, R. G., & Morton J. (1969) Pre-categorical acoustic storage (PAS). *Perception & Psychophysics, 5*, 365-373.

DeFrancis, J. (1991). *Visible speech: The diverse oneness of writing systems.* Honolulu: University of Hawaii Press.

Diehl, R., & Kluender, K. (1989). On the objects of speech perception. *Ecological Psychology, 1*, 121-144.

Fujisaki, M., & Kawashima, T. (1970). Some experiments on speech perception and a model for the perceptual mechanism. *Annual Report of the Engineering Research Institute* (Faculty of Engineering, University of Tokyo), *29*, 207-214.

Geiger, G., & Lettvin, J. Y. (1988). Dyslexia and reading as examples of alternative visual strategies. In C. von Euler, I. Lundberg, & G. Lennerstrand (Eds.), *Wenner-Gren Symposium Series 54, Brain and Reading*. London: The Macmillan Press Ltd.
Goodman, K. S. (1976). Reading: A psycholinguistic guessing game. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading*. Newark, DE: International Reading Association.
Goodman, K. S., & Goodman, Y. M. (1979). Learning to read is natural. In L. B. Resnick & P. A. Weaver (Eds.), *Theory and practice of early reading*, Vol. 1. Hillsdale, NJ: Erlbaum.
Goodman, K. S. (1986). *What's whole in whole language: A parent-teacher guide*. Portsmouth, NH: Heinemann.
Kuhl, P. K. (1981) Discrimination of speech by nonhuman animals: Basic auditory sensitivities conducive to the perception of speech-sound categories. *Journal of the Acoustical Society of America, 70*, 340-349.
Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science, 190*, 69.
Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*, 1-36.
Liberman, A. M., & Mattingly, I. G. (1989). A specialization for speech perception. *Science, 243*, 489-494.
Liberman, I. Y., & Liberman, A. M. (1990). Whole language vs. code emphasis: Underlying assumptions and their implications for reading instruction. *Annals of Dyslexia, 40*.
Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle*. Research Monograph Series. Ann Arbor: University of Michigan Press.
Mattingly, I. G., & A. M. Liberman. (1988). Specialized perceiving systems for speech and other biologically significant sounds. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Functions of the auditory system* (pp. 775-793). New York: Wiley.
Mattingly, I. G. & A. M. Liberman. (1990). Speech and other auditory modules. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Signal and sense: Local and global order in perceptual maps*. New York: Wiley.
Miller, J. D. (1977). Perception of speech sounds in animals: Evidence for speech processing by mammalian auditory mechanisms. In T. H. Bullock (Ed.), *Recognition of complex acoustic signals*. (Life Sciences Research Report 5). Berlin: Dahlem Konferenzen.
Oden, G. C., & D. W. Massaro (1978). Integration of featural information in speech perception. *Psychological Review, 85*, 172-191.
Orton, S. J. (1937). *Reading, writing and speech problems in children*. New York: W. W. Norton & Co.
Pavlides, G. T. (1985). Eye movement differences between dyslexics, normal, and retarded readers while sequentially fixating digits. *American Journal of Opto and Physiological Optics, 62*, 820-832.
Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics, 13*, 253-260.
Stein, J., Riddell, P., & Fowler, S. (1988). Disordered right hemisphere function in developmental dyslexia. In C. von Euler, I. Lundberg, & G. Lennerstrand (Eds.), *Brain and Reading*. England: The Macmillan Press Ltd.
Stevens, K. N. (1975) The potential role of property detectors in the perception of consonants. In G. Fant & M. A. Tatham (Eds.), *Auditory analysis and perception of speech*. New York: Academic Press.
Wang, W. S-Y. (1981). Language structure and optimal orthography. In O. J. L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 10

On the Relations between Learning to Spell and Learning to Read

Donald Shankweiler
Eric Lundquist
University of Connecticut, Storrs

The study of spelling is oddly neglected by researchers in the cognitive sciences who devote themselves to reading. Experimentation and theories concerning printed word recognition continue to proliferate. Spelling, by contrast, has received short shrift, at least until fairly recently. It is apparent that in our preoccupation with reading, we have tended to downgrade spelling, passing it by as though it were a low-level skill learned chiefly by rote. However, a look beneath the surface at children's spellings quickly convinces one that the common assumption is false. The ability to spell is an achievement no less deserving of well- directed study than the ability to read. Yet spelling and reading are not quite opposite sides of a coin. Though each is party to a common code, the two skills are not identical. In view of this, it is important to discover how development of the ability to spell words is phased with development of skill in reading them, and to discover how each activity may influence the other. Thus, this chapter is concerned with the relationship between reading and writing.

It is appropriate to begin by asking what information an alphabetic orthography provides for a writer and reader, and to briefly review the possible reasons why beginners often find it difficult to understand the principle of alphabetic writing and to grasp how spellings represent linguistic structure. In this connection, would an orthography best suited for learning to spell differ from an orthography best suited for learning to read? The second section discusses how spelling and reading are interleaved in a child newly introduced to the orthography of English. Here, one central question is precedence: Does the ability to read words precede the ability to spell them, or, alternatively, might some children be ready to apply the alphabetic principle in writing before they can do so in reading? A related question is strategy. Do children sometimes approach the two tasks in very different ways? Finally, the last section discusses how analysis of children's spellings may illuminate aspects of orthographic learning that are not readily accessible in the study of reading.

How writers and readers are equipped to cope with the information provided by an alphabetic system

Writing differs from natural and conventional signs in that it represents linguistic units, not meanings directly (DeFrancis, 1989; Mattingly, this volume). The question of how the orthography maps the language is centrally relevant to the course of acquisition of reading and spelling. All forms of writing permit the reader to recover the individual words of a linguistic message. Given that representation of words is the essence of writing, it is important to appreciate that words are phonological structures. To apprehend a word, whether in speech or in print, is thus to apprehend (among other things) its phonology. But in the manner of doing this, A. M. Liberman (1989; this volume) notes that there is a fundamental difference between speech on the one hand and reading and writing on the other. For a speaker or listener who knows a language, the language apparatus produces and retrieves phonological structures by means of processes that function automatically below the conscious level. Thus, Liberman notes that to utter a word one does not need to know how the word is spelled, or even that it can be spelled. The speech apparatus that forms part of the species-specific biological specialization for language "spells" the word for the speaker (that is, it identifies and orders the segments). In contrast, writing a word, or reading one, brings to the fore the need for some *explicit* understanding of the word's internal structure. Since in an alphabetic system, it is primarily phonemes that are mapped, those who succeed in mastering the system would therefore need to grasp the phonemic principle and be able to analyze words as sequences of phonemes.

The need that alphabetic orthographies present for conscious apprehension of phonemic structure poses special difficulties for a beginner (see Gleitman & Rozin, 1977; I. Y. Liberman, 1973; Liberman, Shankweiler, Fischer, & Carter, 1974). The nub of the problem is this: phonemes are an abstraction from speech, they are not speech sounds as such. Hence, the nature of the relation between alphabetic writing and speech is necessarily indirect and, as we now know, often proves difficult for a child or a beginner of any age to apprehend. In order to understand why this is so it will pay us to dwell for a moment on the ways in which it is misleading to suppose that an alphabetic orthography represents speech sounds (see Liberman, Rubin, Duques, & Carlisle, 1985; Liberman et al., 1974).

First, the letters do not stand for segments that are acoustically isolable in the speech signal. So, for example, one does not find consonants and vowels neatly segmented in a spectrogram in correspondence with the way they are represented in print. Instead phonemes are co-articulated, thus overlappingly produced, in syllable- sized bundles. Accordingly, apprehension of the separate existences of phonemes and their serial order requires that one adopt an analytic stance that differs from the stance we ordinarily adopt in speech communications, in which the attention is directed to the content of an utterance, not to its phonological form. In view of this, it is not surprising to discover that preschool children have difficulty in segmenting spoken words by phoneme (see Liberman et al., 1989; Morais, 1991 for reviews).

Without some awareness of phonemic segmentation, it would be impossible for a beginning reader or writer to make sense of the match between the structure of the printed word and the structure of the spoken word. So, for example, writers and readers can take advantage of the fact that the printed word CLAP has four segments only if they are aware that the spoken word "clap" has four (phonemic) segments. Accordingly, in order to master an alphabetic system it is not enough to know the phonetic values of the letters. That knowledge, necessary though it is, is not sufficient. In order to fully grasp the alphabetic principle, it is necessary, in addition, to have the ability to decompose spoken words phonemically. Indeed, experience shows that there are many children who know letter-phoneme correspondences yet have poor word decoding skills (Liberman, 1971).

Considerable evidence now exists that children's skill in segmenting words phonemically and their progress in reading are, in fact, causally linked (e.g., Adams, 1990; Ball & Blachman, 1988; 1991; Bradley & Bryant, 1983; Byrne & Fielding-Barnsley, 1991; Goswami & Bryant, 1990; Lundberg, Frost, & Petersen, 1988). One would also expect to find that the same kind of relationship prevails between phoneme segmentation abilities and spelling. And, indeed, the data are consistent with that expectation. Studies by Zifcak (1984) and Liberman et al. (1985) have shown substantial correlations between performance on tests of phoneme segmentation of spoken words and the degree to which all the phonemes are represented in children's spellings. The findings of Rohl and Tunmer (1988) confirm this association. They compared matched groups of older poor spellers with younger normal ones and found that the poor spellers did significantly less well on a test of phoneme segmentation. (See also Bruck & Treiman, 1990; Juel, Griffith, & Gough, 1986, and Perin, 1983).

The complex relation between phonemic segments and the physical topography of speech is one sense in which alphabetic writing represents speech sounds only remotely. This, we have supposed, constitutes an obstacle for the beginning reader/writer to the extent that it makes the alphabetic principle difficult to grasp and difficult to apply. Two further sources of the abstractness of the orthography should also be mentioned, which may be especially relevant to the later stages of learning to read and to spell.

First, alphabetic orthographies are selective in regard to those aspects of phonological structure that receive explicit representation in the spellings of words (Klima, 1972; Liberman et al., 1985). No natural writing system incorporates the kind of phonetic detail that is captured in the special-purpose phonetic writing that linguists use. Much context-conditioned phonetic variation is ignored in conventional alphabetic writing,[1] in addition to the variation associated with dialect and idiolect. Hence, conventional writing does not aim to capture the phonetic surface of speech, but aims instead to create a more generally useful abstraction. It is enlightening to note, in this connection, that young children's "invented spellings"[2] often differ from the standard system in treating English writing as though it were more nearly phonetic than it is (Read, 1971; 1986).

A second source of abstractness stems from the fact that the spelling of English is more nearly morphophonemic than phonemic. English orthography gives greater weight to the morphological structure of words than is the case with some other alphabetic orthographies, for example, Italian (see Cossu, Shankweiler, Liberman, Katz & Tola, 1988) and Serbo-Croatian (see Ognjenović, Lukatela, Feldman, & Turvey, 1983). Examples of morphological penetration in the writing of English words are easy to find. A ubiquitous phenomenon is the consistent use of s to mark the plural morpheme, even in those words, like DOGS, in which the suffix is pronounced not [s], but [z]. The

---

[1]For example, it has often been noted that aspirate and inaspirate /p/, /t/ and /k/ are not distinguished in English spelling. In the word COCOA, for example, both the initial and medial consonant are spelled alike although phonetically and acoustically they are different.

[2]Often children who have had little or no formal instruction attempt to write words using the letters that they know, together with their their conceptions of the phonetic values of the letters and the segmental composition of the words they wish to write. This phenomenon has been studied extensively by Read (1986). The question of whether invented spellings can regularly be elicited from children with varied educational and family backgrounds was addressed by Zifcak (1981). In a study of 23 inner-city six year olds from blue-collar families, it was found that nearly all the children were willing to make up spellings for words though most had little knowledge of the standard orthography.

morphemic aspect of English writing appears also in spellings that distinguish words that are homophones, for example, CITE, SITE; RIGHT, WRITE.

The knowledge that spellings of some English words may sacrifice phonological transparency to capture morphological relationships brings into perspective certain seeming irregularities, as several writers have noted (Chomsky & Halle, 1968; Klima, 1972; Venezky, 1970). Homophone spellings are instances in which the two modes of representation, the phonemic and the morphemic, are partially in conflict (DeFrancis, 1989). In these spellings the principle of alphabetic writing is compromised to a degree, but it is not abandoned, since most letters are typically shared between words that have a common pronunciation. A lexical distinction in homophone pairs is ordinarily indicated by the change of only a letter or two. Thus, homophone spellings in English present an irregularity from a narrowly phonological standpoint, while nonetheless keeping the irregularity within circumscribed limits.

Such examples are telling. They led DeFrancis (1989) to make a novel and stimulating suggestion: that the needs of readers and writers may actually conflict to some degree. The convention of distinct spellings for homophones would benefit readers by removing lexical ambiguity in cases in which context does not immediately resolve the matter. Writers, on the other hand, would perhaps be better served by a system that minimizes inconsistencies in mapping the surface phonology. For writers, the presence of homophones which are distinguished by their spellings increases the arbitrariness of the orthography, and hence the burden on memory. Because it has to serve for both purposes, the standard system can be regarded as a compromise, in some instances favoring readers and in other instances favoring writers.

Scrutiny of the words that users of English find difficult to spell confirms that morphologically complex words are among those most often misspelled (Carlisle, 1987; Fischer, Shankweiler, & Liberman, 1985). Carlisle (1988) notes that in derived words the attachment of a suffix to the base may involve a simple addition resulting in no change in either pronunciation or spelling of the base (ENJOY, ENJOYMENT). Alternatively, the addition may result in a pronunciation change in the base (HEAL, HEALTH), a spelling change but not a pronunciation change (GLORY, GLORIOUS) or a case in which both spelling and pronunciation change (DEEP, DEPTH). Difficulties in spelling morphologically complex words appear to stem in part from their phonological complexity and irregular spellings. But they may also stem from failure to recognize and accurately partition derivationally related words. Carlisle (1988) tested school children aged 8 to 13 for morphological awareness. They were asked to respond orally with the appropriate derived form, given the base followed by a cueing sentence designed to prompt a derivative word (e.g., "Magic. The show was performed by a ________"). It was found that awareness of derivative relationships was very limited in the youngest children, especially in cases in which the base undergoes phonological change in the derived form (as in the above example). Moreover, the ability to produce derived forms has proven deficient in children and adults who are poor spellers (Carlisle, 1987; Rubin, 1988). All in all, the evidence supports the expectation that both phonologic and morphologic aspects of linguistic awareness are relevant to success in spelling and reading.

So far we have discussed the common basis of reading and writing, pointing first to the great divide that separates speech processes on the one hand from orthographic processes on the other. Then we proceeded to identify the factors that make learning an alphabetic system difficult. The idea was also introduced that reading and spelling may tax orthographic knowledge in somewhat different ways. It is to these differences that we turn next.

Can children apply the alphabetic principle in spelling before they are able to apply it in reading?

The possibility that the needs of readers and writers may differ with respect to the kind of orthographic mapping that is easiest to learn raises the broader issue of the relation between learning to write and learning to read. Does one precede the other? Do children adopt different strategies for the one than for the other? To answer these questions we will want to examine what is known about how spelling articulates with reading in new learners.

As to the first question, one may wonder whether precedence is really an issue. Just as in primary language development, where it is often noted that children's perceptual skills run ahead of their skills in production, so in written language, too, it would seem commonsensical to suppose that a new learner's ability to read words would exceed the ability to spell them. Most users of English orthography have probably had the experience of being unsure how to spell some words that they recognize reliably in reading. Contributing to the difficulty is the fact that there is usually more than one way for a word to be spelled that would equivalently represent its phonological structure. (Consider, for example, "clene" and "cleen" as equivalent transcriptions of the word *clean*). The reader's task is to recognize the correspondence between a letter string that stands for a word (i.e., its morphophonological structure) and the corresponding word in the lexicon. It is not required that the reader know exactly how to spell a word in order to read it—only that the printed form (together with the context) should provide sufficient cues to prompt recognition of the represented word and not some other word. In contrast, the writer must generate the one (and ordinarily only one) spelling that corresponds to the conventional standard. So it is natural to assume that spelling words requires greater orthographic knowledge than reading them. We therefore might expect that a beginner would have the ability to read many words before necessarily being able to spell them correctly.

Nonetheless, questions about precedence in the development of reading and writing have arisen repeatedly. Some writers have suggested that, contrary to the view that reading is easier, children may indeed be ready to write words, in some fashion, before they are able to use the alphabetic principle productively in reading. Montessori (1964) expressed this view, and it has more recently been articulated by several prominent researchers. In part, these claims are based on experiences with preschool children who were already writing using their own invented spellings. Carol Chomsky (1971; 1979) stressed that many young writers do this at a time when they cannot read, and, indeed, may show little interest in reading what they have written. Others who have proposed a lack of coordination between spelling and reading in children's acquisition of literacy are Bradley and Bryant (1979), Frith (1980), and Goswami and Bryant (1990).

In order to discuss the question of precedence we must first consider how we are going to define spelling and reading. By spelling, do we mean spelling a word according to conventional spelling? To adopt that criterion would ignore the phenomenon of children's invented spelling. That would seem unwise since it is well-established that some children are able to write more or less phonologically before they know standard spellings (Read, 1971; 1986). It would be appropriate for some purposes to credit a child for spelling a word if the spelling the child produces approximates the word closely enough that it can be read as the intended word.

The criterion of reading is in one sense less problematical, but in another sense it is more so. For someone to be said to have read a word, that word, and not some other word (or nonword) must have been produced in response to the printed form. It is also relevant to ask how the response was arrived at. Words written in an alphabetic system can be approached in a phonologically analytic fashion or, alternatively, they can be learned and remembered holistically (i.e., as though they were logographs). As Gough and Hillinger (1980) stress, the difficulty with the logographic strategy is that it is self-limiting because it does not enable a reader to read new words. Moreover, as the vocabulary grows and the number of visually similar words increases, the memory burden becomes severe and the logographic strategy becomes progressively more inaccurate. Should we therefore consider someone a reader if she can identify high frequency words, but cannot read low frequency words or nonwords? There is some consensus that we should not (e.g., Adams, 1990; Gleitman & Rozin, 1977; Gough & Hillinger, 1980; Liberman & Shankweiler, 1979). The possibility of reading new words, not previously encountered in print, is a special advantage conferred by an alphabetic system. It is reasonable to suppose that someone who has mastered the system will possess that ability.

However, in the view of some students of reading, most children when they begin to read, and perhaps for a considerable time afterward, read logographically, and only later learn to exploit the alphabetic principle (Bradley & Bryant, 1979; Byrne, 1992; Gough & Hillinger, 1980). Given the absence of agreement as to what is to be taken as sufficient evidence of reading ability, the question of whether spelling or reading comes first is less the issue than whether children initially employ discrepant strategies for reading and writing.

The strategy question is brought into focus by Goswami and Bryant (1990). As noted above, they suppose that the child's initial strategy in reading (the default strategy) is to approach alphabetically written words as though they were logographs. They contend that children tend to do this even when they have had instruction designed to promote phonemic awareness. Reading analytically might require more advanced word analysis skills than are available to most beginning readers. Writing, on the other hand, forces the child to think in terms of segments. The process of alphabetic writing is by its nature segmental and sequential: The writer forms one letter at a time and must order the letters according to some plan. Thus, Goswami and Bryant suppose that children's initial approaches to writing would tend to be phonologically analytic. Goswami and Bryant (1990) find it paradoxical that children's newly found phonological awareness, which most often is introduced in the context of instruction in reading, has an immediate effect on their spelling, but not on their reading. "So at first there is a discrepancy and a separation between children's reading and spelling. It is still not clear why children are so willing to break up words into phonemes when they write, and yet are so reluctant to think in terms of phonemes when they read (p. 148)."

Bryant and his colleagues (see especially Bradley and Bryant, 1979) deserve much credit for grasping the need for a coordinated approach to the study of reading and spelling. They recognized that this undertaking would require testing children on reading and spelling the same words. It is well known that performance on reading and spelling tests are highly correlated, at least in older children and adults (Perfetti, 1985; Shankweiler & Liberman, 1972). Bradley and Bryant stressed that the correlation between reading and spelling scores depends on the words chosen. They proposed that the words that children at the beginning stages find difficult to read are not always the words that are difficult to spell, and vice versa. Words that tended to be read correctly but misspelled were words whose spellings presented some irregularity, like EGG or LIGHT, whereas words spelled and not read tended to be regular words, like MAT and BUN (Bradley & Bryant, 1979).

The finding that the spell-only words and the read-only words did not overlap very much in the beginning would lend support to the hypothesis that children at this stage use different strategies for spelling and reading. The greater difficulty in spelling irregular words is what one would expect if the children were attempting to spell according to regular letter-to-phoneme correspondences. They would tend to regularize the irregular words and thus get them wrong. Moreover, the failure to read regular words suggests that the children were using some nonanalytic strategy for reading, responding perhaps to visual similarity. That would make them prone to miss easy words whenever their appearance is confusable with other words that look similar. If they were reading analytically they would read these words correctly. Thus, Bryant and his colleagues cite findings that seem to underscore the differences between early reading and spelling.

Should we, then, accept Goswami and Bryant's paradox and suppose that reading and writing are cognitively disjunct at the early stages, even in children who have received training in phonological awareness? We think not. First, as the succeeding section shows, some data (to which we turn next) point to concurrent development of reading and spelling skills. Secondly, it is too early to assess fully the impact on children's reading and spelling of the several experimental approaches to instruction in phonological awareness (e.g., Ball & Blachman, 1988; 1991; Blachman, 1991; Byrne & Fielding-Barnsley, 1991; in press). Therefore, we believe that the question must remain open.

A new research study, which coordinated the investigation of spelling and reading in six year olds (the subjects were selected only for age), does not find evidence that incompatible strategies are employed by beginners (Shankweiler, 1992). Unlike the Bradley and Bryant study, the test words in this experiment included no words with irregular spellings. The test words did contain phonological complexities, however. Each contained a consonant cluster at the beginning or the end.

There was a wide range in level of achievement within this group of six year olds. Nine of the 26 children were unable to read and spell more than one word correctly. The remaining 17 were able to read a mean of 70 percent of the words correctly but were able to correctly spell only 39 percent. These findings show that the spelling difficulties of beginners are not confined to irregular words.[3] Regularly spelled words can cause difficulty if they are phonologically complex, as when they contain consonant clusters. With the exception of one child, all read more words correctly than they were able to spell. Finally, analytic skill in reading, as indexed by ability to read nonwords, was almost perfectly correlated ($r = .93$) with spelling performance (on a variety of real words).[4] These data do not sit well with the conclusion that early reading and spelling are

[3]These results are in full agreement in this respect with those of Treiman (1993), who carried out a comprehensive study of spelling in six year olds. The findings of both studies support the caveat that one should not be too quick to attribute children's spelling errors to the irregularities of English orthography.

[4]Spelling was correlated with reading real words, .91 and .81, respectively, based on two independent measures of reading.

cognitively dissociated. On the contrary, the findings lend support to the idea that skill in reading and spelling tend to develop concurrently over a wide range of individual differences in attainment.

It is notable that spelling accuracy consistently lagged somewhat behind reading. Only 6 percent of the words were spelled correctly and read incorrectly, whereas 37 percent were read and not spelled. Thus the children showed what might be expected to be true generally: that spelling the words would prove to be more difficult than reading them, if by reading we mean correct identification of individual words, and by spelling we mean spelling these words according to standard conventions.

Interpreting error patterns in spelling and reading

So far we have been comparing spelling and reading at a coarse level of analysis. To address more rigorously the question of whether new learners use similar or dissimilar strategies for spelling and reading we would wish to make a detailed comparison between the error pattern in spelling words and reading them. But, as it happens, this turns out to be a difficult thing to do.

Problems of comparability

Most of the published information on the correlations between reading and spelling scores is based simply on right/wrong scoring. This approach has the disadvantage of throwing away much of the potential information in the incorrect responses. It fails to distinguish reading errors that are near misses from errors that are wild guesses, and it does not distinguish misspellings that capture much of a word's phonological structure from those that capture little of it. If we give partial credit for wrong responses, we must create a scheme to evaluate the many possible ways of misspelling a word and assign relative weights to each.

As an illustration of how we might proceed, we turn again to the research study last described (Shankweiler, 1992). In this study, reading was assessed by the Decoding Skills Test (DST, Richardson, & Di Benedetto, 1986). The test consists of 60 real words, chosen to give representation to the major spelling patterns of English, and, importantly, it also includes an equal number of matching nonwords, the latter formed by changing one to three letters in each of the corresponding words. For the purposes at hand, phonotactically legal nonwords constitute the best measure of reading for assessing the skills of the beginning readerbecause only these can provide a true measure of decoding skill. Because they are truly unfamiliar entities, nonwords test whether a reader's knowledge of the orthography is productive. As noted earlier, only that kind of knowledge enables someone to read new words not previously encountered in print (see Shankweiler, Crain, Brady, & Macaruso, 1992). Responses to the Decoding Skills Test were recorded on audiotape and transcribed in IPA phonemic symbols for later comparison with the spelling measures.

To gain a fine-grained measure of spelling for comparison with the reading error measures, the children's written spellings were scored phoneme by phoneme, using the following categories:

Correct spelling
Phonologically acceptable substitute (e.g., k for ck)
Phonologically unacceptable substitute (e.g., c for ch)
Phoneme not represented

When we try to compare the error pattern in reading and spelling, we encounter a further difficulty: Reading is a covert process that is assessed only by its effects. One cannot directly infer what goes on in the head when someone attempts to read a word. When we ask the child to read aloud unconnected words in list form, we encounter an obstacle: children are often unwilling to make their guesses public. Of course, a beginning reader who is stuck on a particular word may be entertaining a specific hypothesis about the word's identity, but in the absence of an overt response, we cannot discover the hypothesis and use it as a basis for inferring the source of the difficulty.

Writing, on the other hand, leaves a visible record of the writer's hypothesis about how to spell a word. The findings of the study we have been discussing bear this out. Many of the children declined the experimenter's invitation to guess at the words they were having difficulty in reading. Yet the same children produced a spelling for nearly every word they were asked to write. The upshot is that we have nearly a complete set of responses to the spelling test, but many gaps in the record occur on the corresponding items on the reading test. This yields an unsatisfactory data base for comparing the error pattern in spelling and reading. Thus, the kind of word-by-word comparison we would like to make may be unattainable.

Nonetheless, there is much to be gained by a linguistic analysis of children's spellings. Indeed, it is chiefly through their writing, and not through their reading, that children reveal their hypotheses about the infrastructure of words.

Children's conceptions of the infrastructure of words as revealed in their spellings

When encouraged to invent spellings for words, young children invent a system that is more compatible with their linguistic intuitions than the standard system. Whether the result corresponds to standard form is simply not a question that would occur to the child at this stage. In Carol Chomsky's words, creative spellers "appear to be more interested in the activity than the product (1979, p. 46)." There is evidence that children's invented spellings tend to be closer to the phonetic surface than the spellings of the standard system (Read, 1986). The standard system of English, as we noted, maps lexical items at a level that is highly abstract, both because the conventional system is morphophonemic, and because it tends not to transcribe phonetic detail that is predictable from general phonological rules.

In the comparative study of reading and writing in six year olds which we have discussed (Shankweiler, 1992), even the least-advanced beginners, who wrote only a single letter to represent an entire word, usually chose a consonant that could represent the first phoneme in the word. A child who does this is apparently aware that letters represent phonological entities even though she is not yet able to analyze the internal structure of the syllable. Altogether, first consonants were represented in 95% of cases. There was a strong tendency to omit the second segment of a consonant cluster: that is, the L in CL, the T in ST, the M in SM, the R in CR, and so forth. These were omitted in 56% of occurrences, yet when these consonants occurred alone in initial position, they were rarely omitted. Bruck and Treiman (1990) report the same trends, both in normal children and dyslexics. The tendency to omit the second segment from an initial cluster fits with Treiman's idea (1992) that children may initially use letters to represent syllable onsets and rimes rather than phonemes.[5]

[5]The onset consists of the string of consonants preceding the vowel nucleus. When the onset consists of a single consonant, as in the example of CAR, Treiman (1985) showed that children may treat it as a segment

The ability to represent the second segment of initial consonant clusters was a very good predictor of overall spelling achievement. It was also a good predictor of the accuracy of word reading. Regression analysis showed that this part score accounted for 45 percent of the variance in either spelling or reading when a different set of words is tested, after age, vocabulary (Dunn, Dunn, & Whetton, 1982) and a measure of phonemic segmentation skill (Kirtley, 1989) had already been entered. Representation of the interior segment in final clusters does almost as well when entered in the regression. The results of fine scoring give further support to the view that reading and spelling skill are closely linked even in beginners.

Why are consonant clusters a special source of difficulty? Two possibilities might be considered, each related to the phonetic complexity of clusters. First, it is well known that clusters cause pronunciation difficulties for young children. Perhaps the spelling error signals a general tendency to simplify these consonant clusters - a failure to perceive and produce them as two phonemes. But there was no indication that this was the case. All the children could pronounce the cluster words without difficulty.

An alternative possibility is that the children had difficulty in conceptually breaking clusters apart and representing them as two phonemes. In that case, the difficulty in spelling could be seen as a problem in phonological awareness. So, also, could the problems in reading the cluster words. Reading analytically would require the reader to decompose the word into its constituent segments, and the presence of clusters would increase the difficulty of making this analysis.

Research conducted during the past two decades has shown that phonological awareness is not all of a piece. Full phoneme awareness is a late stage in a process of maturation and learning that takes years to complete (Bradley & Bryant, 1983; Liberman et al., 1974; Morais, Cary, Alegria, & Bertelson, 1979; Treiman & Zukowski, 1991). Although the order of acquisition is not completely settled, there is evidence that before they can segment by phoneme children are able to segment spoken words using larger sublexical units—onsets and rimes, and syllables, particularly stressed syllables that rhyme (Brady, Gipstein, & Fowler, 1992; Liberman et al., 1974; Treiman, 1992).

The role of literacy instruction in fostering the development of phonological awareness has been much discussed in the research literature (See chapters in Brady & Shankweiler, 1991, and in Gough, Ehri, & Treiman, 1992). In this connection, Treiman (1991) urges that an analysis of spelling is the best route by which to study those aspects of phonological awareness that depend on experience with reading and writing. We would tend to agree. This is not to say, however, that *writing, but not reading* would feed this development in young children. It is to be expected that a child's interest and curiosity about the one activity would encourage and nourish an interest in the other.[6]

To sum up, because reading and writing are secondary language functions derived from spoken language, they display a very different course of acquisition than speech itself: unlike speech, mastery of alphabetic writing requires facility in decomposing words into phonemes and morphemes. Since both reading and writing depend upon grasp of the alphabetic principle, it could be expected that both would develop concurrently, though spelling, being the more difficult, would progress more slowly. Several researchers,

---

distinct from the remainder of the syllable, which corresponds to the rime. At the same time, they are unable to decompose the rime into separable components. An invented spelling, like CR for CAR or BL for BELL is consistent with such partial knowledge of the internal structure of the syllable.

[6]Adams (1990), Ehri (1989; Ehri & Wilce, 1987) and Treiman (in press) reach a similar conclusion.

however, have raised challenging questions about the order of precedence, suggesting that spelling, due to the inherently segmental nature of writing words alphabetically, emerges earlier than the ability to decode in reading. At present, the evidence is mixed. It is significant that recent research comparing children's reading and spelling errors indicates that in both spelling and reading, regularly spelled words present difficulties to beginners when the words contain phonologically-complex consonant clusters. Thus, beginners' difficulties in reading and spelling do not necessarily involve different kinds of words, as had been suggested earlier. This undercuts the claim of incompatible strategies.

Whether a child initially adopts a logographic or an analytic strategy for reading may depend in large part on the kind of pre-reading instruction the child was provided with. There is evidence that both phonological awareness and knowledge of letter-phoneme correspondences are important to promote grasp of the alphabetic principle, and are thus important to skill in spelling and decoding (Ball & Blachman, 1988; 1991; Bradley & Bryant, 1983; Byrne & Fielding-Barnsley, 1991; Gough, Juel & Griffith, 1992). Neither is sufficient alone. The phasing of these two necessary components of instruction may turn out to be critical in determining the child's initial approach to the orthography.

Acknowledgment
We are indebted to Professor Peter Bryant of the University of Oxford for making it possible for one of us (DS) to study a group of children whose data are discussed in this chapter, and for help with the analysis and much valuable discussion. Thanks are due to Stephen Crain, Anne Fowler, Ram Frost, Leonard Katz, Alvin Liberman, and Hyla Rubin for their comments on earlier drafts. This work was supported in part by Grant HD-01994 to Haskins Laboratories from the National Institute of Child Health and Human Development.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Ball, E. W., & Blachman, B. A. (1988). Phoneme segmentation training: Effect on reading readiness. *Annals of Dyslexia, 38,* 208-225.

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.

Blachman, B. A. (1991). Phonological awareness: Implications for prereading and early reading instruction. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Bradley, L., & Bryant, P. E. (1979). The independence of reading and spelling in backward and normal readers. *Developmental Medicine and Child Neurology, 21,* 504-514.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read—a causal connection. *Nature, 30,* 419-421.

Brady, S. A., Gipstein, M., & Fowler, A. E. (1992). The development of phonological awareness in preschoolers. Presentation at AERA annual meeting, April 1992.

Brady, S. A., & Shankweiler, D. P. (1991). *Phonological processes in literacy: A tribute to Isabelle Y. Liberman.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Bruck, M., & Treiman, R. (1990). Phonological awareness and spelling in normal children and dyslexics: The case of initial consonant clusters. *Journal of Experimental Child Psychology, 50,* 156-178.

Byrne, B. (1992). Studies in the Acquisition Procedure for Reading: Rationale, hypotheses, and data. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 1-34). Hillsdale, NJ: Lawrence Erlbaum Associates.

Byrne, B., & Fielding-Barnsley, R. (1990). Acquiring the alphabetic principle: A case for teaching recognition of phoneme identity. *Journal of Educational Psychology, 82,* 805-812.

Byrne, B., & Fielding-Barnsley, R. (1991). Evaluation of a program to teach phonemic awareness to young children. *Journal of Educational Psychology, 83,* 451-455.

Carlisle, J. F. (1987). The use of morphological knowledge in spelling derived forms by Learning Disabled and Normal Students. *Annals of Dyslexia, 37,* 90-108

Carlisle, J. F. (1988). Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics, 9,* 247-266.

Chomsky, C. (1971). Write first, read later. *Childhood Education, 47,* 296-299.

Chomsky, C. (1979). Approaching reading through invented spelling. In L. B. Resnick & P. A. Weaver (Eds.), *Theory and practice of early reading, vol. 2* (pp. 43-65). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Cossu, G., Shankweiler, D., Liberman, I. Y., Tola, G., & Katz, L. (1988). Phoneme and syllable awareness in Italian children. *Applied Psycholinguistics, 9,* 1-16.

DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems.* Honolulu: University of Hawaii Press.

Dunn, L. M., Dunn, L. M., & Whetton, C. (1982). *British picture vocabulary scale.* Berks., England: NFER-Nelson.

Ehri, L. C. (1989). The development of spelling knowledge and its role in reading acquisition and reading disability. *Journal of Learning Disabilities, 22,* 356-365.

Ehri, L. C., & Wilce, L. S. (1987). Does learning to spell help beginners learn to read words? *Reading Research Quarterly, 22,* 47-64.

Fischer, F. W., Shankweiler, D., & Liberman, I. Y. (1985). Spelling proficiency and sensitivity to word structure. *Journal of Memory and Language, 24,* 423-441.

Frith, U. (1980). Unexpected spelling problems. In U. Frith (Ed.), *Cognitive processes in spelling.* London: Academic Press.

Gleitman, L. R., & Rozin, P. (1977). The structure and acquisition of reading I: Relations between orthographies and the structure of the language. In A. S. Reber & D. L. Scarborough (Eds.), *Toward a psychology of reading.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Gough, P. B., Ehri, L. C., & Treiman, R. (1992). *Reading acquisition.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Gough, P. B., & Hillinger, M. L. (1980). Learning to read: An unnatural act. *Bulletin of the Orton Dyslexia Society, 30,* 179-196.

Gough, P. B., Juel, C., & Griffith, P. L. (1992). Reading, Spelling, and the orthographic cipher. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 35-48). Hillsdale, NJ: Lawrence Erlbaum Associates.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78,* 243-255.

Kirtley, C. L. M. (1989). Onset and rime in children's phonological development. Unpublished doctoral dissertation, University of Oxford.

Klima, E. S. (1972). How alphabets might reflect language. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye.* Cambridge, MA: MIT Press.

Liberman, A. M. (1989). Reading is hard just because listening is easy. In C. von Euler, I. Lundberg & G. Lennerstrand (Eds.), *Wenner-Gren Symposium Series 54, Brain and Reading.* London: Macmillan.

Liberman, I. Y. (1971). Basic research in speech and lateralization of language: Some implications for reading disability. *Bulletin of the Orton Society, 21,* 71-87.

Liberman, I. Y. (1973). Segmentation of the spoken word. *Bulletin of the Orton Society, 23,* 65-77.

Liberman, I. Y., Rubin, H., Duques, S., & Carlisle, J. (1985). Linguistic abilities and spelling proficiency in kindergartners and adult poor spellers. In D. B. Gray and J. F. Kavanagh (Eds.), *Biobehavioral measures of dyslexia* (pp. 163-176). Parkton, MD: York Press

Liberman, I. Y., & Shankweiler, D. (1979). Speech, the alphabet, and teaching to read. In L. B. Resnick & P. A. Weaver (Eds.), *Theory and practice of early reading, vol. 2* (pp. 109-134). Hillsdale, NJ: Lawrence Erlbaum Associates.

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.

Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle.* IARLD Monograph Series, Ann Arbor, MI: University of Michigan Press.

Lundberg, I., Frost, J., & Petersen, O.-P. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly, 23,* 263-284.

Montessori, M. (1964). *The Montessori method.* New York: Schocken Books.

Morais, J. (1991). Constraints on the development of phonemic awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phonemes arise spontaneously? *Cognition, 7,* 323-331.

Ognjenović, V., Lukatela, G., Feldman, L. B., & Turvey, M. T. (1973). Misreadings by beginning readers of Serbo-Croatian. *Quarterly Journal of Experimental Psychology, 35A,* 97-109.

Perin, D. (1983). Phonemic segmentation and spelling. *British Journal of Psychology, 74,* 129-144.

Perfetti, C. A. (1985). *Reading ability.* New York: Oxford University Press.

Read, C. (1971). Pre-school children's knowledge of English phonology. *Harvard Educational Review, 41,* 1-34.

Read, C. (1986). *Children's creative spelling.* London: Routledge and Kegan Paul.

Richardson, E., & DiBenedetto, B. (1986). Decoding skills test. Parkton, MD: York Press.

Rohl, M., & Tunmer, W. E. (1988). Phonemic segmentation skill and spelling acquisition. *Applied Psycholinguistics, 9,* 335-350.

Rubin, H. (1988). Morphological knowledge and early writing ability. *Language and Speech, 31,* 337-355.

Shankweiler, D. (1992). Surmounting the Consonant Cluster in Beginning Reading and Writing. Presentation at AERA annual meeting, April 1992.

Shankweiler, D., Crain, S., Brady, S., & Macaruso, P. (1992). Identifying the causes of reading disability. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 275-305). Hillsdale, NJ: Lawrence Erlbaum Associates.

Shankweiler, D., & Liberman, I. Y. (1972). Misreading: A search for causes. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationships between speech and reading* (pp. 293-317). Cambridge, MA: MIT Press.

Treiman, R. (1985). Onsets and rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Child Psychology, 39,* 161-181.

Treiman, R., & Zukowski, A. (1991). Levels of phonological awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A Tribute to Isabelle Y. Liberman.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Treiman, R. (1991). Children's spelling errors on syllable-initial consonant clusters. *Journal of Educational Psychology, 83,* 346-360.

Treiman, R. (1992). The role of intrasyllabic units in learning to read and spell. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 65-106). Hillsdale, NJ: Lawrence Erlbaum Associates.

Treiman, R. (1993). *Beginning to spell: A study of first grade children.* New York: Oxford University Press.

Venezky R. L. (1970). *The structure of English orthography.* The Hague: Mouton.

Zifcak, M. (1981). Phonological awareness and reading acquisition. *Contemporary Educational Psychology, 6,* 117-126.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 11

Phonological Awareness, Reading, and Reading Acquisition: A Survey and Appraisal of Current Knowledge

Shlomo Bentin
Department of Psychology and School of Education
Hebrew University, Jerusalem, Israel

Phonetic production and perception are part of the natural endowment of the human race. As soon as infants can be tested, they show an ability to distinguish between phonetic categories (e.g., Kuhl, 1987; Kuhl & Meltzoff, 1982; Molfeese, & Molfeese, 1979) and very early in life they are able to use phonetic elements and a few rules of combination to form phonologic structures that represent words. Children's phonological perception ability is, in fact, admirable. Even though the several phonetic gestures that are included in a phonological structure are co-articulated and therefore their acoustic effects overlap, very young children are able to decipher the phonetic code and distinguish between words on the basis of single phonemes (Eimas, 1975; Eimas, Miller, & Jusczyk, 1987; Eimas, Sequeland, Jusczyk, & Vigorito, 1971; Morse, 1972). Moreover, the deciphering of the phonetic code requires very little attention and effort. These findings lead several investigators to propose that the perception of speech is accomplished by a precognitive process controlled by a distinct biological module which is specialized to recover the coarticulated gestures from the acoustic stream and provide the cognitive system with unequivocal phonological information (Liberman & Mattingly, 1989; Mattingly & Liberman, 1990).

In contrast to their well developed phonological ability young children cannot reflect on or intentionally manipulate structural features of spoken language. Most four-to-five year-old children will not be able, for example, to tell what a word's first phoneme is, or how the word ends. Putting it differently, young children do not have the metalinguistic ability that would enable them to manipulate sub-word phonological elements (Bruce, 1964; for a recent review and an alternative perspective see Goswami & Bryant, 1990). This metalinguistic ability has been labeled *phonological awareness* (Liberman, 1973; Mattingly, 1972). The study of phonological awareness is important because the last two decades of research have provided ample evidence for its intimate relationship with reading acquisition and skill. In the present chapter I examine the nature of phonological awareness, its acquisition and development, and its role in reading acquisition.

Forms and levels of phonological awareness

By definition, awareness should be an all-or-none aptitude. In support of this view, studies in our laboratory (Leshem, unpublished doctoral dissertation), as well as in others (Calfee, Chapman, & Venezky, 1972; Stanovitch, Cunningham, & Cramer, 1984) revealed that the distribution of children's performance on tests of phonemic segmentation is bimodal: on a particular test, individual scores were either very high or very low. Additional support to this view was provided by several authors who have shown that pre-school children as well as illiterate adults can learn initial consonant deletion within a single session if they are provided with corrective feedback (Content, Kolinsky, Morais, & Bertelson, 1986; Morais, Content, Bertelson, Cary, & Kolinsky, 1988). Other authors, however, postulated that the development of explicit representation of phonemic structures could well be gradual (Content et al., 1986). This view was based on results showing that children's performance on different tests of phonological awareness varied considerably (e.g., Stanovitch Cunningham, & Cramer, 1984). For example, preschool children are relatively successful in rhyme detection tasks (Bradley, & Bryant, 1983; Lenel & Cantor, 1981; Maclean, Bryant, & Bradley, 1987), can accurately count the number of syllables in words (Liberman, Shankweiler, Fisher, & Carter, 1974), but they cannot isolate single phonemes (Liberman, Shankweiler, Liberman, Fowler, & Fisher, 1977; Rosner & Simon, 1971). The "all-or-none" view of awareness and the variability in performing different tests of phonological awareness can be reconciled by assuming that phonological awareness is a heterogenic metalinguistic competence involving abilities that differ in developmental trends and origins. Indeed, several recent reports emphasized the heterogeneous nature of phonological awareness (Bertelson & de Gelder, 1989; Bertelson, de Gelder, Tfouni, & Morais, 1989). In order to understand what the different forms of phonological awareness might be, we should first survey the ways phonological awareness has been assessed.

Because phonological awareness refers to the phonological structure of spoken words, phonological awareness tests require the ability to either detect, isolate, or manipulate sub-word phonological segments (or some combination of the above). Some tests require these aptitudes explicitly. These are, for example, phoneme isolation ("What is the first/last sound in desk?"; e.g., Bentin, Hammer, & Cahan, 1991; Wallach & Wallach, 1976), phoneme segmentation ("What sounds do you hear in the word hot?"; e.g., Fox & Routh, 1975; Williams, 1980), phoneme counting ("How many sounds do you hear in the word cake?"; Liberman et al., 1974; Yopp, 1985), and specifying a deleted phoneme ("What sound do you hear in cat, that is missing from at?"; Bentin & Leshem, in press; Stanovich et al., 1984). In other tests, correct performance requires sensitivity to sub-word phonological segments, although awareness of those segments is not explicitly tested. Such tests are, for example, the detection and/or production of rhyme ("Does sun rhyme with run?"; e.g., Calfee et al., 1972; Maclean et al., 1987), word-to-word matching ("Do pen and pipe begin the same?"; e.g., Bentin et al., 1991; Wallach & Wallach, 1976), phoneme reversal ("Say on with the first sound last and the last sound first"; Alegria, Pignot, & Morais, 1982), and phoneme deletion ("What would be left if you took out the /t/ from told?"; e.g., Bruce, 1964; Rosner, 1975; Morais, Cary, Alegria, & Bertelson, 1979). Tests also differ in the size of the segment they refer to. Some tests require awareness of single phonemes while others require awareness of sub-syllabic segments such as the word's onset or rime[1] (Kirtley, Bryant, Maclean, & Bradley, 1989; Treiman, 1985) or of syllabic segments (e.g., syllable counting; Liberman et al., 1974). Hence, phonological awareness was tested in many ways and, apparently, the observed level of "phonological awareness" was determined to some extent by the particular tests used.[2] The above survey suggests that tests of phonological awareness may differ along at least three dimensions: 1) operation required (detection, isolation, or manipulation of the phonological segment); 2) manner of testing awareness of phonological codes (indirect or explicit); and 3) size of the relevant phonological segment (syllabic, sub-syllabic, phonemic). Although the above dimensions are not entirely orthogonal (most detection tests, for example, are also indirect), a detailed examination of previous reports shows that the performance on different tests of phonological awareness varied systematically along all three dimensions.

Regardless of their size, detection of phonological segments was better than isolation, while the manipulation of segments was the poorest and latest accomplished task. For example, 29 out of 66 four-year-old children were able to detect the one word (out of three) which did not rhyme with the others, but only 8 where able to produce rhymes to target words (Maclean et al., 1987). Similarly, most studies revealed that children in kindergarten are usually very poor at isolating one phoneme of a word (Bentin et al., 1991; Lundeberg, Frost, & Petersen, 1988) or repeating an utterance after deleting one phoneme (e.g., Bruce, 1964; Rosner & Simon, 1971; Content et al., 1986), but they are more successful when they have to match words or detect oddity among words on the basis of only one phoneme (Content et al., 1986; Stanovich, Cunningham, & Crammer, 1984; Yopp, 1988). Children are more aware of syllabic and subsyllabic segments than they are of phonemic segments. For example, children start detecting rhymes and common phonemic clusters at the onset of a word much before they can match words on the basis of single phonemes (e.g., Bradley & Bryant, 1983; 1985). Similarly, preschool children are considerably more accurate in counting the syllables than the phonemes included in words (Cossu, Shankweiler, Liberman, Katz, & Tola, 1988; Liberman et al., 1974; Treiman & Baron, 1981), and the same is true for more sophisticated manipulations of syllables vs. phonemes such as segmentation (Fox & Routh, 1975; Lundberg et al., 1988) and reversal (Content, Morais, Alegria, & Bertelson, 1982; Mann, 1984). Finally, it appears that children's performance is better when phonological awareness is tested indirectly than in explicit tests. For example, counting the number of phonemes (particularly when tokens or wooden blocks are used) is more accurate than "spelling out" the sounds of a word (Yopp, 1988).

The considerable variation in pre-schoolers' performance on different tests of phonological awareness was mentioned and discussed by several authors. However, most of these authors were concerned primarily with the selection of the tests that were most reliable and best correlated with reading skills (e.g., Golnikoff, 1978; Lewkowicz, 1980; Torneous, 1984). Other authors simply partitioned the different tests into coherent groups

[1]The "onset" and the "rime" are, respectively the consonant (or consonants) that precede the vowel, and the rest of the syllable (Halle & Vergnaud, 1980; MacKay, 1972). I will elaborate on these segments and their relationship to phonological awareness later in the text.

[2]An additional test that was often used is blending, i.e. the ability to form a word by synthesizing syllables or phonemes uttered by the experimenter (e.g., Fox & Routh, 1975; Lundberg, Frost, & Petersen, 1988). I think, however, that although this test requires the manipulation of phonological units, it does not require explicit deciphering of the phonological code, and therefore it examines a skill that is basically different from phonological awareness.

(e.g., Content et al., 1986) or established a hierarchy of tests according to relative difficulty (e.g., Roberts, 1975; Stanovich et al., 1984). Only a few authors used this variation to analyze the nature and components of phonological awareness. A notable exception is the factor analysis that was recently performed by Yopp (1988) on the results of kindergarten children in ten tests of phonological awareness. The factor analysis revealed that only two factors accounted for most of the variance. Tests of phonemic segmentation, sound isolation, and phoneme counting had high loadings on Factor 1 and low loadings on Factor 2. Tests requiring the deletion of phonological segments and tests of word matching on the basis of single phonemes had moderate to high loadings on Factor 2 and low loadings on Factor 1. Because the tests that loaded Factor 2 required more steps to completion and placed a greater burden on short-term memory than the tests that loaded Factor 1, Yopp suggested that Factor 2 reflects a *Compound Phonemic Awareness* whereas Factor 1 reflects a *Simple Phonemic Awareness*. Hence, Yopp explained the variation in performance along the "operation" dimension by assuming that different levels of operation vary in the number of steps required for test completion. A stepwise regression analysis of reading scores on phonological awareness showed that both factors were good predictors of reading ability. This, however, is not surprising, since it turns out that simple phonemic awareness is in fact included in compound phonemic ability. Therefore, although the description of the two factors might help explain the variability of phonological awareness measures, it adds very little to the explanation of the relationship between phonological awareness and reading.

One of the most reliable sources of variation in phonological awareness performance is the size of the test-relevant segment. Most studies of phonological awareness showed that most preschool children can segment words into syllables but cannot manipulate or isolate single phonemes (Bruce, 1964; Calfee, 1977; Calfee, Lindamood, & Lindamood, 1973; Fox & Routh, 1975; Hakes, 1980; Liberman et al., 1974; Lundberg et al., 1988; Rosner & Simon, 1981; Treiman & Baron, 1981; Zhurova, 1963). Other studies found that four-year-old children can detect rhymes and can match words on the basis of common subsyllabic segments (e.g., Bradley & Bryant, 1985; Kirtley et al. 1989) but are unable to match words on the basis of single phonemes (Maclean et al., 1987). A possible explanation of the difference between children's ability to detect, count, and manipulate syllables or sub-syllabic clusters and their performance with single phonemes is to assume, as Content et al., (1986) did, that phonological awareness is a gradually developing ability, or that there are "levels" of phonological awareness (e.g., Goswami & Bryant, 1990). However, it is also possible that there is a qualitative distinction between the awareness of single phonemes and the awareness of multi-phonemic structures which accounts for the observed difference in performance with the two types of segment. In other words, it is possible that awareness and manipulation of single phonemes and detection and sensitivity to syllabic or intrasyllabic structures are qualitatively different forms of phonological awareness rather than two levels along a continuum of one ability. I will try to defend this qualitative distinction.

As a consequence of the process of coarticulation that characterizes speech production, the sound frequency patterns forming acoustic segments in speech reflect the combined contribution of several complex gestures, each intended to produce a different phone. Moreover, because a phone can be coarticulated with different phonetic contexts, there can be no direct correspondence in segmentation between the acoustic signal and the phonetic message it conveys. Therefore, speech perception cannot be based on a simple translation from a set of auditory representations to a set of perceptual phonetic categories. Consequently, awareness of each of the phonemes conveyed by one acoustic segment probably follows a more basic and automatic process of phonetic deciphering. This is probably why, although phonetic distinctions in speech are easy and natural, awareness of phonetic categories appears much later in ontogenetic development and probably requires more than simple cognitive maturation. This awareness require the ability to break up the coarticulated phonological segments and isolate their individual phonemic constituents.

The above analysis implies that segmentation should be relatively easy when the required phonological units correspond to perceived acoustic segments but difficult when the disentangling of coarticulated phones is required. Coarticulated phonological units usually include a highly resonant nucleus (a vowel) flanked by one or several consonants, together forming a syllable. Therefore, syllabic segmentation can be based on simple auditory perception and might not reflect genuine phonological awareness. This view also suggests that the isolation of stop consonants should be significantly more difficult than the isolation of steady-state vowels because the former have no independent acoustic existence—they are always coarticulated. The latter hypothesis, however, is only partly supported by empirical evidence. Previous studies of initial phoneme isolation (Bentin & Leshem, in press) and initial phoneme deletion (Content et al., 1982; Content et al., 1986) suggested that the performance of pre-school children was better with vowels than with consonants; that order of difficulty was reversed, however, when the last (rather than the first) phoneme had to be isolated: Final consonants were easier to isolate than final vowels (Bentin & Leshem, in press). A similar pattern was found when performance with stop consonants was compared to performance with fricatives (Content et al., 1986).

In contrast to the commonly reported failure of pre-literate children to isolate and manipulate single phonemes which are perceived in coarticulated form, most studies demonstrate that children are considerably more successful in detecting and producing rhymes. The sensitivity to rhymes might be taken as evidence for a second form of phonological awareness because it also requires the breaking of coarticulated phonetic clusters. For example, the recognition that the monosyllabic words "beg" and "leg" rhyme involves breaking them into b-eg and l-eg segments, and recognizing that the end segments of each syllable sound alike. Because the same children cannot usually tell that /b/ is the first and /g/ is the last phone in "beg," it is conceivable that rhyme detection and phonemic segmentation require different phonological skills. The most outstanding attempt to explain this difference was made by Peter Bryant, Lynette Bradley and their collaborators at Oxford.

The basic idea advocated by the Oxford group is that there are linguistically valid segments intermediate between single phonemes and syllables. These segments were labeled *onset* and *rime*. The onset is the consonant or string of consonants that precedes the vowel in a syllable, and the rime is the rest of the syllable. For example the onset of the monosyllabic word "black" is /bl/ and its rime is /ack/. Note, that although the onset and the rime are phonologically defined units, the validity of this distinction was based either on observing the nature of errors in speech (MacKay, 1972), or on linguistic constraints on sequences of phonemes (Halle & Vergnaud, 1980). Hence, the validity of this phonological categorization is not based on phonetic considerations and is very different from the distinction between phonetic categories that was discussed above. Nevertheless, awareness of this intrasyllabic segmentation and the ability to manipulate these segments require, as mentioned above, breaking the coarticulated unit of perception, and they may therefore be considered a form of phonological awareness.

Words that rhyme share, by definition, the same rime. Therefore, the reliable demonstrations that four-year-old children can detect and produce rhymes proves that they are aware of the rimes of syllables. Are they also similarly aware of the onsets? That evidence is less compelling. Kirtley et al. (1989) attempted to demonstrate such an awareness, however most of their evidence is based on negative findings and their interpretations are speculative. In their study they used oddity tasks with different word sets. First they replicated the finding that it is easier to find an "odd" word among four when the commonality is based on the initial consonant than on the final consonant. Their interpretation of this phenomenon was that the initial consonant formed the whole onset of the word whereas the final consonant was only a part of the rime. However, in order for this interpretation to hold unequivocally they would have had to show that when the initial consonant was only a part of the onset (such as the /s/ in "string") its detection should be more difficult. Unfortunately such a comparison has not been attempted. Moreover, our own observations (Leshem, unpublished doctoral dissertation) suggest that the opposite is true. Our five-year-old subjects were more successful in isolating the initial consonant in words that began with a CCV string than in words that began with a CV string.

In a second experiment Kirtley et al. (1989) used different oddity combinations aimed at distinguishing between situations in which the odd word could be detected on the basis of a full intrasyllabic segment or required the breaking of such segments. In all the conditions the results supported the prediction that it is easier to detect oddity on the basis of intact intrasyllabic structures. However, the same results could be interpreted solely on the basis of special sensitivity to the rime, without making any assumptions about the onset. Moreover, as in the first experiment, no multi-phonemic segments were used, and so the "onset" was always confounded with a single initial phoneme. Nevertheless, it should be stressed that the authors' interpretation is not counter-intuitive and may be right. At this time, however, all we can say is that there is strong evidence for a particular sensitivity to the rime of syllables which may have been induced by extensive experience with rhymes. Moreover, this form of sensitivity was shown only in detection tasks. To the best of my knowledge, awareness of onset and rimes has never been shown in tests of segmentation or isolation. Hence, it is possible that children who are able to detect rhymes and correctly select the odd word in an oddity tests are sensitive to sub-syllabic units but still unable to point out the phonological segment on which their decision is based. Consequently, it is possible that sensitivity to sub-syllabic segments, either as suggested by the Oxford group or limited to rimes only, reflects a qualitatively different form of phonological awareness than sensitivity to single phonemes. Later in this chapter I discuss the relevance of both forms of phonological awareness to reading.

In conclusion, this section shows that there are, in fact, only two forms of phonological awareness: One which is demonstrated by the ability to isolate segments and manipulate single phonemes, and one demonstrated by sensitivity to the rime and perhaps to the onset of syllables. The first requires explicit knowledge about the phonemic segment and, therefore I will label it "phonemic awareness"; the second is reflected indirectly in the detection of oddity and commonality between words on the basis of subsyllabic segments. I will label this second form "early phonological awareness." Other tests vary along different dimensions but do not reflect any separate ability.

Factors influencing the development of phonological awareness

Clearly, phonological awareness is not an innate aptitude; it is probably triggered by some experience. Therefore, the questions of "how" and "when" this skill appears have frequently been raised and have been the subject of much controversy. Some authors claimed that phonological awareness is triggered, or at least considerably enhanced, by exposure to the alphabet (e.g., Bertelson et al., 1985; Bertelson & de Gelder, 1990). Others proposed that phonological awareness develops a long time before children learn to read, through experiences which at the time have nothing to do with reading (e.g., Bryant & Bradley, 1985). As we will see, however, these are not mutually exclusive theories, because each of the proponents is actually talking about of a different form of phonological awareness.

There is ample evidence that learning to read affects phonological awareness skills. For example, using consonant addition and deletion tasks, Read, Zhang, Nie, and Ding (1986) found well-developed phonological awareness in Chinese subjects who learned to read a recently developed Chinese alphabetic system (Pinyin) but not among subjects who read only the logographic system (Kanji). The mean percentage of correct performance was 83% in the former group but only 21% in the later. Along the same lines, Mann (1986) reported that first-graders in Japan who learned how to read a syllabary (Kana) were good at manipulating syllables but significantly inferior to American first-graders in manipulating phonemes. Equivalent results were found with Belgian children in the first grade; those who learned to read according to the "analytic" (segmental) method performed better on tests of phonemic segmentation than those who learned to read by the "global" (holistic) method (Alegria, Pignot, & Morais 1982). However, the strongest support for the view that in the absence of reading acquisition phonemic segmentation skills do not develop spontaneously is provided by a series of studies by Morais and his colleagues showing that illiterate adults perform very poorly on tests of phoneme deletion, although they may manipulate phonology at syllabic and word levels (Morais et al. 1979, 1986, 1987). Similar results were also found with semi-literate adults (Read & Ruyter, 1985) and with the reading disabled (Bryne & Ledez, 1983). The ability of the illiterate subjects to manipulate multi-phonemic units as opposed to single phonemes is congruent with findings in preschool children.

As reviewed in the previous section, there is no doubt that some three-year-old children and most four-year-old children recognize and play with rhymes (e.g., Chukovsky, 1963). Formal testing has shown that when either detection through oddity or the production of rhyme and alliteration was involved, many three- and four-year-old children could make judgments about the component sounds (particularly rimes) in words that they heard or uttered (Maclean et al., 1987). The significant ability of pre-literate children to detect rhymes as well as to perform above chance in oddity tests based on sensitivity to subsyllabic segments (Bradley & Bryant, 1985; Kirtley et al., 1989) lead to the conclusion that phonological awareness exist before reading acquisition (Maclean et al., 1987). Note, however, that the apparent disagreement between the two views stems from a different definition of phonological awareness. Those who found signs of phonological awareness in three and four-year-old children refer primarily to what we called the early form of phonological awareness—the one which focuses on subsyllabic segments and is tested indirectly. In contrast, the defenders of the alternative view, (i.e., that phonological awareness is triggered by the exposure to the alphabetic principle), refer to phonemic awareness—the one reflected in the ability to explicitly manipulate single phonemes deciphered from the coarticulated unit. This fact was recently recognized by both parties (Bertelson et al., 1989; Goswami & Bryant, 1990).

Having resolved the above controversy, we are still left with several important questions. What are the factors that affect the development of the two forms of phonological awareness? Is the early form a precursor of the later form? Does the early form develop spontaneously or is it the result of explicit or implicit instruction? Is exposure to the alphabetic principle the only factor influencing the development of phonemic awareness? Can the development of phonemic awareness be accelerated and achieved prior to exposure to the alphabetic principle? The available literature may provide answers to some of these questions.

The impressive studies reported by Bradley and Bryant (1985) prove beyond any reasonable doubt that explicit training with sound categorization improves performance on oddity tests based on rime and onset. In other words, the early form of phonological awareness can be significantly improved in kindergarten by explicit training. This, however, does not prove that this metaphonological ability cannot occur spontaneously. A direct answer to this question requires a rigid control of children's pre-test experience with rhymes. Obviously, it is practically impossible to control children's experience in life, and so we are forced to address this question only indirectly. For example, in a longitudinal study by Maclean et al. (1987), young children's performance on different tests of rhyme and alliteration detection as well as their knowledge of nursery rhymes, was related to their socio-economical background and their parents' education. Although it is a rough estimate, it would not be completely wrong to assume that children of middle-class highly educated parents had more opportunities to be exposed to nursery poems and other forms of rhymes than chidden coming of lower-class poorly educated parents. Therefore, a significant difference between the performance of the two groups might indicate that experience with rhymes is a critical trigger of the early phonological awareness. The results of this comparison suggested that at the earliest age tested (3 years old) children coming from the "privileged" homes were more successful in the detection of alliteration than the other children, but there were no differences in the detection of rhymes, and both groups were equally knowledgeable about nursery rhymes. Moreover, even the small difference did not last. There was no sign of influence of family background after the initial tests. On the basis of these results, and considering that illiterate adults who showed no phonemic awareness were nevertheless sensitive to rhyme judgments and vowel deletion (e.g., Bertelson et al., 1989), we may safely conclude that the early phonological awareness, which does not require awareness of single phonemes, can be easily triggered without explicit instruction and may develop independently of reading acquisition.

In contrast to their sensitivity to syllabic and sub-syllabic phonological segments, evidence from studies with illiterates suggests that the ability to isolate and manipulate single phonemes that are coarticulated in speech (i.e., phonemic awareness) does not develop spontaneously (Morais et al., 1979; Morais, Bertelson, Cary, & Alegria, 1986). These authors proposed that learning to read an alphabetic orthography provides most children (and adults) with the opportunity to develop full phonemic awareness. In contrast to speech, where individual phonemes are coarticulated, in writing the phonemes are represented by clearly defined orthographic segments, the letters. Assuming that children learn about these letter-sound correspondence when they learn to read, it seems likely that during the acquisition of reading skills they become explicitly aware that words are formed of the sounds which the letters represent. Indeed, most studies revealed a significant gap between the phonemic segmentation skills of first-graders and of kindergarten children. For example, Liberman et al. (1974) found that none of the pre-kindergartners and only 17% of the kindergartners tested were able to parse words into phonemes, while 70% of the first-graders tested succeeded in doing so.

A caveat about interpreting developmental studies of phonological awareness, and particularly the striking improvement in phonemic segmentation ability during the first grade, is that all such studies share the serious problem of the possible confounding of differences in the extent or method of reading acquisition with other age-related variables that may have influenced phonological awareness (e.g., the amounts of informal linguistic experience and general cognitive development). In addition, the comparison of illiterate and ex-illiterate adults may be compromised, for example, because the choice to join a literacy program in adulthood was probably not arbitrary. Therefore, before definite claims about a causal relationship between reading acquisition and the emergence of phonemic awareness could be made, it was still necessary to isolate the effect of reading acquisition on the appearance and development of awareness of individual phonemic segments.

Owing to the impossibility of experimenting with elementary school attendance, previous attempts to control for general age-related effects on phonological awareness were based on comparisons between the youngest and the oldest children within one grade level (Bowey & Francis, 1991), or between the oldest children in the kindergarten and the youngest children in the first grade (Bowey & Francis, 1991; Morrison, 1988). Although suggestive, this approach suffers from a serious shortcoming of selection, because the cutoff date for school admission is never strictly imposed. Moreover, the exceptions are not random: Intellectually advanced children who are slightly younger than the official school age are often admitted, while children who are somewhat older than the cutoff point but insufficiently developed may be held back an additional year (Cahan & Davis, 1987; Cahan & Cohen, 1989). This creates a situation of "missing" children in each grade, particularly among children at the extreme ages. Such selective misplacement usually leads to overestimation of the schooling effect (Cahan & Cohen, 1989).

In a recent study Bentin, Hammer, & Cahan (1991) proposed a solution to this problem. Rather then comparing empirically obtained data from children at the extreme ages in each grade, the authors predicted these data on the basis of the best fitting regression of test scores on chronological age, across the entire legal age range in each grade, with the exclusion of the selection-tainted birth dates near the cutoff point. The separate effects of schooling and one year of age were estimated by means of a regression discontinuity design (Cook & Campbell, 1979) involving the regressions of phonemic segmentation scores on chronological age. The effect of age was reflected by the slope of the within-grade regressions, whereas the effect of schooling was reflected in the discontinuity between the two regression lines. The results of this analysis are presented in Figure 1.

As evident in Figure 1, the percentage of correct responses on the phonemic segmentation battery was higher in school children (76%, SD=14%) than in the kindergarten group (35%, SD=23%) ($t(674)=29.12$, $p<.0001$). However, this difference reflected the combined effects of age and schooling. The comparison of the independent schooling and age effects revealed that, although both effects were statistically significant, the effect of schooling (reading acquisition) was four times as large as the effect of one year of chronological maturation.

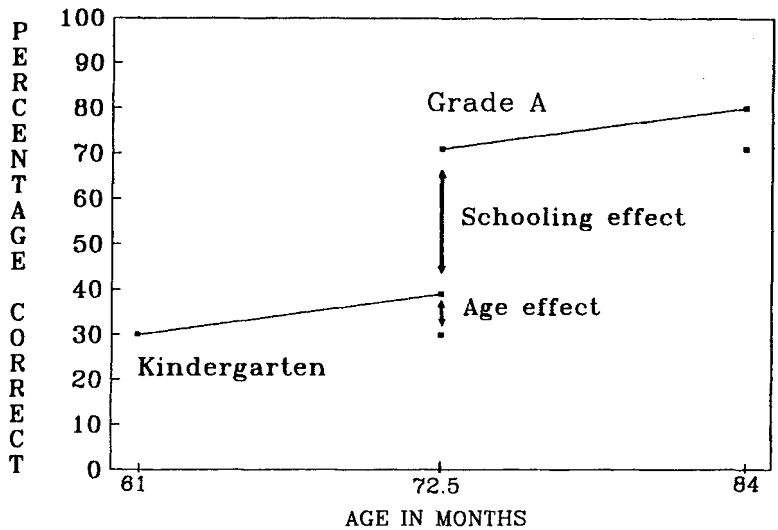

Figure 1. **Schooling and age effects on the development of phonological awareness in kindergarten and Grade A children.**

The results of the Bentin et al. (1991) study pointed to schooling (learning to read) as a major factor affecting the development of phonological awareness. This is not to say, however, that exposure to an alphabetic orthography is the only way to trigger phonemic awareness. There is ample evidence for the efficiency of tuition in metaphonological skills outside the context of reading acquisition (e.g., Lundberg et al., 1988). Significant improvement in phonemic segmentation skills were obtained using different training methods such as the use of visual aids to represent phonemes (Elkonin, 1973; Lindamood & Lindamood, 1969), the designing of speech-correction games played with puppets that impersonated human speakers (Content et al., 1982), simply using corrective information during testing in successive blocks (Content et al., 1986), and designing speech-sound oriented group games (Olofsson & Lundberg, 1983). In all these studies, however, the experimental groups were selected from a normal population of children. Moreover, in many of the previous studies the control groups were not trained for other language abilities (but see Ball & Blachman, 1991). Therefore, the specific effect of training in phonological awareness may have been confounded with the positive effects that training in general linguistic skills may have on reading acquisition and might have been limited to linguistically well-developed children. In a recent study we rectified these problems, finding that training in segmentation skills significantly improved phonemic segmentation ability in five-year-old children who were initially at the lower end of the distribution of scores on a battery of phonemic segmentation tests (Bentin & Leshem, in press). Moreover, in that study we found that children who had been trained in phonemic segmentation were able to apply their newly acquired metaphonological skills in other tests of phonological awareness.

In conclusion, the presently available evidence suggests that early phonological awareness, i.e., the ability to detect and produce rhymes and the sensitivity to subsyllabic segments, develops differently from phonemic awareness (i.e., the ability to isolate and manipulate individual phonemes in speech). The former appears to emerge almost automatically and instantaneously in the great majority of children when they are first exposed to nursery rhymes or other forms of phonological word games and develops independently of reading instruction. The latter, on the other hand, is triggered in most children when they come to understand the alphabetic principle during the acquisition of reading in an alphabetic orthography. However, phonemic awareness can be also be triggered and full phonemic awareness can be developed in pre-readers by explicit training of phonemic segmentation skills. There is no direct evidence for interdependence between the two forms of phonological awareness. It is conceivable, however, that well-developed awareness of rhymes and subsyllabic segments is necessary for a smooth acquisition of phonemic awareness during reading instruction. In other words, it is possible that a well-developed early phonological awareness is a prerequisite for the emergence of phonemic awareness without explicit instruction. Indirect evidence for this hypothesis is reviewed in the next section.

Reading and phonological awareness: It's a two-way street

Although studying the development of metaphonological skills is important in its own right, the significance of phonological awareness is considerably enhanced by its well-established relationship with the acquisition of reading skills. Many studies have demonstrated that children's performance in various phonological awareness tests highly correlates with their reading skill in the early school grades in English (Bradley & Bryant, 1985; Calfee et al., 1973; Fox & Routh, 1975; Liberman et al., 1977; Rosner & Simon, 1971; Treiman & Baron, 1981; Tunmer & Nesdale, 1986), as well as in other languages such as Italian (Cossu et al., 1988), Swedish (Lundberg, Olofsson, & Wall, 1980), Spanish (de Manrique & Gramigna, 1984), French (Bertelson, 1987), and Hebrew (Bentin & Leshem, in press). Correlative studies were applied in developing tools for predicting success in reading (Blachman, 1984; Juel, Griffith, & Gough, 1986; Lundberg et al., 1980; Mann, 1984; Share, Jorm, MacLean, & Matthews, 1984); however, they tell us very little about the nature of the relationship. A high positive correlation might exist between two independent skills if they are similarly affected by a third factor. On the other hand, it is also possible that the correlation reflects a causal relationship, as, for example, when one skill is a pre-requisite or trigger for the second.

Theoretical considerations suggest that phonological awareness and the acquisition of the alphabetic principle are directly interdependent, and that the positive correlation might reflect mutual influence and even causal relations between these two skills. The alphabet is the latest and probably the most advanced form of writing (DeFrancis, 1989). One of its most important virtues is that, like speech, it uses a relatively small set of well-defined symbols (the letters) that can be combined in a practically infinite number of ways to represent all the possible words in a language. The representation of words by orthographic patterns is efficient only because the basic units of writing, the letters, are mapped onto the basic units of speech, the phones. Thus, words are not represented in writing by arbitrary and holistically distinguished patterns but rather, the combination of letters that represents a particular word is fully determined by the sequence of phonemes of which the word is composed. Hence, in order to understand a written word the reader must be able to decipher the phonological unit from its written form. Even assuming that a fluent reader may form direct associations between some written patterns and their meanings, and use these associations to access the semantic information directly, the ability to decipher phonology from writing is a prerequisite for reading and understanding written words at the first encounter, and needs to be mastered before efficient reading can occur. This is the essence of the alphabetic principle, and this is the reason why reading and writing require a reasonable awareness of the internal phonological structure of spoken words. (For detailed discussion of these considerations see, for example, Ehri, 1979; Leong, 1986; Liberman, 1989; Liberman & Liberman, 1990; Liberman, Shankweiler, & Liberman, 1989; Rozin & Gleitman, 1977.)

The above account for the reading process implies that, regardless of the particular teaching method adopted by the teacher, in the process of learning to read children learn the basic mapping rules from the domain of letters to the range of phonemes. Obviously, the acquisition of mapping rules requires explicit knowledge of the members of the domain and of the range. The items in the domain (the letters) are explicitly taught by the teacher. On the other hand, the members of the range (the phonemes) are not explicitly taught in the classroom. When children start learning to read they are expected to be aware of the phonological structure of spoken words, or at least to become aware of it very quickly. Indeed, as reviewed in the previous section, most children become aware of the phonemic structure of spoken words fairly easily, as a consequence of exposure to the alphabet, which leads to the understanding of the alphabetic principle. Unfortunately, for a significant proportion of children mere exposure to the alphabet is not sufficient, and they consequently develop a reading disability. Several studies have demonstrated that these children may be helped by explicit training in phonological awareness in parallel to reading acquisition (Perfetti, Beck, Bell, & Hughes, 1987; Wallach & Wallach, 1976; Williams, 1980) or preferably during kindergarten (Ball & Blachman, 1991; Bentin & Leshem, in press; Bradley & Bryant, 1983, 1985; Lundberg, Frost, & Peterson, 1988; Vellutino & Scanlon, 1984). A survey of these studies may shed additional light on the metaphonological prerequisites of reading acquisition.

In their initial longitudinal study, Bradley and Bryant (1985) trained four- and five-year-old children to categorize words on the basis of initial sound, and to be aware of that common sound. Some of the children were also given experience with plastic letters. Children in control groups were trained for conceptual categorization or received no training whatsoever. When they reached school, the reading, spelling and mathematical ability of the children in the four groups were compared. The results of these comparisons showed that the children who had been trained to categorize words on the basis of initial phonemes were better in reading and spelling than the children in the control groups, whereas the mathematical skills of all four groups were equal. It was also found that the reading and spelling performance of children who were given experience with plastic letters in addition to phonemic categorization surpassed that of children who were trained only in sound categorization. Finally, a follow-up of this study (Bradley, 1989) revealed that the advantage gained by the experimental groups was maintained five years later: At the age of 13 years, their reading performance was still better than that of the control groups. It is important to note that in this early study the children were trained to make phonemic distinctions, because in more recent publications the Oxford group seems to be convinced that the form of phonological awareness important for reading acquisition is sensitivity to the onset and rime of syllables (Goswami & Bryant, 1990; Maclean et al., 1987).

Although a substantial positive correlation was found between the early phonological awareness and reading acquisition, I doubt that awareness of subsyllabic segments alone is sufficient for understanding the alphabetic principle. First, although several letter-strings frequently appear together (for example */ing/*), many do not. In fact, in reading, as in speech, the distinction among words is frequently based on one letter. Therefore, I am in greater agreement with the following conclusion drawn by the same group: "... a major step in learning to read may take place when the child learns to break the rime into its constituent sounds by detaching ... the preceding vowel from the final consonant" (Kirtley et al., 1989). In fact, there is ample evidence that in training phonemic segmentation facilitates reading acquisition (e.g., Ball & Blachman, 1991; Cunningham, 1988, in press; Lundberg et al., 1988) but not a single study in which training *only* in rhyming skills facilitated reading acquisition. In this context it is interesting to mention our own training study (Bentin & Leshem, in press), because the structure of Hebrew orthography, in which vowels are represented by diacritical marks appended to the consonants (see Frost & Bentin, this volume), would be the ideal orthography to make use of onset and rimes rather than phonemes in reading.

In our study we trained four groups of five-year-old children selected from the lower end of the distribution of scores on a phonemic segmentation test-battery. Group I was trained in phonemic segmentation; group II was trained in phonemic segmentation and also in recognizing letters of the alphabet and relating them to their sound; group III was trained in general linguistic abilities such as vocabulary enhancement, sentence comprehension, etc.; group IV received no training. Training, in groups of four children, lasted for 10 weeks with two 1/2-hour sessions per week. A year later, the reading performance of these children was assessed and compared with the performance of children who were comparable to the four training groups, except for being at the higher end of the distribution of scores of the initial phonemic awareness battery (Group V). The children were tested after four months and nine months of reading instruction. Each test consisted of lists of items that the children were instructed to read aloud. Two lists included words and two lists nonwords. The lists included an equal number of monosyllabic and disyllabic items. Table 1 presents the percentage of correctly read words in each group, for each stimulus type.

As evident in Table 1, reading skills in the first grade were significantly correlated with the phonemic segmentation skills that were assessed in the kindergarten before training, and were influenced by training segmentation skills. Because there are no standardized reading tests in Hebrew (except for reading comprehension) it is difficult to interpret the absolute scores. Note, however, that these tests were constructed in collaboration with the teachers in the respective schools and were designed to reflect the expected level of reading at each testing time. Therefore, it is suggestive to observe that the reading performance of children who were initially low in phonemic awareness and received no training in phonemic segmentation was about 40%, which according to school standards means failure. In contrast, children from the same population who received training and improved their phonemic awareness scored around 70%, almost as well as children who were initially high in phonemic awareness. These data are particularly important because we tested children who learn to read an orthography in which, because of its specific characteristics, the basic segment usually used by teachers for reading instruction is a consonant-vowel combination.

Table 1

Percentage of correctly read items (SEm) of each stimulus type after 4 months and 9 months of reading instruction. Note that different tests were given each time, to correspond with the respective reading level.

STIMULUS TYPE	FIRST READING TEST				
	GROUP I	GROUP II	GROUP III	GROUP IV	GROUP V
WORDS					
One Syllable	85.5 (4.7)	76.7 (8.5)	63.4 (8.5)	59.9 (6.3)	94.5 (2.0)
Two Syllables	72.9 (4.9)	66.1 (7.9)	46.2 (7.8)	46.3 (7.4)	87.1 (3.7)
NONWORDS					
One Syllable	64.7 (7.9)	43.4 (9.7)	26.1 (7.0)	28.3 (8.2)	75.0 (5.2)
Two Syllables	58.4 (7.8)	45.3 (9.7)	25.5 (7.2)	24.4 (8.5)	64.9 (7.7)
	SECOND READING TEST				
WORDS					
One Syllable	66.6 (6.0)	70.7 (5.7)	35.2 (6.1)	38.6 (11.)	68.2 (4.8)
Two Syllables	63.9 (5.6)	66.2 (7.8)	21.8 (4.2)	29.6 (10.)	64.4 (4.5)
NONWORDS					
One Syllable	63.4 (6.3)	59.6 (7.7)	27.0 (5.3)	28.8 (11.)	67.4 (6.8)
Two Syllables	52.7 (5.8)	47.5 (9.7)	16.3 (3.5)	19.6 (8.6)	57.5 (7.5)

In conclusion, the evidence relating reading acquisition to phonological awareness is robust. It suggests that the alphabetic principle requires the ability to isolate and manipulate single phonemes in coarticulated speech. The major factor that triggers this ability is exposure to the alphabet. However, phonemic awareness cannot be triggered by the alphabet unless the early form of phonological awareness is well developed. Children who do not meet this prerequisite must be explicitly trained for phonemic segmentation. Our data show that training phonemic segmentation in kindergarten for a relatively short period is effective in inducing the metaphonological skills required for easy acquisition of reading. With younger children, however, or with children who are language-delayed the training program should probably begin with the establishment or improvement of sensitivity to rhymes and the ability to detect the onset and rime of the syllables.

Acknowledgment
This study was supported by the Israel Foundations Trustees and by NICHHD Grant HD-01994 to Haskins Laboratories.

References

Alegria, J., Pignot, E., & Morais, J. (1982). Phonetic analysis of speech and memory codes in beginning readers. *Memory & Cognition, 10*, 451-556.

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*, 49-65.

Bentin, S., Hammer, R., & Cahan, S. (1991). The effects of aging and first year schooling on the development of phonological awareness. *Psychological Science, 2, 271-274.*

Bentin, S., & Leshem, H. (in press). On the interaction of phonological awareness and reading acquisition: It's a two-way street. *Annals of Dyslexia.*

Bertelson, P., & De Gelder, B. (1990). The emergence of phonological awareness: Comparative approaches. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the Motor Theory of Speech Perception.* Hillsdale, NJ: Erlbaum.

Bertelson, P., de Gelder, B., Tfouni, L. V., & Morais, J. (1989). Metaphonological abilities of adult illiterates: New evidence of heterogeneity. *European Journal of Cognitive Psychology, 1*, 239-250.

Bertelson, P., Morais, J., Alegria, J., & Content, A. (1985). Phonetic analysis capacity and learning to read. *Nature, 313*, 73-74.

Blachman, B. A. (1984). Language analysis skills and early reading acquisition. In G. Wallach & K. Butler (Eds.), *Language learning disabilities in school-age children* (pp. 271-287). Baltimore, MD: Williams and Wilkins.

Bowey, J. A., & Francis, J. (1991). Phonological analysis as a function of age and exposure to reading instruction. *Applied Psycholinguistics, 12*, 91-121.

Bradley, L. (1989). Predicting learning disabilities. In J. J. Dumont & H. Nakken (Eds.), *Learning disabilities: Cognitive, social and remedial aspects* (pp. 1-17). Amsterdam: Swets & Zeitlinger.

Bradley, L., & Bryant, P. (1983). Categorizing sounds and learning to read: A causal connection. *Nature, 301*, 419-421.

Bradley, L., & Bryant, P. (1985). *Rhyme and reason in reading and spelling.* Ann Arbor: University of Michigan Press.

Bryant, P., & Bradley, L. (1985). *Children's reading problems*. Oxford: Blackwell.

Bruce, D. J. (1964). The analysis of word sounds. *British Journal of Educational Psychology, 34*, 158-170.

Byrne, B., & Ledez, J. (1983). Phonological awareness in reading disabled adults. *Australian Journal of Psychology, 35*, 185-197.

Cahan, S., & Cohen, N. (1989). Age vs. schooling effects on intelligence development. *Child Development, 60*, 1236-1249.

Cahan, S., & Davis, D. (1987). A between-grades-level approach to the investigation of the absolute effects of schooling on achievement. *American Educational Research Journal 24*, 1-13.

Calfee, R. C., Chapman, R., & Venezky, R. (1972). How a child needs to think to learn to read. In L. W. Gregg (Ed.), *Cognition in learning and memory.* New York: John Wiley & Sons.

Calfee, R. C., Lindamood, P., & Lindamood, C. (1973). Acoustic-phonetic skill and reading—kindergarten through twelfth grade. *Journal of Educational Psychology, 64*, 293-298.

Chukovsky, K. (1963). *From two to five.* Berkeley: University of California Press.

Content, A., Kolinsky, R., Morais, J., & Bertelson, P. (1986). Phonetic segmentation in pre-readers: Effects of corrective information. *Journal of Experimental Child Psychology, 42*, 49-72.

Content, A., Morais, J., Alegria, J., & Bertelson, P. (1982). Accelerating the development of segmentation skills in kindergartens. *Cahiers de Psychologie Cognitive, 2*, 259-269.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design analysis issues for field settings.* Boston: Rand McNally.

Cossu, G., Shankweiler, D., Liberman, I. Y., Katz, L., & Tola, G. (1988). Awareness of phonological segments and reading ability in Italian children. *Applied Psycholinguistics, 9,* 1-16.

Cunningham, A. E. (1988). *A developmental study of instruction in phonemic awareness.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, April 1988.

Cunningham, A. E. (in press). Explicit vs. implicit instruction in phonemic awareness. *Journal of Experimental Child Psychology.*

DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems.* Honolulu: University of Hawaii Press.

Eimas, P. D. (1975). Speech perception in infancy. In L.B. Cohen & P. Salapatek (Eds.). *Infant perception: From sensation to cognition* (Vol. 2). New York: Academic Press.

Eimas, P. D., Sequeland, E. R., Jusczyk, P. W., & Vigorito, T. (1971). Speech perception in infants. *Science, 171,* 303-306.

Eimas, P. D., Miller, J. L., & Jusczyk, P. W. (1987). On infant speech perception and the acquisition of language. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 161-195). Cambridge University Press.

Ehri. L. C. (1979). Linguistic insight: Threshold of reading acquisition. In T. G. Waller & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice.*(Vol 1, pp. 63-114). New York: Academic Press.

Elkonin, D. B. (1963). The psychology of mastering the elements of reading. In B. Simon & J. Simon (Eds.), *Educational psychology in the U.S.S.R.* London: Routledge & Kagan Paul.

Fox, B., & Routh D. K. (1975). Analyzing spoken language into words, syllables and phonemes: A developmental study. *Journal of Psycholinguistic Research, 4,* 331-342.

Golnikoff, R. M. (1978). Phonemic awareness skills and reading achievement. In F. B. Murrary & J. J. Pikulski (Eds.), *The acquisition of reading: Cognitive, linguistic, and perceptual prerequisites.* Baltimore: University Park Press.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read.* East Sussex: Erlbaum.

Halle, M., & Vergnaud, J. (1980). Three dimensional phonology. *Journal of Linguistic Research, 1,* 83-105.

Juel, C., Griffith, P., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in the first and second grade. *Journal of Educational Psychology, 78,* 243-255.

Kirtley, C., Bryant, P., MacLean, M., & Bradley, L. (1989). Rhyme, rime, and the onset of reading. *Journal of Experimental Child Psychology, 48,* 224-245.

Kuhl, P. K. (1987). Perception of speech and sound in early infancy. In A. Salapateck & L. Cohen (Eds.), *Handbook of infant perception: Vol. II, From perception to cognition* (pp. 274-282). New York: Academic Press.

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science, 218,* 1138-1141.

Lenel, J. C., & Cantor, J. H. (1981). Rhyme recognition and phonemic perception in young children. *Journal of Psycholinguistic Research, 10,* 57-68.

Leong, C. K. (1986). The role of language awareness in reading proficiency. In G. T. Pavlidis & D. F. Fisher (Eds.), *Dyslexia: Its neuropsychology and treatment* (pp. 131-148). London: Wiley.

Lewkowicz, N. K. (1980). Phonemic awareness training: What to teach and how to teach it. *Journal of Educational Psychology, 72,* 686-700.

Liberman, A. M. (1989). Reading is hard just because listening is easy. In C. von Euler (Ed.), *Wenner-Gren International Symposium Series: Brain and reading* (pp. 197-205). Hampshire, England: Macmillan.

Liberman, A. M., & Mattingly, I. G. (1989). A specialization for speech perception. *Science, 243,* 489-494.

Liberman, I. Y. (1973). Segmentation of the spoken word and reading acquisition. *Bulletin of the Orton Society, 23,* 65-77.

Liberman, I. Y. & Liberman, A. M. (1990). Whole word vs. code emphasis: Underlying assumptions and their implications for reading instruction. *Bulletin of the Orton Society, 40,* 51-76.

Liberman, I. Y., Shankweiler, D. Fischer, F. W., & Carter, B. J. 1974. Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18,* 201-212.

Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1-34). Ann Arbor: the University of Michigan Press.

Liberman, I. Y., Shankweiler, D., Liberman, A. M., Fowler, C., & Fisher, F. W. (1977). Phonetic segmentation and recording in the beginning reader. In A. S. Reber & D. L. Scarborough (Eds.), *Toward a psychology of reading* (pp. 207-226). Hillsdale, NJ: Erlbaum.

Lindamood, C. H., & Lindamood, P. C. (1969). *The A.D.D. program: Auditory discrimination in depth.* Boston: Teaching Research Corporation.

Lundberg, I., Frost, J., & Peterson, O-P. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly, 23,* 263-284.

Lundberg, I., Olofsson, A., & Wall, S. (1980). Reading and spelling skills in the first school years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology, 21,* 159-173.

MacKay, D. G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology, 3,* 210-227.

Maclean, M., Bryant, P., & Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly, 33,* 255-281.

Mann, V. A. (1984). Longitudinal prediction and prevention of early reading difficulty. *Annals of Dyslexia, 34,* 117-136.

Mann, V. A. (1986). Phonological awareness: The role of reading experience. *Cognition, 24,* 65-92.

de Manrique, A. M. B., & Gramigna, S., (1984). La segmentacion fonologica y silabica en ninos de preescolar y primer grado. *Lectura y Vida 5,* 4-13.

Mattingly, I. G. (1972). Reading, the linguistic process, and linguistic awareness. In J. F. Kavenagh & I. G. Mattingly (Eds.), *Language by Ear and by Eye.* Cambridge Mass: MIT Press.

Mattingly, I. G. & Liberman, A. M. (1990). Speech and other auditory modules. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Signal and sense: Local and global order in perceptual maps* (pp. 501-519). New York: Wiley.

Molfeese, D. L., & Molfeese, D. L. (1979). VOT distinctions in infants: Learned or innate? In H. A. Whitaker, & H. Whitaker (Eds.), *Studies in neurolinguistics* (Vol 4). New York: Academic Press.

Morais, J., Content, A., Bertelson, P., Cary, L., & Kolinsky, R. (1988). Is there a sensitive period for the acquisition of segmental analysis? *Cognitive Neuropsychology, 5,* 347-352.

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.

Morse, P. A. (1972). The discrimination of speech and non-speech stimuli in early infancy. *Journal of Experimental Child Psychology, 14*, 447-492.

Morrison, F. J. (1988). Development of phonemic awareness: A natural experiment. Paper presented at the *Annual Meeting of the Psychonomic Society*, Chicago.

Olofsson, A., & Lundberg, I. (1983). Can phonemic awareness be trained in kindergarten? *Scandinavian Journal of Psychology, 24*, 35-44.

Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. 1987. Phonemic knowledge and to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly, 33*, 283-319.

Read, C. A., & Ruyter, L. (1985). Reading and spelling in adults of low literacy. *Remedial and Special Education, 6*, 43-52.

Read, C. A., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic reading. *Cognition, 24*, 31-44.

Roberts, T. (1975). Skills of analysis and synthesis in the early stages of reading. *British Journal of Educational Psychology, 45*, 3-9.

Rosner, J. & Simon, D. P. (1971). The auditory analysis test: An initial report. *Journal of Learning Disabilities, 4*, 384-392.

Rozin, P., & Gleitman, L. (1977). The structure of the acquisition of reading: II. The reading process and the acquisition of the alphabetic principle. In A. S. Reber and D .L. Scarborough (Eds.), *Toward a psychology of reading* (pp. 55-141). Hillsdale, NJ: Erlbaum.

Share, D. L., Jorm, A. F., MacLean, R., & Matthews, R. (1984). Sources of individual differences in reading achievement. *Journal of Educational Psychology, 76*, 1309-1324.

Stanovitch, K. E., Cunningham, A., & Cramer, B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology, 38*, 175-190.

Torneous, M. (1984). Phonological awareness and reading: Chicken and egg problem? *Journal of Educational Psychology, 76*, 1346-1358.

Treiman, R. (1985). Onset and rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Child Psychology, 39*, 161-181.

Tunmer, W. E., & Nesdale, A. R. (1986). Phonemic segmentation skill and beginning reading. *Journal of Educational Psychology, 77*, 417-427.

Vellutino, F. R., & Scanlon, D. (1987). Phonological coding, phonological awareness, and reading ability: Evidence from a longitudinal and experimental study. *Merrill-Palmer Quarterly, 33*, 321-363.

Wallach, M., & Wallach, L. (1976). *Teaching all children to read.* Chicago: University of Chicago Press.

Williams, J. P. (1980). Teaching decoding with an emphasis on phoneme analysis and phoneme blending. *Journal of Educational Psychology, 72*, 1-15.

Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly, 23*, 159-177.

Zhurova, L. E. (1963). The development of analysis of words into their sounds by preschool children. *Soviet Psychology and Psychiatry, 2*, 17-27.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 12

Can Theories of Word Recognition Remain Stubbornly Nonphonological?

Claudia Carello
M. T. Turvey
University of Connecticut, Storrs

Georgije Lukatela
University of Belgrade

The issue of how readers get from the printed word to its lexical representation is a hotly contested one (see Carr & Pollatsek, 1985; Humphreys & Evett, 1985; Van Orden, Pennington, & Stone, 1990, for reviews). Candidate routes are the visual and the phonological. In the visual route, lexical entries are said to be accessed directly on the basis of orthographic properties. The phonological route requires that lexical access be mediated by the recoding of graphemes into their corresponding phonemes. Considerable experimental data have been offered in support of both types of routes. The bulk of research on word identification using English language materials has been taken to implicate the dominance of a visual access route with, perhaps, an optional but not preferred phonological route (e.g., Coltheart, Besner, Jonasson, & Davelaar, 1979; Humphreys & Evett, 1985). Data on word identification using Serbo-Croatian language materials point unequivocally to a nonoptional phonological access route (e.g., Lukatela & Turvey, 1990a, b; Lukatela, Carello, & Turvey, 1990).

We assume that the basic mechanism of written language processing is the same for all languages. Different data patterns among languages, therefore, are to be taken as evidence of how that mechanism can be fine-tuned by the structure of a particular language. We will use some differences and similarities between Serbo-Croatian, English, and Hebrew to elucidate possible features of a written language processing mechanism that would allow such patterns to arise. Given the nature of the data that have been obtained with Serbo-Croatian, such a mechanism must allow for automatic prelexical phonology. Therefore, we must begin with the assumption that all writing systems are phonological—they provide a system for transcribing phonologically any possible word of the language (A. M. Liberman, in press; Mattingly, 1985). The variety of orthographies do this in more or less straightforward ways, resulting in their being phonologically shallow or deep (I. Y. Liberman, A. M. Liberman, Mattingly, & Shankweiler, 1980). How orthographic depth has been interpreted mechanistically will be addressed. We will ultimately claim that the stubborn rejection of phonology in the prevailing theories of reading cannot be sustained within a consistent theory of language processing that accommodates all of the facts, not just those that are convenient (nor, we might add, just those obtained with English).

Why writing systems must be phonological

As will be shown in later sections, the evidence from experiments in Serbo-Croatian overwhelmingly favors phonological mediation. But our argument begins at a more fundamental level—why that should be an expected outcome. The theoretical backdrop concerns the relation of reading to speech. The basic fact is that phonological structures are the raw materials on which syntactic processes normally work in comprehending speech. These processes are well in place by the time one has to commence the less natural task of learning to read. Given that reading, in contrast to speaking and understanding, is something that must be learned explicitly, how might that be accomplished?

> The seemingly sensible strategy for the reader is to use the optical shapes to access phonological structures early in the reading process. Once the reader has done that, he has put the hard part of reading behind him, for everything else will be done automatically by language processes that he commands by virtue of his humanity (A. M. Liberman, 1991, pp. 242-243).

The alternative to this seemingly sensible strategy is that readers concoct nonlinguistic processes that bar them from the ordinary language processor for as long as possible. The effect is that the route that they take to the lexicon and the base representations that they find there are kept—for whatever reason—"stubbornly nonphonological" (A. M. Liberman, 1991, p. 242).

We suspect, instead, that it is reading theorists rather than readers who remain stubbornly nonphonological, both in denying the plausibility of taking advantage of extant phonological processes and overlooking the phonological basis of writing systems. The ultimate constraint on an orthography is that it permit any *possible* utterance in the language to be transcribed—it must respect the allowable phonological forms of the spoken language (determined by its articulatory gestures and their combinations). We note three aspects of how orthographies accomplish this openness (see Mattingly, 1985, this volume). First, orthographies transcribe linguistic units rather than acoustic or phonetic properties which are too context-sensitive. Second, the linguistic units that are transcribed seem to be words, irrespective of how the graphemic units are framed (cf. Wang, 1981). Third, words are transcribed by exploiting the phonological structure of the language, not by using word-specific symbols. This last point is critical for productivity; it allows a systematic way to transcribe novel utterances.

The morphological and phonological structure of a language determine what form this phonological exploitation takes. Alphabetic (or, more generally, segmental) writing systems are more appropriate for languages that "have fairly elaborate syllable structures, large and rather inefficiently exploited inventories of morphemes, and little homophony" (I. Liberman et al., 1980. p. 149). Syllabic writing systems, in contrast, are more appropriate for languages with a small number of syllables (usually with a regular CVCV... structure without consonant clusters). It should be noted that Chinese, despite its reputation in the folklore of orthographies, reflects phonological constraints as well. Its mislabeling as pictographic or ideographic has more to do with socio-cultural agendas than with what is represented by its orthography, namely, syllables (Mattingly, this volume). Logograms are a secondary accompaniment as they must be in a productive, complete writing system (DeFrancis, 1989; Mattingly, this volume; Wang, 1981).

The orthographic depth hypothesis

Serbo-Croatian, English and Hebrew differ in how straightforwardly their orthographies transcribe the sounds of the spoken language. In Serbo-Croatian, a grapheme such as G is pronounced /g/ regardless of the context in which it appears. There are no irregular pronunciations, silent letters, doubled letters, and so on. In English, G might be pronounced /g/, /j/, or /zh/, or not pronounced at all, depending on whatever else is in the letter string. In Hebrew, vowels are not even represented in 90% of the written material that adults encounter. Homographs are common; the pronunciation of an isolated word depends on which vowels are elected by a reader. This kind of difference has been referred to as orthographic depth[1] (Frost, Katz, & Bentin, 1987; I.Liberman et al., 1980; Lukatela, Popadić, Ognjenović, & Turvey, 1980; Sebastián-Gallés, 1991).

Orthographic depth has been considered relevant to reading because it seems to imply that getting from script to sound is more or less dependable for different languages and, therefore, should be more or less apparent in reading processes. While more or less dependable is fairly well agreed upon, more or less apparent has been subject to some interpretation which we would like to clarify in the context of our recently developed network formulation of Serbo-Croatian word recognition. The theme is that the easier it is to "get to" the sound from the spelling the more likely the reader is to do so in ordinary reading, using that as a basis for getting to other things as well—such as the lexicon. To some, this suggests that readers of a phonologically shallow orthography will access the lexicon phonologically whereas readers of a deep orthography will access the lexicon visually. The reasoning concerns how efficiently articulatory codes are provided. If the translation is complex and takes a long time, they won't be used (e.g., Frost et al., (1987).

Although not formulated with orthographic depth in mind, dual route theories are consistent with this reasoning. A phonological route to the lexicon is not used in English, it is argued, because the irregularity of script-to-sound makes the translation take too long; visual access, in contrast, is achieved rapidly. A phonological influence will be felt only for those letter strings for which visual access is slowed. This would include nonwords and, perhaps, low frequency words. Under a dual route interpretation, a phonological route to the lexicon would be impossible in Hebrew because the letter strings are so ambiguous.

More recently we have tried to take care in referring to *how apparent* the involvement of phonology is in accessing the lexicon as opposed to *whether or not* it is involved. We are trying to finesse two issues here. One issue has to do with the ease of demonstrating phonological involvement in Serbo-Croatian due to particular methodological advantages (versus processing differences between Serbo-Croatian and deeper orthographies). As noted in detail elsewhere, Serbo-Croatian is not only shallow, it is shallow in two largely distinct but partially overlapping scripts. The nature of the overlap is such that some letters are pronounced the same in the two alphabets while others are pronounced differently depending on which alphabet the reader uses. This allows the construction of letter strings in which a host of properties (semantics, syntax, frequency, associative relatedness) can be controlled experimentally while distinguishing graphemic from phonemic similarity. If experiments in English or Hebrew could be similarly contrived, they might, in principle, show unequivocal phonological involvement as well.

[1]Orthographic depth, in fact, has a second aspect and that is the relative remoteness of the phonetic representation from the morphophonological representation (I. Liberman et al., 1980). Experimental investigations that deal with orthographic depth tend to focus only on how easily the orthography approximates the phonetic representation.

The second issue concerns the mechanistic interpretation of orthographic depth. The tradition has been to couch it in terms of discrete grapheme-phoneme correspondence rules or GPCs. More recent models, such as those pioneered by McClelland and Rumelhart (1986) and envisioned by Van Orden et al. (1990), could accommodate (in principle) orthographic depth with respect to the strength and number of connections in a parallel distributed network. Distinctions between these two kinds of approaches will be considered in some detail before turning to experimental demonstrations of phonological involvement in word recognition.

Mechanistic interpretations of orthographic depth

Grapheme-phoneme correspondence rules constitute the more familiar treatment of how one gets from spelling to sound (e.g., Coltheart, 1977, 1978). They specify how particular letters or clusters of letters are to be pronounced. And they do so discretely; the rules do not vary in strength. Thus, under this treatment, a shallow orthography is one that has relatively few rules and whose words can be relied upon to follow them (e.g., in Serbo-Croatian, the rules are defined at the level of the individual grapheme because their pronunciation is not changed by being combined with different combinations of graphemes). A deep orthography may have numerous rules or exceptions to its rules (e.g., in English, when a word ends in E, the E is silent and the preceding vowel is long; CAVE obeys this but HAVE does not) or, perhaps, application of its rules is simply inadequate to allow a reader to settle on a single pronunciation (e.g., in Hebrew, the standard printed form omits vowel marks so that a particular letter string can be pronounced as different words depending on which vowels are elected). However reliable they are, GPCs more or less embody what is orthographically legal in a given language. (This fact is responsible for the easy link between GPCs and pseudowords: GPCs may be useful at least insofar as they permit one to pronounce a novel letter string.)

Under the discrete symbol, rule-based characterization of assembling phonology, it is possible to consider that readers of different kinds of orthographies are engaging in different kinds of processes. Those for whom GPCs are reliable would be well-served to try them since a straightforward translation might be faster than a lexical search for a visual match to the orthographic pattern. But those for whom GPCs are unreliable or inadequate might be forced to be visual readers since using the rules would be slow and error prone. Visual readers, then, would be engaged in a search for a word-specific match in the lexicon.[2] Novel letter strings might allow them to apply GPCs but don't require it; a pronunciation can be generated by (visual) analogy to a real word (e.g., Glushko, 1979; Kay & Marcel, 1981).

But the possibility of a persisting visual route (i.e., all the way to the lexicon) ignores the phonological foundation of writing systems and delays the reader's tapping into the language machinery set up to understand speech. Simply on logical grounds, it is unsavory. Rather than considering that orthographic depth contributes to the formation of different language processing devices, one that is rule-based and one that is word-specific, let us consider that orthographic depth serves to modulate the same basic device. A parallel distributed network is a candidate device amenable to such modification. The network consists of successive, connected layers of subsymbolic nodes (see Lukatela,

[2] Van Orden et al. (1990) point out that the way in which the debate has been framed has had the insidious effect of turning psycholinguistics into what it had originally criticized: An account of verbal behavior rooted in specific stimulus-response connections.

Feldman, Turvey, Carello, & Katz, 1989, and Lukatela Turvey, Feldman, Carello, & Katz, 1989, for detailed descriptions of the network for Serbo-Croatian). The connections reflect covariation (e.g., between orthographic features and phonological features) and statistical regularity (they are weighted) rather than discrete rules. In other words, they are modifiable. Orthographic depth modulates this device with respect to the number and strength of these connections. Processing differences among languages, then, are not so much qualitative as quantitative: In Serbo-Croatian, the letter-phoneme connections are few and strong; in English, they are many and weak. The ultimate effect on response time, of course, may be qualitative because the pattern of activation emerges from a dynamical system in which linear changes can have nonlinear consequences.

This interpretation of orthographic depth means that there is only one kind of processing. Indeed, parallel distributed networks destroy the basis for considering "routes" to the lexicon as if they were independent pathways with no mingling of activation. At bottom, they allow us to reject the dual route model altogether, including its logic for inferring nonphonology.

The logic of inferring nonphonology in the absence of a dual route theory

In the logic of dual-route theory, lexical context effects are thought to undermine the case for assembled phonology. Since rules are discrete, not graded, things like frequency should not matter: K should be pronounced /k/ whether it appears in KICK or KALE. If the higher frequency word is pronounced faster, it must be because its visual form is more familiar. But in an interactive network, the lower threshold of high frequency word units means that they would be activated sooner by the pattern of excitation arising through the phoneme unit level. Indeed, with communication between levels, many lexical properties can be expected to influence pronunciation. Automatic involvement of the lexicon is inevitable given the bi-alphabetic nature of Serbo-Croatian. The presence of phonologically ambiguous letters generates activity along two letter-phoneme connections. If there are, say, two phonologically ambiguous letters in a four letter word, four pronunciations of that letter string are assembled. Each gives rise to some activation at the word unit level depending on how closely the phonemes match the word unit (with respect to number and order) and on the word units' frequencies. Activation of certain word units is strengthened by interaction between word and phoneme levels and continues until above-threshold activation of a single word unit emerges. Of course, this interactive processing is not limited to phonologically ambiguous words; it is characteristic of the ordinary language processor. Phonologically unique letter strings generate a single code but it partially activates a number of word units. Although a single word unit emerges quickly from interaction between word and phoneme levels, interactive processes nonetheless provide the opportunity for lexical influences on pronunciations assembled on the basis of prelexical phonology. That is to say, phonological codes are assembled by the prelexical phonological connections but a single pronunciation is settled on out of the global pattern of activation. Lexical involvement does not contravene prelexical phonology (Carello, Lukatela, & Turvey, under review).

Relatedly, we argue that the distinction between assembled and accessed phonology has been cut too sharply, as if phonological information came only from one source or the other. To arrive at the lexicon phonologically does not mean that the assembled code carries every phonological nuance. Linguistic features such as stress and prosody must be derived from the phonological representation in the lexicon which has itself been accessed via prelexical phonological connections. For example, the stress pattern of Serbo-Croatian words—which is not marked in the orthography—can be rising or falling and long or short. In addition, although the first syllable is usually stressed, occasionally the second syllable is stressed instead. A correct pronunciation requires information about the stress pattern which can only be had at the word unit level. But information about the stress pattern is only made available once the word unit has been activated by the phonological code. That is to say, the existence of accessed phonology does not contravene prelexical phonology (see Lukatela & Turvey, 1990, for experimental ramifications of differing stress patterns between contexts and targets).

The logic of inferring phonology

The case for prelexical phonology is, at the very least, not undercut by the existence of lexical context effects. But what kind of evidence would make the case for prelexical phonology? If we invert the logic that has been established by those advocating primacy of the word-specific visual route, we can look for several things. Phonological influences should be observed on acceptance latencies (which are, by and large, faster than rejection latencies), especially on high frequency words. Such evidence would support the claim that the influence is felt in ordinary word recognition, not just for letter strings that have no lexical entries (cf. Coltheart, Davelaar, Jonasson, & Besner, 1977; Kay & Marcel, 1981). Phonological effects should be apparent in both naming and lexical decision. Naming is important because it is supposed to be free of post-lexical influences (Balota & Chumbley, 1985; West & Stanovich, 1982). The effect ought to depend on the number of constituents (letters, syllables) in a word. A phonologically analytic process would reflect the burden of decoding details of the orthographic structure (Green & Shallice, 1976). Phonological effects should persist in the face of experimental conditions that discourage the use of prelexical phonology. Strategic insensitivity would suggest that the phonological route is nonoptional (cf. Hawkins, Reicher, Rogers, & Peterson, 1976). Finally, phonological effects must appear over and above effects due to graphemic similarity. Orthographically similar rhyming items should behave the same as orthographically dissimilar rhyming items but different from orthographically similar nonrhyming items (cf. Evett & Humphreys, 1981).

The case for prelexical phonology in Serbo-Croatian

The case for prelexical phonology has been made on each of these points using the Serbo-Croatian language. These results can be organized around three general manipulations permitted by exploiting the two alphabets: (1) comparisons between phonologically unique letter strings, composed exclusively of unique and common letters, and phonologically ambiguous letter strings, composed exclusively of common and ambiguous letters; (2) comparisons of phonemically and graphemically similar pairs, written in the same alphabet, and phonemically similar but graphemically dissimilar pairs, with the context and target written in different alphabets; and (3) comparisons of phonologically ambiguous pseudowords in which a mixed interpretation of the letters in a single letter string either is or is not a word.

Manipulations of the first type produce the so-called Phonological Ambiguity Effect—letter strings with more than one phonological interpretation are associated with longer latencies and higher errors than letter strings with only one phonological interpretation *even though they are the same words.* For example, VETAR and BETAP are Roman and Cyrillic, respectively, for "the wind." VETAR has one phonological interpretation, /vetar/, because V and R are uniquely Roman letters and E, T, and A are common. BETAP, in contrast, has four phonological interpretations because B and P can be read (differently) in Roman and Cyrillic. The Cyrillic interpretation of both yields /vetar/. The Phonological Ambiguity Effect occurs in naming and lexical decision (e.g., Lukatela, Feldman et al., 1989; Lukatela, Turvey et al., 1989), is larger for words (independent of frequency) than pseudowords (e.g., Feldman & Turvey, 1983; Lukatela, Feldman, et al., 1989), increases with more phonologically ambiguous letters (Feldman, Kostić, Lukatela, & Turvey, 1983; Feldman & Turvey, 1983), decreases with more unique letters (Lukatela, Feldman, et al., 1989), and persists despite instructions (Lukatela, Savić, Gligorijević, Ognjenović, & Turvey, 1978) or experience favoring one alphabet (Feldman & Turvey, 1983) or discouraging phonological coding (Lukatela, Feldman, et al., 1989).

Manipulations of the second type produce phonemic similarity effects. Naming latencies to the word target PUŽIĆ (/puzhich/) and the pseudoword target PUDIĆ (/pudich/) are facilitated to the same degree by phonemically similar contexts, whether those are graphemically similar (PUTIĆ, /putich/) or dissimilar (ПУТИЋ, /putich/) (Lukatela & Turvey, 1990a). For lexical decision latencies, the direction of the phonemic similarity effect depends on target frequency, the ordinal position of the distinguishing phoneme, and lexicality. Phonemic similarity effects persist even when the context is masked (for both word-pseudoword and pseudoword-word sequences, Lukatela & Turvey, 1990a), and when graphemic similarity is further reduced by writing contexts in lower case and targets in upper case, for example, pasus-ПАСУЉ, /pasus-pasulj/ (Lukatela et al., 1990). Finally, target identification under conditions of backward masking—a target followed by a pseudoword mask which is itself followed by a pattern mask—is enhanced when the pseudoword mask is phonologically similar to the target (Lukatela & Turvey, 1990b). The mask presumably continues activation at the phoneme unit level that had been initiated by the target (Naish, 1980; Perfetti, Bell, & Delaney, 1988).

Manipulations of the third type produce "virtual word" effects. BEMAP and HAPEM both differ from a real word by one letter (BETAP and XAPEM, respectively). B, P, and H have different interpretations in Roman and Cyrillic so that a phonologically analytic processing of each string would produce four codes. For BEMAP none of these is a word, whereas for HAPEM one is a word. HAPEM-type strings produce a much larger false positive error rate: 30% vs. 3%. When a HAPEM-type follows a context associatively related to the virtual word interpretation, false positives increase to 55%, compared to 7% for BEMAP-types following associates of their source words (Lukatela, Feldman, et al., 1989; Lukatela, Turvey, et al., 1989). In naming, the mixed alphabet (virtual word) interpretation of HAPEM-type strings occurred 3-4 times more often than the mixed alphabet interpretation of BEMAP-types. These differences arise even though the two types of pseudowords are equally similar visually to a real word—they differ by one letter. But whereas every code for BEMAP is also one phoneme different from a real word, one code for HAPEM shares all phonemes with a real word. Virtual word effects derive from prelexical phonology.

For Serbo-Croatian, in sum, the requisite patterns of results have been obtained to allow the conclusion of prelexical phonology. Phonological involvement has been demonstrated on "yes" responses, with high frequency words, on words more than pseudowords, in naming as well as lexical decision; it is sensitive to the number of constituents, and persists despite experimental conditions that might discourage it; finally, phonological effects are independent of graphemic effects which, in fact, do not occur. The results from English and Hebrew do not permit quite the same point by point confirmation. But there are what we might consider "existence proofs" for a number of them.

The case for prelexical phonology in English

The supporting English results can be organized around four general manipulations, the first three of which exploit the deep orthography: (1) comparisons between pseudohomophones, nonwords that are pronounced the same as real words but spelled differently, and spelling controls, nonwords that differ from the targets by the same number of letters as the pseudohomophone but are pronounced differently; (2) comparisons between homophones, words that are pronounced the same as target words but spelled differently, and spelling controls; (3) comparisons of phonologically consistent pairs in which a given stem receives the same phonological interpretation, and phonologically inconsistent pairs in which a given stem receives different phonological interpretations; and (4) comparisons of phonemically and graphemically similar pairs, phonemically similar and graphemically dissimilar pairs, and graphemically similar but phonemically dissimilar pairs.

Manipulations of the first type provided some of the earliest suggestions of phonological involvement in lexical decision (e.g., Rubenstein, Lewis, & Rubenstein, 1971). But since this effect was on "no" responses which are already slow, delayed rejections of pseudohomophones was soon interpreted as implicating phonological involvement only when the direct route hadn't worked fast enough (see Van Orden et al., 1990, for a rebuttal of the logic behind the so-called "delayed phonology hypothesis"). Not prone to such an indictment are recent experiments showing associative priming by and of pseudohomophones: TABLE facilitated the naming of the pseudohomophone CHARE relative to the spelling control CHARK; the pseudohomophone prime TAYBLE facilitated the naming of CHAIR relative to the spelling control prime TARBLE (Lukatela & Turvey, 1991). In both of these instances, in order for the associative relationship to have had an effect, the lexicon must have been accessed and it must have been accessed through phonology. In four experiments, the graphemic control did not produce a significant effect (and the numerical difference was always in the wrong direction). In contrast, TAYBLE did not differ from TABLE in its effect on naming CHAIR. Moreover, the word targets (and source words of the pseudohomophones) were of relatively high frequency (217 according to the norms of Francis & Kučera, 1982).

Other experiments have demonstrated additional dimensions of equivalency in the processing of pseudohomophones and their real word counterparts (Lukatela & Turvey, in press). Between the presentation and recall of one or five digits, subjects performed a secondary task of naming a visually presented letter string—a pseudohomophone (e.g., FOLE, HOAP) or its lexical counterpart (FOAL, HOPE). If nonwords are named by a slow (resource expensive) process that assembles the letter string's phonology and words are named by a fast (resource inexpensive) process that accesses lexical phonology (see Paap & Noel, 1991), then memory load should interact with lexicality (HOPE vs. HOAP, FOAL vs. FOLE). To the contrary, three experiments found that load interacted only with frequency (HOPE vs. FOAL, HOAP vs. FOLE), suggesting that pseudohomophones and their word counterparts are processed similarly, namely, phonologically. An example of the form of the interaction is shown in Figure 1. In a fourth experiment the associative priming-of-naming task described above was secondary to the memory task.

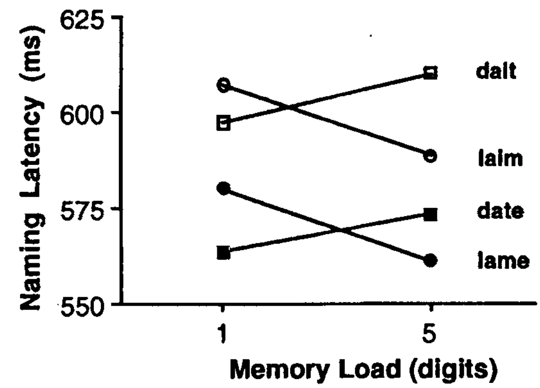

Figure 1. **High frequency words and their pseudohomophones (closed and open squares, respectively) are hindered by increased memory load. The opposite pattern is obtained with low frequency words and their pseudohomophones (closed and open circles, respectively). That is, words are more similar to their nonlexical but phonologically identical counterparts than they are to each other.**

In elaboration of Lukatela and Turvey's (1991) observations, associative priming (HOPE-DESPAIR, FOAL-HORSE) was equaled by pseudohomophone associative priming (HOAP-DESPAIR, FOLE-HORSE) with memory load affecting both kinds of priming in the same way.

Manipulations of the second type show homophony effects on rejection latencies, this time in semantic categorization tasks: BEATS takes longer to reject as a member of the category VEGETABLE than do other foils (e.g., Meyer & Ruddy, 1973). Finer analyses, however, reveal homophony to be influential on faster yes responses as well: The false positive error rate is higher for homophones than for spelling controls (18.5% vs. 3.0%, Van Orden, 1987) and the false positive "yes" latencies are comparable to the correct "yes" latencies (Van Orden, Johnston, & Hale, 1988). Moreover, although orthographic similarity of homophones (BEATS is more like BEETS than ROWS is like ROSE) matters under unmasked conditions, the orthographic effect disappears when targets are pattern-masked while the homophony effect remains strong (Van Orden, 1987). Van Orden argues that this supports the role of phonological mediation as an early source of constraint on word identification (Van Orden, 1987; Van Orden et al., 1990).

Manipulations of the third type show differences in priming effects between graphemically similar pairs that are also phonologically similar (BRIBE-TRIBE) and graphemically similar pairs that are phonologically dissimilar (TOUCH-COUCH). Generally, phonological consistency is beneficial and phonological inconsistency is detrimental (Hanson & Fowler, 1987; Meyer, Schvaneveldt, & Ruddy, 1974). Where there are priming effects for both types of pairs, the effect with phonologically similar pairs is greater (Hanson & Fowler, 1987). Even the results of Evett and Humphreys (1981), who did not find differences due to consistency when the primes were masked, have been interpreted as supportive of phonological mediation by "noisy phonologic codes" (Van Orden et al., 1990, p. 495). The epithet noisy is applied on the assumption that TOUCH would give rise to a code that had elements of both /tutch/ and /towtch/.[3] Therefore, priming of COUCH by TOUCH is, in fact, a phonological effect. Van Orden (1987; Van Orden et al., 1990) argues further that sometimes noisy codes are sufficient to distinguish words from nonwords (e.g., when the pseudoword foils are illegal nonwords), in which case there would be no advantage for phonologically consistent pairs. Phonological inconsistency will be detrimental when noisy codes must be cleaned up, viz., for foils that are legal nonwords. These are the results reported by Shulman, Hornak, and Sanders (1978) and Hanson and Fowler (1987). Interestingly, detrimental phonological inconsistency effects—those historically taken to demonstrate phonological mediation—are most likely under experimental conditions that ought to discourage phonology were it optional (Van Orden et al., 1990). That is to say, with legal nonword foils, words would be better distinguished by a graphemic code were it an option.

Manipulations of the fourth type have produced inconsistent results. While facilitation for phonemically similar, graphemically dissimilar pairs has been reported (Hillinger, 1980), this has not been replicated, either in lexical decision (Martin & Jensen, 1988) or naming (Peter, Turvey, & Lukatela, 1990). But graphemic priming was not found either. As an important aside, we note that this latter result would appear to be in sharp contradiction of the major expectation from the hypothesized visual, word-specific route. If lexical items are coded visually (more precisely, orthographically), then preceding words that are visually similar to immediately subsequent words should facilitate decisions on the immediately subsequent words. That such visually based facilitation is difficult to obtain (ordinarily investigators have to impose a number of additional manipulations, such as severe forward masking of the prime, to reveal slight effects [e.g., Forster, 1987]) should be taken as prima facie evidence that visual access is neither prominent nor particularly straightforward. Curiously, proponents of the visual, word-specific route have been mute on this failure to prime the lexicon visually.

More reliable than the results from forward phonemic priming are results from masked backward priming (a target followed by a pseudoword mask which is itself followed by a pattern mask): Targets are more likely to be identified when the pseudoword mask is phonemically rather than graphemically similar to it (Naish, 1980; Perfetti et al., 1988). Manipulations of this fourth type can be combined with those of the second type. ROWS is more likely to be recognized as a member of the category FLOWER when followed by a phonemically similar rather than graphemically similar pseudoword mask (Peter & Turvey, 1992).

In sum, the results for English are accumulating to allow the conclusion of prelexical phonology. Phonological involvement has been demonstrated on "yes" responses, with high frequency words, and in naming as well as lexical decision; it has occurred despite experimental conditions that might discourage it; and, finally, phonological effects have been obtained that are over and above graphemic effects which are, in fact, unreliable.

The case for prelexical phonology in Hebrew

Our assertion that the underlying processing is the same across languages requires that there be at least some evidence of prelexical phonology in the deepest orthographies. The phoneme layer still exists even though the letter to phoneme connections might be multiple and very weak. Support for phonological involvement in Hebrew comes from

[3]This is not unlike what we have proposed for phonologically ambiguous letter strings in Serbo-Croatian—all possible pronunciations of a string are generated before one is settled on through competitive processes.

two general manipulations that exploit the fact that vowels are not represented in ordinary text: (1) comparisons of pointed and unpointed letter strings, and (2) comparisons of phonologically ambiguous and unambiguous words.

Manipulations of the first type provided the earliest suggestion of phonological mediation in Hebrew. For some consonant strings, there is only one phonological interpretation with a single lexical entry. Adding the proper vowels redundantly specifies the same pronunciation. Adding certain incorrect vowels specifies other particular pronunciations that are phonotactically legal even though they are without a lexical entry. Adding other incorrect vowels that are allophonic with the correct vowels will specify the correct pronunciation even though that orthographic pattern has no lexical counterpart (it is a pseudohomophone). Navon and Shimron (1981) found that allophonically voweled letter strings (essentially pseudohomophones) did not differ in naming time from ordinary unpointed or correctly pointed letter strings. That is, the correct phonological interpretation accessed its lexical entry even though its orthographic form was novel. Naming was slower when the added vowels specified a pronunciation without a lexical entry. More recently, it has been shown that readers will wait for the vowel marks in a delayed presentation paradigm (consonant string followed at some lag by the diacriticals) even though the orthographic form has only one lexical entry (Frost, 1992). This was true for both high and low frequency words in both lexical decision and naming.

Manipulations of the second type are somewhat similar to manipulations of phonological ambiguity in Serbo-Croatian in that a given letter string can be pronounced in more than one way. In this case, the phonological options come not from choice of alphabet but from choice of vowels to assign to an unpointed letter string. Here we consider only those pronunciations that constitute words (rather than all pronunciations that might be generated by the random assignment of vowels). Consonant strings with three or more phonemic realizations are named more slowly than consonant strings with only one (Bentin, Bargai, & Katz, 1984). When semantic priming contexts are consonant strings with two phonemic realizations and two meanings, one a high frequency word and one a low frequency word, lexical decision is facilitated more by the phonological interpretation associated with the higher frequency word (Frost & Bentin, 1992). When these same letter strings are pointed (and, therefore, phonologically unambiguous), the amount of facilitation by the low and high frequency versions is the same. Relatedly, contexts with both a high and low frequency meaning but with a single phonological interpretation (like the English word RUN, for example) also produce equivalent facilitation in targets semantically related to either of the two meanings. Taken together, these findings suggest that the ambiguity effect found with heterophonic homographs is phonological rather than semantic in origin (Frost & Bentin, 1992). Delaying the onset of vowel marks after the presentation of ambiguous letter strings with two phonemic realizations slows the naming of words (both high and low frequency) and pseudowords equally (Frost, 1992). This lag effect is larger than that for unambiguous words.

The results for Hebrew suggest at least some involvement of prelexical phonology. It has been demonstrated on "yes" responses, with high frequency words, and in naming as well as lexical decision; it has occurred despite experimental conditions that do not require it; and one phonological effect has been obtained that is over and above a graphemic effect. But the body of data from Hebrew are equivocal, perhaps epitomized by the fact that lexical decision to targets either orthographically or phonemically similar to pseudoword primes are facilitated to the same degree (Bentin et al., 1984).

Nonetheless, the extent of parallel evidence in Serbo-Croatian, English, and Hebrew is impressive. The script-sound relationships in the three languages constitute very different experimental settings. Some of the classes of experiments that we have discussed are not possible in the other language. English and Hebrew have no mixed alphabets; Serbo-Croatian has no phonological inconsistency. For the most part, the differences favor Serbo-Croatian as a vehicle for demonstrating prelexical phonology (Lukatela et al., 1990; Lukatela & Turvey, 1990 a, b). Despite these differences, early nonoptional phonological involvement is apparent in all. Differences that remain are arguably due to differences in covariant learning particularly with respect to letter-phoneme connections.

Concluding remarks: The primacy of phonological "dynamics"

We have chosen to build our arguments for reading's natural phonological basis around a hypothesis of prelexical phonology as primary. Roughly interpreted, this hypothesis is that processes intimately connected to those by which speech is produced and perceived constitute the major constraint on the mapping from print to lexicon. The now classic dual-route theory has provided a fairly simple (and empirically fruitful) framework within which to deliberate how a person's knowledge about words might be tapped by letter strings: It is tapped either by the letter strings described in the predicates of the visual system, or by letter strings described in the predicates of the speech system, or both. As the theory tends to go, the visual predicates are more prominent than the speech predicates. Our arguments in this chapter were phrased very much in the context of the dual-route theory, and in reaction to the proposed primacy of visual predicates. The strategy we adopted was chosen because, in many respects, it is the most convenient and the most conducive to communication (relying as it does upon the most conventional understanding). In these final remarks, however, we would like to take a more critical and circumspect stance. We explore the implications of a continuous dynamical perspective on word-recognition processes, the perspective adumbrated in much of the foregoing criticism of the "stubbornly nonphonological" accounts.

Our departure point is an assertion: Learning to read is largely an autonomous process. By this assertion we intend to mean several things. First, reading is achieved by a system capable of attuning to mappings between orthographic and linguistic structures, however arbitrarily complex those mappings might happen to be (that is, it does not require that the mappings be orthogonal or linearly separable). Second, the structures mapped between are characterized by distinguishable features or substructures at many grain sizes; there is, however, no biasing of the system toward any particular grain size. Consequently, attunement may occur to mappings that vary considerably in the sizes of the substructures comprising their domains and codomains. Third, the system's attunement is eventually most pronounced (but not exclusively restricted) to the mappings significant to reading without having to be informed explicitly as to what those particular significant mappings might be. Fourth, the enhanced attunement to reading-significant mappings follows from a generic selection principle: Those mappings are selected that are single-valued, or most nearly so. That is, the more invariant the relation between particular substructures of the orthography and particular linguistic substructures, the more likely is it that that mapping will be selectively enhanced.

In dynamical terms, what are the consequences of invariance—of single-valuedness? An approximate answer, one highlighted by Van Orden et al. (1990), is that resonance or self-consistency is achieved rapidly within the connective matrix binding (the processing units of) the domain's and codomain's substructures. Borrowing from adaptive resonance theory (Grossberg, 1987; Grossberg & Stone, 1986), a resonant mode is achieved when the activity excited in a given layer of processing units from below matches that excited from above. A closely related answer is that the pattern of activity engendered in the network instantiation of the mapping is stable. Consequently, where a mapping deviates from single-valuedness, the time course of achieving resonance is slower and/or the final-state stability is less.

In most languages, if not all, the invariance is greatest between orthography and phonology, roughly speaking, between the spellings of words and the names of words. Patently, the mappings between orthography and the meanings of words, and orthography and the syntactic functions of words, are considerably less consistent. Phonological representations will, therefore, achieve resonance faster, and reach states of stability greater, than other linguistic representations. Again, in terms of adaptive resonance theory, a greater match is achieved, and achieved at a more rapid pace, between the activity patterns in the phonological layer excited by (a) the lexical layer above, and (b) the graphemic layer below. The upshot is that even if many activations of linguistic substructures by orthographic substructures occur concurrently in word recognition, it is the phonological activation that stabilizes earliest, providing a basis for stabilizing the other patterns of linguistic activation (Van Orden et al., 1990).

In these final remarks we have pursued a line of argument constrained by the notions of autonomy and invariant. We have been led to conclude that, in word recognition, the dynamics associated with phonological processes are primary. It will be interesting to see in what directions a theory grounded in dynamics might evolve (along the lines, perhaps, of recent efforts in movement coordination, e.g., Kugler & Turvey, 1987; Schmidt, Beek, Treffner, & Turvey, 1991; Turvey, 1990; Turvey, Schmidt, & Beek, in press) and the kinds of experimental hypotheses to which it might give rise. A benchmark for evaluating such a theory's worth is the dual process theory, which has been the dominant source of stimulation for research on word recognition in recent times. Will a dynamically based theory be as fruitful?

Acknowledgment
We wish to dedicate this chapter to our dear friend and colleague Alvin Liberman who has provided the occasional gentle reminder that reading and speech must be intimately connected. Preparation of this manuscript was supported in part by National Institutes of Child Health and Human Development Grants HD-01994 and HD-08495 to Haskins Laboratories and the University of Belgrade, respectively.

References

Balota, D. A., & Chumbley, J. I. (1985). The locus of word frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language, 24*, 89-106.

Bentin, S., Bargai, N., & Katz, L. (1984). Orthographic and phonemic coding for lexical access: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 353-368.

Carello, C., Lukatela, G., & Turvey, M. T. (under review). Lexical involvement in naming does not contravene prelexical phonology: Comment on Sebastián-Gallés.

Carr, T. H., & Pollatsek, A. (1985). Recognizing printed words: A look at current models. In D. Besner, T. G. Waller, & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 1-82). Orlando, FL: Academic Press.

Coltheart, M. (1977). Critical notice of Gibson, E. J. and Levin, H., "The psychology of reading." *Quarterly Journal of Psychology, 29*, 157-167.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). London: Academic Press.

Coltheart, M., Besner, D., Jonasson, J. T., & Davelaar, E. (1979). Phonological encoding in the lexical decision task. *Quarterly Journal of Experimental Psychology, 31*, 489-507.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535-555). New York: Academic Press.

DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems.* Honolulu: University of Hawaii Press.

Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology, 33*, 325-350.

Feldman, L. B., Kostić, A., Lukatela, G., & Turvey, M. T. (1983). An evaluation of the "basic orthographic syllable structure" in a phonologically shallow orthography. *Psychological Research, 45*, 55-72.

Feldman, L. B., & Turvey, M. T. (1983). Word recognition in Serbo-Croatian is phonologically analytic. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 288-298.

Forster, K. (1987). Form-priming with masked primes: The best match hypothesis. In M. Coltheart (Ed.), *The psychology of reading: Attention and performance XII* (pp. 127-146). Hillsdale, NJ: Lawrence Erlbaum Associates.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage.* Boston: Houghton Mifflin.

Frost, R. (1992). Prelexical and postlexical strategies in reading: Evidence from a deep and a shallow orthography. Paper presented at the Fifth Annual Meeting of the European Society for Cognitive Psychology. Paris, September, 1992.

Frost, R., & Bentin, S. (1992). Processing phonological and semantic ambiguity: Evidence from semantic priming at different SOAs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 58-68.

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.

Glushko, R. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 361-379

Green, D. W., & Shallice, T. (1976). Direct visual access in reading for meaning. *Memory & Cognition, 4*, 753-758.

Grossberg, S. (1987). Competitive learning: from interactive activation to adaptive resonance. *Cognitive Science, 11*, 23-63.

Grossberg, S., & Stone, G. (1986). Neural dynamics of word recognition and recall: Priming, learning, and resonance. *Psychological Review, 93*, 46-74.

Hanson, V. L., & Fowler, C. A. (1987). Phonological decoding in word reading: Evidence from hearing and deaf readers. *Memory & Cognition, 15*, 199-207.

Hawkins, H. L., Reicher, G. M., Rogers, M., & Peterson, L. (1976). Flexible coding in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 380-385.

Hillinger, M. L. (1980). Priming effect with phonemically similar words: The encoding bias hypothesis reconsidered. *Memory & Cognition, 8,* 115-123.

Humphreys, G. W., & Evett, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual route theory of reading. *The Behavioral and Brain Sciences, 8,* 689-739.

Kay, J., & Marcel, A. (1981). One process, not two, in reading aloud: Lexical analogies do the work of nonlexical rules. *Quarterly Journal of Experimental Psychology, 33A,* 397-413.

Kugler, P. N., & Turvey, M. T. (1987). *Information, natural law, and the self-assembly of rhythmic movement.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Liberman, A. M. (1991). Observations from the sidelines. *Reading and Writing: An Interdisciplinary Journal, 3,* 429-433.

Liberman, A. M. (in press). The relation of speech to reading and writing. To appear in B. de Gelder & J. Morais (Eds.), *Language and literacy: Comparative approaches.* Cambridge, MA: MIT Press.

Liberman, I. Y., Liberman, A. M., Mattingly, I., & Shankweiler, D. (1980). Orthography and the beginning reader. In J. F. Kavanagh & R. Venezky (Eds.), *Orthography, reading, and dyslexia* (pp. 137-153). Baltimore, MD: University Park Press.

Lukatela, G., Carello, C., & Turvey, M. T. (1990). Phonemic priming by words and pseudowords. *European Journal of Cognitive Psychology, 2,* 375-394.

Lukatela, G., Feldman, L. B., Turvey, M. T., Carello, C., & Katz, L. (1989). Context effects in bi-alphabetical word perception. *Journal of Memory and Language, 28,* 214-236.

Lukatela, G., Popadić, D., Ognjenović, P., & Turvey, M. T. (1980). Lexical decision in a phonologically shallow orthography. *Memory & Cognition, 8,* 124-132.

Lukatela, G., Savić, M., Gligorijević, B., Ognjenović, P., & Turvey, M. T. (1978). Bi-alphabetical lexical decision. *Language and Speech, 21,* 142-165.

Lukatela, G., & Turvey, M. T. (1990a). Phonemic similarity effects and prelexical phonology. *Memory & Cognition, 18,* 128-152.

Lukatela, G., & Turvey, M. T. (1990b). Automatic and prelexical computation of phonology in visual word identification. *European Journal of Cognitive Psychology,* 2, 325-343.

Lukatela, G., & Turvey, M. T. (in press). Similar attentional, frequency, and associative effects for pseudohomophones and words. *Journal of Experimental Psychology: Human Perception and Performance.*

Lukatela, G., & Turvey, M. T. (1991). Phonological access of the lexicon: Evidence from associative priming with pseudohomophones. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 951-966.

Lukatela, G., Turvey, M. T., Feldman, L. B., Carello, C., & Katz, L. (1989). Alphabetic priming in bi-alphabetical word perception. *Journal of Memory and Language, 28,* 237-254.

Martin, R. C., & Jensen, C. R. (1988). Phonological priming in the lexical decision task: A failure to replicate. *Memory & Cognition, 16,* 505-521.

Mattingly, I. (1985). Did orthographies evolve? *RASE, 6,* 18-23.

Mattingly, I. (this volume). Linguistic awareness and orthographic form. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning.* Amstederm, The Netherlands: Elsevier.

McClelland, J. L., & Rumelhart, D. E. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructures of cognition: Vol. 2. Psychological and biological models.* Cambridge, MA: MIT Press.

Meyer, D. E., & Ruddy, M. G. (1973, November). *Lexical memory retrieval based on graphemic and phonemic representation of printed words.* Paper presented at meeting of the Psychonomic Society, St. Louis, MO.
Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). Loci of contextual effects on visual word-recognition. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 98-118). London: Academic Press.
Naish, P. (1980). The effects of graphemic and phonemic similarity between targets and masks in a backward visual masking paradigm. *Quarterly Journal of Experimental Psychology, 32,* 57-68.
Navon, D., & Shimron, J. (1981). Does word naming involve grapheme-to-phoneme translation? Evidence from Hebrew. *Journal of Verbal Learning and Verbal Behavior, 20,* 97-109.
Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research, 53,* 13-24.
Perfetti, C. A., Bell, L. C., & Delaney, S. (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language, 27,* 59-70.
Peter, M., & Turvey, M. T. (April, 1992). Semantic categorization of backward masking. Paper presented at the Eastern Psychological Association.
Peter, M., Turvey, M. T., & Lukatela, G. (1990). Phonological priming: Failure to replicate in the rapid naming task. *Bulletin of the Psychonomic Society, 28,* 389-392.
Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 10,* 645-657.
Schmidt, R. C., Beek, P. J., Treffner, P. J., & Turvey, M. T. (1991). Dynamical substructure of coordinated rhythmic movement. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 635-651.
Sebastián-Gallés, N. (1991). Reading by analogy in a shallow orthography. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 471-477.
Shulman, H. G., Hornak, R., & Sanders, E. (1978). The effect of graphemic, phonetic, and semantic relationships on access to lexical structures. *Memory & Cognition, 6,* 115-123.
Turvey, M. T. (1990). Coordination. *American Psychologist, 45,* 938-953.
Turvey, M. T., Schmidt, R. C., & Beek, P. J. (in press). Fluctuations in inter-limb rhythmic coordinations. In K. Newell (Ed.), *Variability in movement control.* Champaign, IL: Human Kinetics Press.
Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound and reading. *Memory & Cognition, 15,* 181-198.
Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 371-385.
Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97,* 488-522.
Wang, W. S-Y. (1981). Language structure and optimal orthography. In O. J. L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology* (pp. 223-236). Hillsdale, NJ: Erlbaum.
West, R. F., & Stanovich, K. E. (1982). Source of inhibition in experiments on the effect of sentence context on word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 5,* 385-399.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 13

Reading in English and Chinese: Evidence for a "Universal" Phonological Principle

Charles A. Perfetti
Sulan Zhang
Iris Berent
University of Pittsburgh

The role of phonology in reading has not been without controversy, if we may begin with an understatement. Some researchers have concluded that phonology plays an inevitable role in reading, arguing that meaning and print are not related in quite the same way as meaning and speech—that reading is partly dependent on the association of print with spoken language (Gough & Hillinger, 1980; Liberman & Shankweiler, 1979; Mattingly, 1972; Perfetti & McCutchen, 1982). In contrast, expressing a point of view that represents a dominant school of thought, at least in American reading education, Frank Smith (1985) observed that "...meaning and print are related in the same manner as meaning and speech, and neither language form is dependent upon the other" (p. 57).

Part of Smith's argument that reading is a sort of parallel language channel, independent of speech, involved observations of Chinese writing. Since Chinese is a logographic system, it allows no speech mediation between the written form, the character, and its meaning. There is nothing in the character 马 that leads to the Mandarin pronunciation ("ma") of the word meaning "horse." Meaning comes first, then sound. Reading English appears to be quite a different matter, because alphabetic writing systems associate speech segments with print. Smith, however, further claimed that fluent readers of English "...recognize words in the same way that fluent Chinese readers recognize [words]" (p. 103). By this Smith meant that readers of both English and Chinese use only visual information in recognizing words.

There's a sense in which Smith was right about recognition in English and Chinese being similar. There may be more in common between reading Chinese and reading English than meets the eye. This similarity, however, contrary to the view represented by Smith, includes the fact that both writing systems allow a significant role for phonology. Our suggestion, which we develop in the remainder of this chapter, is that the use of phonology is a general characteristic of reading that exists across writing systems.

Generalized phonological activation

Our assumption is that contact with printed words in any writing system automatically arouses phonological properties associated with the words. There are several difficult questions that such a claim raises:

(1) When does this phonological activation occur?
(2) How general is the activation?
(3) What are the phonological properties that are activated?
(4) What function does the activation serve?

When?
Word identification research has been preoccupied with question #1. Is phonological activation pre-lexical or post-lexical? This question is essentially whether phonology mediates the identification of a word (pre-lexical) or follows its recognition (post-lexical). Our preferred answer is "Let's ask another question." The question of timing is important for getting straight the story about how words are recognized. We are not sure, however, that this is the most important question in the long run. If phonological activation always occurs and if it always begins immediately on viewing a word, then the question of pre- vs. post-lexical access reduces to one of specific recognition events, which can vary. Our answer comes to this: Phonological activation always begins "pre-lexically." This activation always plays some part in identifying the word, provided the writing system allows it to do so. Alphabetic systems do, logographic systems, generally speaking, do not.[1]

How general?
There are at least three parts to the generality question. First, phonological activation occurs generally across writing systems, differing mainly in timing. Whether a writing system is alphabetic, logographic, or syllabic affects only the level at which graphic units connect to phonological units. The level of the connections constrains how much early pre-lexical activation occurs. Second, within alphabetic writing systems, phonological activation is general across all elements of the writing system that correspond to phonological objects in the language. Thus, in English, the letter *t* activates the phoneme /t/, the digraph *sh* activates the phoneme /ʃ/, the letter string *-ame* activates the phonemic string /eym/, and the letter string *ate* activates both the phonemic string /eyt/ and the phonological word /ATE/.[2] This does not mean that letter-phoneme connections play a decisive role in every instance of identification. It is hard to imagine that the letter string *the* is processed as other than a visual sign, mainly because of its disproportionally high frequency in English. Based on the Kučera and Francis (1967) word count for printed English, 7 out of every 100 words read by an adult reader are the word "the." By age 20, even a college student who is a very infrequent reader will have encountered "the" over 50,000 times, ample exposure to make it function as a visual sign rather than a linguistic object. On the other hand, to assume that some words serve as signs is not to suggest that there are very many of them.

---

[1]We are tempted to suggest that phonological activation occurs, at latest, at the point of identification. To identify a word is to take a brief accounting of its pronunciation, or its name, perhaps in combination with meaning properties of the word. This suggestion is better captured by the phrase "at-lexically" or just "lexically" rather than "post-lexically." Thus the contrast is between "pre-lexical" and "lexical" phonology.
[2]Alternatively, one might suppose that for letters that spell actual words, there is a chain of activation: *ate* to /eyt/ to the word "ate." This is the standard straw man form of phonological mediation. By assuming that *ate* activates both phonological sequences and phonological words, we simply indicate the need to have word pronunciations as part of a word's representation. This is a commitment to a post-lexical phonology, whatever the exact status of a prelexical phonology.

Third, phonological activation is general across individuals. Although there is evidence suggesting that low ability readers in English are less successful at using phonology (e.g., Brady, 1991; Liberman & Shankweiler, 1991), there is no strong evidence that low ability readers do not engage phonological processes when they read. Attempts to describe some developmental dyslexics as visual-only readers lacking phonological processes seem misguided. We are more inclined to say that most dyslexics are ineffective in phonological processes, rather than lacking them. This generality across individuals seems to extend to skilled deaf readers as well (Conrad, 1979; Hanson & Fowler, 1987; Hanson, Goodell, & Perfetti, 1991). For example, in an experiment with deaf college students, Hanson et al. (1991) had subjects read tongue twister sentences (sentences with initial consonant repetition) while trying to remember strings of digits. They found an interaction between the phonetic contents of the digit memory load and the phonetic contents of a tongue twister sentence, a result that seems to implicate phonological processing during reading for these skilled readers. Less successful deaf readers may be those for whom visual-only is the only possible way to read. More successful deaf readers may be those who have managed to retain some form of phonological processing in their reading.

What phonology?

The phonology activated during reading is the phonology of spoken word forms. On the way to this activation, phonemes are activated unconsciously, whereas the phonological word forms often reach a more conscious state. In an earlier account of a generalized activation model, Perfetti and McCutchen (1982) suggested that the phonological representation was abbreviated, compared with the actual pronunciation of a word. We continue to assume that this is correct in general, because the pace of reading can exceed considerably the rate of speaking. However, the further suggestion that the abbreviated form includes primarily word-initial phonemes, which Perfetti and McCutchen argued allowed unique indexing of many words, may be incorrect. Although the initial segments of a word are always activated, the experience of silent reading is more akin to time-compressed speech than to shorthand. Whatever the exact form of the representation, we assume that it is phonologically specific, containing information on how the word is pronounced. This representation eventually will include information that can be obtained only after other aspects of the word's identity have been established, since *record*, for example, cannot be pronounced until its grammatical class is encoded.

What function?

Phonological activation serves word identification, but it has an equally important and more universal function in comprehension. Phonological word forms are part of the reader's short-term memory, and comprehension depends on this in several ways. First, there is what we have called *reference securing* (Perfetti & McCutchen, 1982), establishing memory representations that are specific enough to allow specific discourse referents to be accessed on demand. Semantic information abstracted from word forms is not enough. The functional difference between *dog* and *canine* is their distinct phonological forms, not their meanings. It is clear that readers in fact generally retain the exact wording for a sentence or at least a clause (Goldman, Hogaboam, Bell, & Perfetti, 1980), just as listeners do (Jarvella, 1971). This verbatim retention allows a representation over which to parse sentences, repair parsing failures, integrate propositions and make inferences.

Whether phonological representations serve comprehension under all conditions or only when memory demands are present has received different answers (Levy, 1978; Slowiazcek & Clifton, 1980). Our assumption is that the activation is fully general, but that the reader can add rehearsal processes to the representation in response to text demands. The starting point is the relatively passive activation of phonological word forms.

Before turning to some of the research that gives shape to this account of phonological processes, we must digress into some unavoidable theoretical thickets.

Principles and mechanisms

The principles of generalized phonological activation can be realized in various ways. Dual route theories of word identification (Besner, 1990; Coltheart, 1978; Meyer, Schvaneveldt, & Ruddy, 1974; Paap & Noel, 1991) assume that phonological conversion processes occur along one route, while direct access, unmediated by phonology, occurs along the other route. Whether our proposal is a form of the dual route model depends on the assumptions made about phonological activation. A dual route model that makes an all-or-none dichotomy between direct visual access and phonologically mediated access is not compatible with our proposal. If phonological activation begins pre-lexically and is fully general across word types, then any model that implies strategic choice of routes is wrong. Any difference between the two routes must be quantitative and not qualitative. If what dual routes amount to is generalized phonological activation along both "routes," as most recent versions of dual route theory seem to allow (Besner, 1990; Paap & Noel, 1991), then they are compatible with our proposal. Such theories can allow a strictly quantitative account of identification, with phonological activation always occurring but being insufficient to cause identification by itself on many occasions, on which occasions a process that tries to access a word based on its appearance wins out. It's not clear what the value of two distinct routes is in such a case, however, since a single route is sufficient to handle merely quantitative processes.

A seemingly very different class of models comes from parallel distributed processing models (PDP), which assume generalized activation as the only mechanism and assume further no independent representation of word forms. Lexical representations are emergent properties of activation systems, having no permanent life. The development of these models, especially those of Seidenberg and McClelland (1989) and Van Orden, Pennington, and Stone (1990), appear to offer exactly the kind of implementation of sublexical phonology that is implied by the general principles of a phonological activation. Van Orden et al.'s (1990) account is especially congenial to the idea of generalized phonological activation because it is developed specifically in the context of a research program that has achieved strong evidence for phonological activation in word recognition (Van Orden, 1987; Van Orden, Johnston, & Hale, 1988).[3]

These two classes of models, dual route and PDP, provide dramatically contrastive approaches to word identification. However, their most dramatic difference is not the mechanism of identification, since activation concepts (PDP) and phonological conversion rules (dual route theory) are essentially translatable, but in whether there is

[3]The Van Orden et al. theory is not actually implemented, but is a principle-based account of word identification that includes pre-lexical phonological activation as an inevitable component of reading words.

any mental representation to be identified. It's the disappearance of the lexicon under PDP assumptions that is the main bone of contention (Besner, 1990; Besner, Twilley, McCann, & Seergobin, 1990; Seidenberg & McClelland, 1990). Lexicon or no lexicon, that is the question, and put that way it's hard to imagine a more critical question in the study of reading, if not the whole of cognitive psychology.

It is useful to realize, however, that this question is embedded in meta-theoretical questions that do not lend themselves to straightforward empirical answers. The debate between connectionist and "classical" approaches to language is largely a matter of the explanatory status of mental concepts—not just "lexicon," but also "grammar," "knowledge," "inference," and "belief," to name a few. Our approach to this issue is to conclude that the debate must run its course informed by the accumulative persuasiveness of the two approaches. Meanwhile, we assume that network models, even old fashioned nondistributed ones (Rumelhart & McClelland, 1981) provide highly productive implementations of important generalized mechanisms. We have assumed that PDP approaches are of limited value in explaining higher level systematic knowledge of language (grammar), but that they find exactly their right level in word recognition, or at least parts of it. The natural domain of these models is implementing mechanisms that allow induction of patterns. And that's what a reader does when it comes to learning how letter strings map onto words. Thus we assume that an implementation interpretation of such models is useful (Fodor & Pylyshyn, 1988).

The disappearance of the lexicon is another matter, however. The problem is that the lexicon is needed for things other than word identification. Even in word identification, post-lexical phonology must be assigned just to get syllabic stress right. And all processes just milliseconds downstream from "lexical access" seem to require quite a bit of information about words. The mechanism of pre-symbolic sublexical activation has a strong appeal as an implementation of fast acting recognition (and word learning) processes. But the disappearance of the lexicon cannot be taken too seriously. It would be removed for word identification only to be reinstated for all subsequent reading processes.

In the context of this meta-theoretical thicket, we believe it is useful to stay at the level of principle rather than at the level of algorithm in a theory of phonological activation, or word identification generally. Implementation of the principles can be along network activation lines. On the other hand, a lexicon may be needed, certainly for other reading processes and probably for identification as well. So ours is a principle-based theory of phonological activation during reading, with lexicon. It consists in three closely related principles.

The central principle of our theory is that, across writing systems, encounters with most printed words (exceptions restricted to a short list of sign-like words) automatically lead to phonological activation, beginning with phoneme constituents of the word and including the word's pronunciation.

A second principle is that writing systems constrain the extent to which this activation includes sub-lexical phonology, but not whether activation occurs.

A third principle is that activated phonology serves memory and comprehension, with phonological rehearsal but not the activation itself under reader control.

We turn now to a discussion of some research that causes us to think these principles may be partly correct, as far as they go. It will turn out that plenty of doubt remains about some things, especially the degree to which activation is automatic under all circumstances.

Lexical experiments

The research on phonological activation in word reading is substantial. The number of paradigms approaches profligacy: word naming, lexical decision, priming (with various measures), category judgment, phrase acceptability judgments, letter identification with masking, misspelling identification (and other search tasks), and backward masking. A number of reviews are useful for this purpose, including the dated-but-still useful McCusker, Hillinger, and Bias (1981) and the more recent argument-reviews of Besner (1987; 1990) and Van Orden et al. (1990). The earlier criticisms we made of some research (Perfetti & McCutchen, 1982) are still valid. We are unable to evaluate the more recent work in all these paradigms in this chapter, so we instead place our work in the following context: The conclusions drawn in the research do depend in part on the paradigms. Each has a characteristic obstacle to interpretation when the hypothesis concerns the fine grain that is entailed by immediate phonological activation. We do not claim that our experiments overturn the conclusions drawn in favor of direct visual access based on other paradigms. However, we do think that our results at least keep open the case for early phonological activation. When coupled with other research, especially that of Van Orden and colleagues (1987; Van Orden et al., 1990) and Lukatela and colleagues (Feldman & Turvey, 1983; Lukatela, Popadić, Ognjenović, & Turvey, 1980; Lukatela & Turvey, 1990a,b), it makes a case for automatic phonological processes prior to complete word identification.

The method of choice for us has been the backward masking paradigm, which we have exploited in studies reported over the past few years (Perfetti & Bell, 1991; Perfetti, Bell, & Delaney, 1988; Perfetti & Zhang, 1991). Naish (1980) was, as far as we know, the first to report phonological effects using this procedure, although his paper was centrally concerned with whether the paradigm was sensitive to semantic effects.

The backward masking experiment

In a typical experiment, a sequence of three stimuli is presented in a 3-field tachistoscope. Each electronic shutter is controlled by a computer program that allows millisecond control of exposure duration. A typical trial sequence consists of (1) a word target of brief exposure, (2) a pseudoword mask of brief exposure, (3) a pattern mask exposed until the next trial, serving also as a fixation area. The ISI=0 in each case. The target word is in lower case, the masking pseudoword in upper case. There are three important variables: the duration of the target, the duration of the mask, and the type of mask. The following sequence represents a trial with a Phonemic Mask: *rate*—RAIT—XXXX. A Graphemic Mask shares the same number of letters with the target as the phonemic mask but is not homophonic to it, e.g., *rate*—RALT; Control Masks share no letters or phonemes with the target, e.g., *rate*—BUSK.

The key data are target report accuracy rates as a function of mask type. Exposure durations are fixed to assure that complete identification is not possible on all trials. All experiments to date have shown two types of mask reduction effect (MRE): a Graphemic MRE: Relative to Control Masks, more targets are reported with Graphemic Masks. A Phonemic MRE: Relative to Graphemic Masks, more targets are reported with Phonemic Masks. The effects are illustrated in Figure 1.

The interpretation of these effects rests on the assumption that partial products of identification (letters and phonemes) are active when the pseudoword mask interrupts processing. To the extent that the pseudoword mask reinstates these partial products, there is a reduction of the deleterious effect of the mask. By this logic, the Graphemic MRE reflects the activation of letters of the target prior to its identification (abstract letter representations, because of the case alternation). The Phonemic MRE reflects the activation of phonemes of the target prior to its identification. It is this latter effect that is central to the conclusion favoring early (pre-lexical) phonological activation. Notice that, since the Graphemic Mask shares phonemes as well as letters with the target, part of its effect reflects phonological processes. Thus, the Phonemic MRE is a fairly conservative estimate of the degree of phonological activation.

The reinstatement explanation is not the only one possible for these effects. Subjects might engage in some form of sophisticated guessing that favors the Phonemic MRE. If subjects believe they are seeing homophones, and, on some percentage of the trials, they identify the mask but not the target word, then they might generate the corresponding real word homophonic to the mask. We addressed this problem in Experiment 2 of Perfetti et al. (1988) by introducing occasional blank trials, in which only a pseudoword mask appeared. If subjects were to guess the target based on viewing only the mask, then we should get a pseudo-phonemic MRE with blank trials, but we did not. Further, masks were re-paired with targets so that a phonemic mask became a control mask. This too produced no pseudo-phonemic MRE. Within the limits of these controls, then, the Phonemic MRE cannot be explained in terms of simple guessing. There may remain the possibility of a more sophisticated strategy in which the subject combines letters from the target with letters obtained from the mask in a way that somehow favors phonemic masks. We think this is unlikely, given that there is no evidence for a simple guessing strategy. That is, any sophisticated strategy involving masks ought to show some effect when the target is the only source of information.

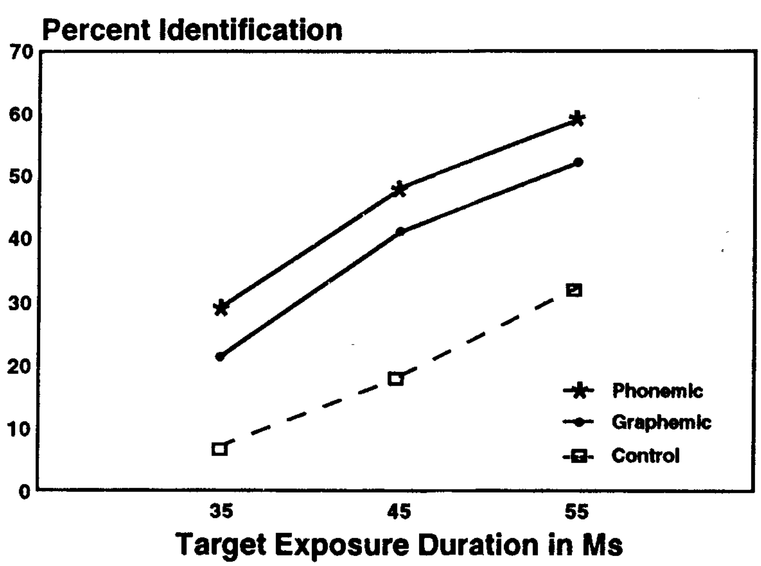

Figure 1. **The backward masking reduction effect. Percent target identification as a function of target duration for 3 pseudoword mask types. (Mask duration = 30 ms.) Adapted from Perfetti and Bell (1991).**

We have addressed a more direct question of interpretation in a recent experiment. Humphreys, Evett, and Taylor (1982) failed to find priming effects in a procedure that was essentially the reverse of our masking experiment, and concluded that prelexical phonology was not activated in word identification. The contradiction appears to have been resolved in Experiment 3 of Perfetti and Bell (1991), who found that such priming *does* occur, provided the prime is viewable for about 45 ms, as shown in Figure 2.[4]

The activation of phonological information in nonwords may simply take longer than in real words, which can benefit from lexical level feedback to sublexical units. In any event, priming and masking now seem to tell the same story about phonological processes.

The general conclusion from these experiments is that prior to word identification, some of the phonemic constituents of the word are being activated. The conclusion is that this is "prelexical" phonology, on the assumption that we are interrupting lexical access and immediately reinstating decaying traces of both graphemic and phonological activation. The phonemic MRE and the phonemic priming effect so far are both entirely general. Perfetti and Bell (1991) manipulated target word frequency and spelling pattern consistency and found that neither affected the MRE (Experiment 1) or the phonemic priming effect (Experiment 3).

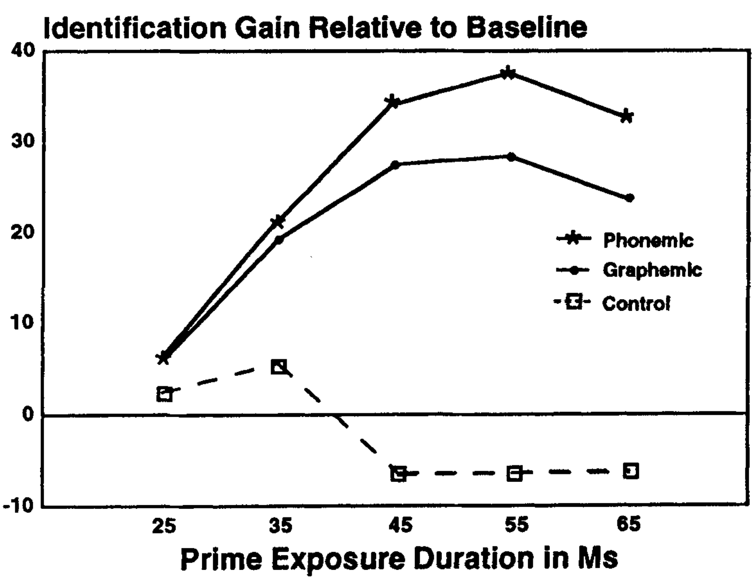

Figure 2. **The priming (or forward masking) effect. Percent gain (loss) in target identification as a function of prime exposure duration for three types of pseudoword "primes." (Target exposure = 30 ms). The gain is relative to an unprimed baseline. Adapted from Perfetti and Bell (1991).**

[4]Although we refer to "priming" effects, they can also be considered forward masking effects. A briefly presented pseudoword forward masks a following real word. Figure 2 actually shows the negative effects of forward masking: As the duration of a control mask increases beyond 35 ms, identification is slightly suppressed relative to baseline identification in which the target is presented with only a following pattern mask. The "primes" work by adding a mask-target relation that overcomes the general masking effect.

The generality of these phonological effects is not predicted by a theory that assumes selective procedures in lexical access. Dual route theory takes both consistency (or regularity) and frequency effects as indicators of which "route" is being used in identification. Words with consistent spellings are supposed to be better able to take advantage of grapheme-phoneme translation routines, favoring the "phonological assembly" route. Word frequency is taken to be a lexical variable, sensitive to the direct access (or "addressed phonology") route. Thus, finding that phonemic mask effects were restricted to words with consistent spelling patterns might be expected on the assumption that only the phonological route can produce phonological effects. Finding that the effects were restricted to *low* frequency words with consistent spelling would be even further in line with the dual route story, which expects words of high frequency to be ordinarily handled by the direct route, with the slower acting phonological route confined to regular words of low frequency. Such an interaction between frequency and regularity has been found with naming tasks (Paap & Noel, 1991; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). However, the phonemic masking effect so far has been general, both across frequency and across consistency.

The generality of these phonological effects, i.e., effects based on nonword primes and masks, is supported by studies in other languages. In a Serbo-Croatian study, Lukatela and Turvey (1990a) found facilitative effects of nonword homophones in the masking paradigm, just as has been found in English. Moreover, they found that this facilitative effect held across alphabets. When the target word was written in Roman and the mask in Cyrillic, or vice versa, there remained a large advantage for a nonword homophone mask. In another study, Lukatela and Turvey (1990b) found facilitative effects of nonword phonological primes in naming and lexical decision in several experiments, including one using a briefly (80 ms) presented masked prime. In a French language study, Ferrand and Grainger (in press) found nonword phonological effects at 64 ms of masked prime presentation. The phonological priming effect, like ours, was independent of target word frequency. Ferrand and Grainger found little or no phonological facilitation, however, when the nonword prime was exposed for only 32 ms. In the Perfetti and Bell study (Experiment 3), phonological priming effects showed themselves at prime durations between 35 and 45 ms. (See Figure 1).[5]

The finding that briefly presented nonwords can produce facilitation of word identification in both priming studies (Ferrand & Grainger, in press; Lukatela & Turvey, 1990b; Perfetti & Bell, 1991) and masking studies (Lukatela & Turvey, 1990a; Perfetti et al., 1988, Perfetti & Bell, 1991) provides considerable evidence for the activation of phonology during word identification. The fact that both perceptual identification (accuracy measures) and speeded response measures display these effects provides further support for the robustness of phonological effects across paradigms.[6]

We have been addressing the "When?" question and part of the "How general?" question in reviewing this research: There is evidence to support the claim that

[5]In the masking experiments, unlike the nonword priming experiments, we have not yet found a point of divergence between a graphemic effect and a phonemic effect. Perfetti and Bell (1991) found both effects at 35 ms, the shortest target duration tested.

[6]There is doubt about whether the phonological activation observed in the masking paradigm is sensitive to variables whose effects reflect strategic processing. Brysbaert, Praet, and d'Ydewalle (1990) found phonemic effects in masking only when the proportion of homophone masks was relatively high, a result that appears to argue against the assumption that mask reduction effects reflect strictly automatic phonological processes.

phonological activation occurs *during* the process of word identification. Evidence within the nonword masking and priming paradigms at least is consistent with the possibility that this activation is general across words in an alphabetic writing system, although this has to be a very tentative conclusion. We turn now to another aspect of the "How general?" question, the extent to which phonology occurs in writing systems that are not alphabetic.

Phonology in reading Chinese

We have already noted the generality of phonological evidence across different languages —French (Ferrand & Grainger, in press) and Serbo-Croatian (Lukatela & Turvey, 1990a,b)—that are encoded by alphabetic orthographies. An important question is what happens for orthographies that are not alphabetic. Chinese provides probably the clearest contrast. It is a logographic system, i.e., one in which the units of written language correspond generally to morphemes (usually words), not to speech segments. In such a system, the opportunities for pre-lexical phonology are severely limited.

The general picture of Chinese as a system in which characters stand for words is roughly correct, but there are a few important details that qualify this general picture. First, with respect to the visual components of characters, the vast majority of characters do not have readily identifiable pictographic values. While the character for rain (雨, *yu*) contains strokes that represent falling rain, the majority of characters have no pictographic component; Zhou (1978) cites an estimate of less than 18% that are either pictographic or ideographic. Nevertheless, the writing system contains cues to meaning. Most words (82%) are written as compounds, one of some 188 radicals combined with a character. The radical often gives semantic information, as when the radical for bird (鸟) combines with two different characters to produce "sea gull" and "goose." The reader's knowledge of written Chinese will tell him that 鸥 is a kind of bird, even if he does not know what kind of bird it is. Thus, in general, the relationship between the character and the meaning of the word is relatively abstract, if not arbitrary; but there are often graphic cues to meaning in compound characters.

Second, with respect to phonology, written Chinese is not without clues to pronunciation. In compounds, while the radical may cue meaning, the character name may cue pronunciation. For example, the character compound for "yak" (牦) combines the radical for "ox" (牜) with the character pronounced "mao" (毛, fur), which tells the reader that "mao" is also the pronunciation of the compound. (This reflects the pervasive homophony of Chinese, which results from its heavy use of monosyllabic words.) However, the vast majority of character names do not give good cues to the pronunciation of a compound. Zhou (1978) estimates that only 39% of compounds actually provide the correct pronunciation. The compound's pronunciation deviates from that of the character in the remaining 61%. For example, the radical pronounced "nu" (女, female) combines with the component pronounced "zi" (子, son) to give the compound pronounced "hao" (好, good).

For both the phonemic and semantic values of components of compound characters, there appears to be an interesting relationship with frequency: Both phonemic and semantic components appear to be more reliable for low frequency compounds than high frequency compounds. This conclusion comes from a sample of 300 compounds taken from three frequency ranges, as indicated in a Chinese frequency dictionary (1986). For each compound, we asked whether its name was the same as one of its components; if so, the component was said to have "phonemic validity." We also asked for each compound whether its meaning was related to one of its components; if so, the component was considered to have "semantic validity." Figure 3 shows the percentage of compounds with phonemic and semantic validity as a function of frequency.

Clearly, for both semantic and phonemic components, there is increasing validity with decreasing frequency. The relationship is even stronger when one takes into account the difference between single characters and compounds. Among a sample of 300 characters (different but overlapping with the sample shown in Figure 3), we found about 84% of the characters were compounds, very much in agreement with Zhou's (1978) estimate of 82%. But we also found the percentage of compounds, relative to single characters, to increase with decreasing frequency: 62% of high frequency characters, 93% of medium frequency characters, and 98% of low frequency characters were compounds. Thus as frequency decreases, the Chinese reader is less likely to encounter a single character (which can give no help in pronunciation), and, when he encounters a compound, it is more likely to give him reliable information for pronunciation and meaning.

The relationship between phonemic validity and frequency has an obvious value for oral reading: An encounter with a rare word has a chance to produce a correct guess as to its pronunciation. In a sense this mimics the state of affairs in English, where irregular words are overrepresented among the most frequent words and regular words overrepresented among the low frequency words.

One general point we wish to make here is that Chinese has the properties commonly attributed to it by non-Chinese writers, but in a probabilistic rather than an all-or-none manner. It is a meaning-based logographic system, but the semantic relations are abstract, more category indicators than pictures. And although the system does not encode phonological segments, it does encode probabilistic cues to pronunciation at the name (single syllable) level. And for both semantic and phonological information within a compound, the validity of the component information increases with decreasing frequency. Nevertheless, the average unreliability of the pronunciation cues makes them poor candidates for anything like facilitative *automatic* phonological activation.

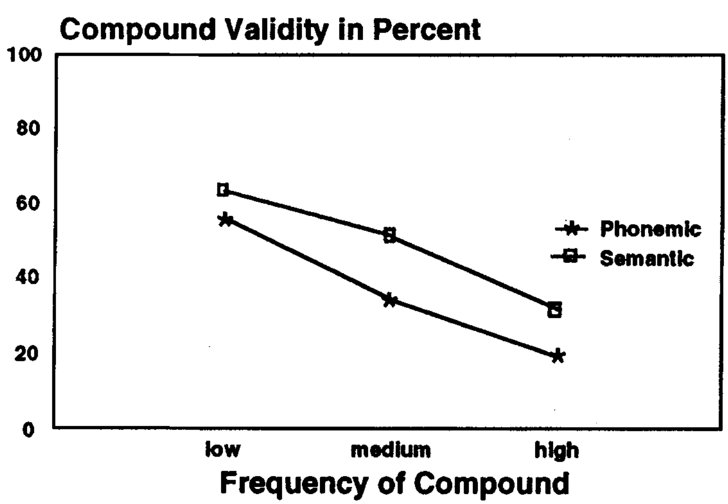

Figure 3. **The validity of both semantic and phonemic components of Chinese characters decreases with the printed frequency of the compound. The difference between semantic and phonemic validity is not meaningful. See text for explanation.**

If there is automatic activation of the names of characters *prior* to the access of semantic information, this activation will be a mix of inhibition (roughly 60%) and facilitation (about 40%) across encounters with the 6452 compounds (Xinhua dictionary, 1971) that contain a potential phonemic component.

Phonology in memory

Although some writers have imagined that Chinese reading allows no role for phonology (e.g., Smith, 1985), the reality has been known to be more complex for some time. Tzeng and colleagues demonstrated that, whatever happened during the initial identification of a character, Chinese readers showed a reliance on phonological codes in memory. Tzeng, Hung, and Wang (1977) found phonological confusions in a memory task and Tzeng and Hung (1980) extended this observation to include comprehension tasks. This seems to place reading Chinese in the same category as reading English, French, or Japanese. Beyond the level of word identification, phonological codes are activated in a working memory system in support of comprehension (Baddeley, 1979; Levy, 1977; Perfetti & McCutchen, 1982; Slowiaczek & Clifton, 1980).

One method used to get at phonology in reading comprehension is the reading of tongue twisters, a procedure in which the reader reads sentences or paragraphs that repeat initial phonemes (Haber & Haber, 1982; McCutchen & Perfetti, 1982). A basic finding is that the time to silently read sentences and make acceptability judgments is longer for tongue twisters than for control sentences matched for syntactic form and semantic content (McCutchen, France, Bell, & Perfetti, 1991; McCutchen & Perfetti, 1982). Tongue twister effects have also been found in text reading, where there is no separate demand for a judgment by the subject, although here the effect showed itself less in reading times than in comprehension failures (France, 1989).

Our interpretation of the tongue twister effect is that it arises from phonological memory interference: As they are read, words are coded as phonological forms. When a string of phonologically similar words is activated, as in reading a sentence, there is some interference among the phonological codes. That this effect arises in memory has been confirmed by McCutchen et al. (1991), who manipulated the phonological contents of memory by requiring subjects to remember digits while they were reading the sentences. When the digits contained the same phonemes as the sentence, there was a cross-over interaction for both reading times and digit recall. For example, when subjects read a sentence repeating the letter *t* and had to recall digits whose name begin with /t/, e.g., 10, 22, 27, 2, 12, both recall of the digits and the reading time for the sentence were adversely affected, compared with a tongue twister paired with digits repeating some different phoneme (e.g., 6, 66, 72, 7, 16).

Although we think such results secure the interpretation of the tongue twister effect, there has been room for doubt, including the possibility that the basic tongue twister effect itself is largely visual. Sentences that repeat phonemes tend also to repeat letters, although McCutchen and Perfetti (1982) unconfounded letters and phonemes to a considerable extent. This is where the Chinese case becomes especially interesting. As a nonalphabetic language, Chinese naturally separates the appearance of the written characters from the phonology of the words. To find a tongue twister effect in Chinese would remove all doubt that its source is visual.

In our recent experiments on the tongue twister effect in Chinese, we have extended the generality of the effect and further secured its interpretation. The subjects, all speakers of Mandarin from the People's Republic of China (residing in Pittsburgh) read tongue twister texts, short stories of 6-8 lines each, and matched controls (matched on word frequency, semantic content and length). The tongue twister texts represented 4 different kinds of phoneme repetitions defined by place of articulation and represented in each case by at least two different phonemes—alveolar stops (/t/,/d/), bilabial stops (/b/,/p/), velar stops (/k/,/g/), and alveolar fricatives (/z/,/s/,/c/). Subjects were instructed to read as quickly and as accurately as possible and to recall as much of the text as possible. The key data were reading times per character: For tongue twister and control texts the mean silent reading times per character were 436ms and 364ms, respectively. The magnitude of this effect is larger than we have observed in English experiments, a fact we think is due to the greater density of phoneme repetition we were able to achieve in Chinese. Syntactic factors in Chinese, especially the absence of articles and rarity of prepositions, allowed the density of phoneme repetition to be 72% in our Chinese experiments, compared with about 50% in the English experiments. The syllable-based density difference was even greater, owing to the monosyllabic structure of Chinese.

In another study with Chinese, we replicated the interaction found by McCutchen et al. (1991) for English. Chinese readers remembering digits while reading tongue twister passages had their reading times slowed even further when the digits they had to remember had the same phonemic content as the passages they were reading. This further localizes the effect of consonant repetition in working memory, a system that must be active during comprehension in reading both nonalphabetic and alphabetic writing systems.

We draw two main conclusions from the Chinese studies. First, finding the tongue twister effect in Chinese confirms the original explanation of this effect (Perfetti & McCutchen, 1982). It is not a visual effect, because Chinese completely unconfounds visual and phonological information. Second, as a phonological effect, it demonstrates that reading Chinese engages phonological codes in working memory, the same as in English.

Phonology in word identification

A role for phonology in remembering and understanding written Chinese is expected by a consideration of a human working memory system that uses speech information. It's at the lexical level that the writing system should matter. Our goal here has been to learn whether there is a role for phonology in identifying characters. According to the universal principles we have suggested, there should be a role for phonology at the earliest point permitted by the writing system. We expect phonology not to mediate the recognition of a character but to immediately accompany its recognition—lexical phonology rather than "prelexical" phonology.

One study that suggests such a role for phonology takes advantage of the multiplicity of spoken Chinese languages and dialects. Mandarin and Cantonese, for example, are different enough in their phonological forms to be mutually unintelligible. The writing system, however, is shared by the various Chinese languages. Thus a speaker of Mandarin and a speaker of Cantonese will both understand a single Chinese written text, although if they were to read the text aloud, two quite different productions would be heard.

Lam, Perfetti, and Bell (1991) exploited this situation by asking subjects who were bidialectical in Mandarin and Cantonese to make homophony judgments for pairs of presented characters. For each pair of characters, subjects had to decide whether they had the same pronunciation in one of the two dialects, either Mandarin or Cantonese. The key condition was one in which the two characters had the same pronunciation in one dialect but not the other. For subjects whose native language was Cantonese, clear phonetic interference effects were observed when they were asked to base their decision on

Mandarin: If two characters had the same pronunciation in Mandarin, the "same" response was slowed when they had different pronunciations in Cantonese. If two characters had different pronunciations in Mandarin, the "different" response was slowed when they had the same pronunciation in Cantonese. This effect was not found when the decision was based on Cantonese, their first language. Thus readers were not able to suppress the pronunciations of the characters provided by their native language.[7]

Although these results implicate automatic phonological activation in reading characters, their generality is limited by the nature of the task. Subjects had to make judgments based on the phonological forms. Whether phonology is activated when it is not needed is the more interesting question. Although still not decisive for this question, the experiments of Perfetti and Zhang (1991) provide some converging evidence.

Perfetti and Zhang (1991) carried out a series of masking and priming experiments with Chinese readers (in Pittsburgh), modeled on the English experiments (Perfetti, Bell, & Delaney, 1988), with the addition of a semantic mask (and prime). In the masking experiment, subjects identified characters briefly exposed then masked by a homophonic, graphic, semantic, or neutral character mask. The target word was presented at the subject's threshold for 50% (40%-60%) identification, with actual duration varying between 30 and 70 ms. The character mask was exposed for 30 ms. Consider what should happen in this situation. Although the same procedure in English produces higher target identification rates when there is a homophone mask, no such benefit should exist for Chinese, if our assumption that the effect in English is pre-lexical is correct. There is no reliable pre-lexical phonology in Chinese. If an effect is observed in Chinese, we should reconsider the assumption that it is pre-lexical in English, figure out how it might be pre-lexical in Chinese after all, or figure out how it could be pre-lexical in English but post-lexical in Chinese.

None of these distasteful alternatives have to be considered. Perfetti and Zhang (Experiment 1) found no Phonemic MRE and no Semantic MRE, with identification rates between 40% and 50% for both conditions as well for as controls. Thus, there appears to be no phonemic mask reduction effect in Chinese, strengthening the pre-lexical interpretation given to the alphabetic results in English and Serbo-Croatian.

The universal phonological principle claims activation of phonology as soon as possible within the limits of the writing system. So there should be phonological activation as the character is actually recognized. Evidence comes from the priming experiments of Perfetti and Zhang (1991), which primed a briefly presented target character with a briefly presented prime character, a procedure opposite to masking. These experiments varied the duration of the character prime and exposed the target for 35 ms under the average 50% threshold. If the pronunciation of the prime is activated as it is recognized, it should be available as the target is exposed, facilitating its identification, "post-lexically." Of course, if the prime is not identified, it produces no post-lexical phonology, so the duration of the prime ought to be important (as it is in English, as Perfetti and Bell (1991) have found). The key results were no priming effects at 20 ms and small but significant effects for both semantic and phonemic primes at 50 ms of prime exposure (Experiment

[7]This is a result reminiscent of the bi-alphabetic studies in Serbo-Croatian carried out by Lukatela and colleagues (Lukatela, Popadić, Ognjenović, & Turvey, 1980; Lukatela, Turvey, Feldman, Carello, & Katz, 1989). Whereas the Chinese case is two languages and one writing system, the Yugoslavian situation is one language and two writing systems. Both create a situation in which a graphic input can activate incompatible phonemic representations, and both produce evidence for automatic phonological activation.

3). Semantic and phonemic effects were virtually identical. These results are consistent with the prediction that phonology will accompany lexical access but not precede it.

Finally, in an experiment that assured that prime duration was sufficient for recognition (180 ms), very large phonemic effects (and smaller semantic effects) were found on naming times. This experiment also produced significant semantic priming, although the effect was smaller than the phonemic priming effect. The results of this series of experiments suggest some interesting possibilities for the role of phonology in reading Chinese. The masking results follow the constraints of the writing system, showing no prelexical phonology. The priming results, however, strongly suggest that phonemic information is immediately available ("lexically") as part of character identification in Chinese. Indeed, the evidence of both masking and priming can be taken as demonstrating "no semantics without phonology." When there was a semantic effect, there was a phonological effect. When there was no phonological effect, there was also no semantic effect.

The results of the Chinese experiments suggest that word identification involves phonology as a part of the process. This is not to say that phonology mediates identification, but rather that it is a component of identification. As in English, to identify a word is to be able to "name" it as well as be able to appreciate its sense. The time course of phonological activation appears to be slower in Chinese than in English. And the activation process is not one of "assembly" but of "retrieval," to take a dual route view of the process. There is more to learn about the time course of Chinese word identification and the role that component identification processes play in identification of compounds.[8] But the principle that writing systems constrain the level at which phonology is activated and not *whether* it is activated seems to hold, when considering the contrast between alphabetic and logographic systems.

Other writing systems

There are, of course, other writing system comparisons to be made. Although our experiments so far have been restricted to English and Chinese, other writing systems have received attention in research with other paradigms, including the Japanese Kana, a syllabary system, and Hebrew, an alphabetic system. Also, an interesting comparison has been added by a recent study of Persian (Baluch & Besner, 1991), an alphabetic orthography with a mix of 3 vowel letters and 3 diacritics. We will briefly discuss the case of Hebrew, since there has been more work on it.

Hebrew is an interesting example, because of its omission of vowels, which is standard in adult texts. Representing a word by only its consonants amounts to an intermediate kind of system, alphabetically indeterminate. The presence of consonants without vowels creates serious ambiguities. To illustrate with English, it's as if one encountered CR and had to figure out whether the word was CAR, CORE, CARE, etc. Hebrew's pervasive homography would seem to discourage phonological processes in reading. In its effect, but not in its source, the Hebrew case is potentially similar to the Chinese case. Spoken Chinese is phonemically homophonous (many homophones at the phoneme level), but written Chinese is graphically unique, resulting in unambiguous lexical meaning. Written Hebrew is phonologically ambiguous, resulting in ambiguous lexical meaning. In

[8]For example, what are the relative time courses for semantic and phonemic processes on the components compounds? Some experiments by Flores d'Arcais (in progress) suggest that when asynchronies are introduced in the viewing of components, presenting the phonemic component first is less disruptive than presenting the semantic component first, at least for naming.

Chinese, the writing system does not support pre-lexical phonology and it provides unique lexical (phonology plus meaning) information, so phonological coding is not necessary. In Hebrew, the writing system supports pre-lexical phonology but it does not provide unique lexical information, so phonology is not helpful.

It is worthwhile to emphasize the important difference between Chinese and Hebrew, however. Because it is graphically unique, the Chinese writing system motivates a "direct" connection between the graphic unit and a lexical unit. Hebrew does not do this, because it has true ambiguity at the lexical level: It's not just that one doesn't know how to pronounce CR, it's that one doesn't know what word CR is. Thus the lexical indeterminacy of written Hebrew does not support a direct route to the lexicon any more than it supports a phonological mediation route. Instead, it requires a heavy use of context in lexical selection. It remains true, nonetheless, that both Chinese and Hebrew writing systems give the reader little motivation to generate phonology from a graphic input if an alternative route exists. We suspect again that there is little choice in the matter, and that phonology is generated, or as we prefer, activated, automatically in Hebrew as well as Chinese.

The processing differences between Hebrew and English will parallel those between Chinese and English. The orthography is connected to a word representation that includes a pronunciation—a "post-lexical" process. However, it is quite possible, even likely, that there is considerable "pre-lexical" phonology in Hebrew. The consonant letters should activate consonant phonemes, even in the absence of a vowel. The problem for word identification is that there may be insufficient activation of a single word candidate from consonants, the activation being widely shared among those words whose spellings are consistent with the consonants. The result is an early phonological activation of little or no help to lexical identification.

Most research on Hebrew so far appears to be consistent with the assumption that prelexical phonology does not mediate word recognition. In a study that compared reading tasks across three writing systems, Frost, Katz, and Bentin (1987) found that frequency effects of lexical status (word frequency, word vs. nonword) and semantic priming were greater in Hebrew (vowel-less) than in Serbo-Croatian and English. Such results are consistent with the assumption that word recognition relies substantially on lexical information in Hebrew. Note, however, that while such results are consistent with the direct use of lexical information, they are insufficient to address the role of pre-lexical phonology in reading the vowel-less Hebrew script.

One might expect that phonological mediation might occur when written Hebrew employs vowel markers, a convention for children's texts and religious texts that fully specifies pronunciation. The conclusions across different studies appear to be mixed, however (Bentin, Bargai, & Katz, 1984; Bentin & Frost, 1987; Koriat, 1984, 1985; Navon & Shimron, 1981). Methodological issues, in our view, cloud the interpretation of these studies, as they do most studies of phonological mediation. A prevailing approach of these studies is to infer the presence of pre-lexical phonology from comparing the reading of vowel-marked scripts with the reading of vowel-less scripts. The assumption is that if word recognition is mediated by a phonological code, then the increased phonological information available in the vowel-marked script will facilitate recognition. There is reason to question this assumption. The problem is that the two scripts differ in other ways: Their appearance is different and so is their use. (Ordinary words are more frequently seen in vowel-less script.) Either factor might lead to computational processes that contribute differentially to reading in the two scripts. For example, consider the allocation of visual attention. In vowel-marked script, it is conceivable that a first cycle of processing encodes letters and a second cycle encodes markers, a possibility also implied by Bentin and Frost (1987) in connection with their lexical decision results. Such a possibility allows phonological processing to be more complete in the vowel-marked script than in the vowel-less script, but at a cost to processing resources. Of course, we are claiming not that this is a correct account of how vowel-marked words are processed, but that there is reason to doubt that comparisons of the two scripts can be decisive about the use of phonology.

Even if the evidence of Hebrew argued more clearly against phonological mediation, that would not imply that there is no pre-lexical phonology associated with Hebrew word identification. The question of whether there is any phonological *activation*—as opposed to phonological *mediation*—prior to identification would remain open. One of the implications of a generalized activation principle is that writing system constituents, including Hebrew consonants, lead to activation automatically whether or not they facilitate identification of a word. For the moment, we might suggest that Hebrew, as it is commonly scripted, is similar to Chinese in delaying definitive phonological information until other evidence accumulates to select a word. But we would not be surprised to learn that there is some pre-lexical activation nonetheless, something that might occur, as we have suggested, even for Chinese compounds.

Orthographic depth

There is a general hypothesis that deals with the issues we have been raising, and that is the orthographic depth hypothesis (Frost et al., 1987; Katz & Feldman, 1983; Turvey, Feldman, & Lukatela, 1984). Shallow orthographies are those such as Italian and Serbo-Croatian, which provide a direct or "shallow" mapping between spellings and pronunciations; deep orthographies, including English, sacrifice this more direct mapping for a "deeper" one that preserves morphological structure. The result is that shallow orthographies have more invariance in their grapheme-phoneme mappings than do deep orthographies, whose mapping is more context dependent. The major claim of the depth hypothesis, as expressed by Frost et al. (1987), is that "in general, in shallow orthographies, phonology is activated directly from print, whereas in deep orthographies, phonology is derived from the internal lexicon" (p. 104). The comparison by Frost et al. (1987) of three writing systems, cited in the previous section, provided evidence for this hypothesis, suggesting that the three systems are ordered, from shallow to deep, as Serbo-Croatian, English, and Hebrew. As the depth of the orthography increases, they report increased effects of lexical variables on naming, of semantic priming on naming latency, and of the presence of nonwords on word naming accuracy.

The central assumption of the orthographic depth hypothesis is that the properties of the writing system as a whole determine the coding mechanism for individual written words. Baluch and Besner (1991) argue that this assumption is incorrect. While allowing for some influence from the writing system as a whole, they claim that across all writing systems, the "addressed" route dominates processing within a flexible multiple-route word-recognition process. Phonology will be used when individual reading circumstances encourage it (e.g., the presence of non-words in a list of words to be read), but not otherwise.

Our view of the orthographic depth hypothesis is that it expresses a useful generalization about how writing systems influence the recognition of printed words. Of course, the point that individual recognition events depend on variation within the local reading environment is also plausible. These are complementary perspectives when stated at this general level. They become contentious only when the question is specifically about the dominance of an "addressed" route vs. an "assembled" route. The depth view appears to be that this dominance is determined by the writing system, with the assembled route dominant for shallow orthographies. The dual route approach appears to be that the route is a matter of specific recognition events, which ordinarily favor the addressed, direct route even for shallow orthographies. It is this claim for a universal dominance of direct access that is the real issue, and one that, as far as we can see, remains to be decided—as does our claim concerning a universal phonological principle.

What is the connection between our universal phonological principle and the orthographic depth hypothesis, on the one hand, and, on the other, the universal direct access hypothesis of Baluch and Besner (1991; also Besner, 1987; Seidenberg, 1985) concerning the dominance of direct access within a dual route theory? Part of the answer is implied by our earlier discussion of principles and mechanisms: A generalized phonological activation process is compatible with dual route theory when the latter stresses that the assembled route is a route always taken, for better or worse. Our key complaint about dual route theory has been the suggestion that there are matters of strategy and choice to decide the route. If, as is increasingly the case, dual route theory amounts to the claim that there is always phonological activation in an alphabetic orthography, then it becomes indistinguishable from a generalized activation account, provided both models have a lexicon. From the generalized phonological activation view the hypothesized dominance of the "addressed" route remains just that, a hypothesis with uncertain empirical validation. If correct, it implies that the generalized phonological activation we claim for all alphabetic systems may be of limited value in many circumstances.[9]

The comparison with the orthographic depth hypothesis involves the same point. On the universal phonological activation principle, generalized phonological activation occurs for all orthographies that provide sublexical mapping to sublexical phonology. The effect of this activation is modulated by the writing system (and by individual recognition events) but activation does not disappear in deeper orthographies. Thus, universal phonology and orthographic depth are quite compatible.

Conclusions

We draw several conclusions, both empirical and theoretical. Empirically, we conclude that reading English and reading Chinese have more in common than has been appreciated when it comes to phonological processes. Our text experiments suggest that readers in both systems rely on phonological processes during the comprehension of written text.

[9]The degree to which this empirical question depends on the interpretation of various experimental tasks cannot be overestimated. Consider naming, a task considered by many researchers to be free of post-lexical influences. To the extent that naming involves systematic post-lexical processes that interact with lexical variables, the experimental work in support of both dual route theory and the depth hypothesis becomes considerably weakened. Take, for example, the assumption that the absence of nonwords within a list makes a subject more likely to use a direct route. Instead, the effect of nonwords might be to create a checking for pronunciation, a process influenced by frequency and other lexical variables, just those assumed to affect only lexical "access." The assumption that any word processing task is free of such problems is wrong, but the consequences of this faulty assumption are typically ignored. Balota (1990) also makes this point in a very different context.

Our lexical experiments show differences just where one might expect them: Evidence for early ("prelexical") phonology in English but not in Chinese; but evidence for still-early ("lexical") phonology in Chinese. The time course of activation appears to be only slightly different in the two cases. Thus, we see the similarity between Chinese and English readers not in their dependence on a visual route, but in their use of phonology as quickly as allowed by the writing system.

More generally, we conclude that phonological processes are pervasive in reading, with respect to various reading processes (from comprehension to word identification), with respect to writing systems (from Chinese to English to Serbo-Croatian), and with respect to individuals (from children to hearing and deaf adults of high reading skill). The pervasiveness of phonological processes is less a function of writing systems, although they are important, than of the human capacity for language. The universality of phonologically referenced language assures that the achievement of reading will make use of it. The acquisition of visually based spelling representations may (or may not) reduce the role phonology plays in recognizing words, but it does not entirely eliminate it. Moreover, the value of phonological representations for memory assures a critical role for phonology in comprehension.

Theoretically, we conclude that there may be more agreement on these matters than seems likely at first glance. Our main claim is that phonological activation is highly general across recognition events. We assume that lexical representations are what are contacted by orthographic and phonological inputs. But we also assume that activation mechanisms that operate, for alphabetic systems, within interconnected networks of word, letter, and phoneme representations, provide a natural explanation of the timing of events along the way to identification. This means that our generalized activation hypothesis is consistent with dual route theories if they make the same assumption about phonological activation always occurring. And our hypothesis is compatible with "old-fashioned" nondistributed parallel activation models, if the latter include phonemes, which they clearly can. The debate between dual route and connectionist models collapses distinctions that survive within our generalized activation model. In our view, what is identified are words (a lexicon exists), and the way they are identified is through automatic activation of words and word constituents, including phonemes as well as letters.

Nevertheless, our main point is one of principle, not one of models. The debate concerning models needs to be informed by the facts about how words are identified across different writing systems. General-level principles such as the universal phonological activation principle serve to reflect some of the facts, point to important research questions, and perhaps constrain models.

Acknowledgment

The research reported in this chapter was supported by National Science Foundation Grant BNS 8811042 to Charles A. Perfetti. We are grateful to Mara Georgi for her help with some of the research.

References

Baddeley, A. D. (1979). Working memory and reading. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Procession of visible language. Vol. 1* (pp. 355-370). New York: Plenum Press.

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.

Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*, 644-652.

Bentin, S., Bargai, N., & Katz, L. (1984). Orthographic and phonemic coding for lexical access: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 353-368.

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

Besner, D. (1987). Phonology, lexical access in reading, and articulatory suppression: A critical review. *Quarterly Journal of Experimental Psychology, 39A*, 467-478.

Besner, D. (1990). Does the reading system need a lexicon? In D. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 73-99). Hillsdale, NJ: Erlbaum.

Besner, D., Twilley, L., McCann, R. S., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review, 97(3)*, 432-446.

Brysbaert, M., Praet, C., & d'Ydewalle, G. (1990). *Phonological recoding in reading isolated words: A secondary and optional channel* (Tech. Rep. No. 110). Leuven: University of Leuven (Belgium), Department of Psychology.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). New York: Academic Press.

Conrad, R. (1979). *The deaf schoolchild.* London: Harper & Row.

Feldman, L. B., & Turvey, M. T. (1983). Word recognition in Serbo-Croatian is phonologically analytic. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 288-298.

Ferrand, X. & Grainger, J. (in press). Phonology and orthography in visual word recognition: Evidence from masked nonword priming. *Quarterly Journal of Experimental Psychology.*

Flores d'Arcais, G. B. (in progress). Research summarized in Max Planck Institute for Psycholinguistics Annual Report 1990. Nijmegen, The Netherlands.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*, 3-71.

France, I. (1989). *The role of post-lexical phonetic processes in reading comprehension.* Unpublished master's thesis, University of Pittsburgh, Pittsburgh, PA.

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.

Goldman, S. R., Hogaboam, T. W., Bell, L. C., & Perfetti, C. A. (1980). Short-term retention of discourse during reading. *Journal of Educational Psychology, 72*, 647-655.

Gough, P. B., & Hillinger, M. L. (1980). Learning to read: An unnatural act. *Bulletin of the Orton Society, 20*, 179-196.

Haber, L. R., & Haber, R. N. (1982). Does silent reading involve articulation? Evidence from tongue-twisters. *American Journal of Psychology, 95*, 409-419.

Hanson, V. L. & Fowler, C. A. (1987). Phonological coding in word reading: Evidence from hearing and deaf readers. *Memory & Cognition, 15(3)*, 199-207.

Hanson, V. L., Goodell, E. W., & Perfetti, C. A. (1991). Tongue-twister effects in the silent reading of hearing and deaf college students. *Journal of Memory and Language, 30*, 319-330.

Humphreys, G. W., Evett, L. J ., & Taylor, D. E. (1982). Automatic phonological priming in visual word recognition. *Memory & Cognition, 10*, 576-590.

Jarvella, R. J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior, 10*, 409-416.

Katz, L., & Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 157-166.

Koriat, A. (1984). Reading without vowels: Lexical access in Hebrew. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 227-242). Hillsdale, NJ: Erlbaum.

Koriat, A. (1985). Lexical access for low- and high- frequency words in Hebrew. *Memory & Cognition, 13,* 37-44.

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

Lam, A., Perfetti, C. A., & Bell, L. (1991). Automatic phonetic transfer in bidialectal reading. *Applied Psycholinguistics, 12,* 299-311.

Liberman, I. Y., & Shankweiler, D. (1979). Speech, the alphabet, and teaching to read. In L. B. Resnick & P.A. Weaver (Eds.), *Theory and practice of early reading, Vol. 2* (pp. 109-132). Hillsdale, NJ: Erlbaum.

Liberman, I. Y., & Shankweiler, D. (1991). Phonology and beginning reading: A tutorial. In L. Rieben & C. A. Perfetti (Eds.), *Learning to read: Basic research and its implications* (pp. 3-17). Hillsdale, NJ: Erlbaum.

Levy, B. A. (1977). Speech and meaning processes. *Journal of Verbal Learning and Verbal Behavior, 16,* 623-638.

Levy, B. A. (1978). Speech processing during reading. In A.M. Lesgold, J. W. Pellegrino, S. D. Fokkema, & R. Glaser (Eds.), *Cognitive psychology and instruction* (pp. 123-151). New York: Plenum Press.

Lukatela, G., Popadić, D., Ognjenović, P., & Turvey, M. T. (1980). Lexical decision in a phonologically shallow orthography. *Memory & Cognition, 8,* 124-132.

Lukatela, G., & Turvey, M. T. (1990a). Automatic and pre-lexical computation of phonology in visual word identification. *European Journal of Cognitive Psychology, 2,* 325-343.

Lukatela, G., & Turvey, M. T. (1990b). Phonemic similarity effects and prelexical phonology. *Memory & Cognition, 18,* 128-152.

Lukatela, G., Turvey, M. T., Feldman, L .B., Carello, C., & Katz, L. (1989). Alphabetic priming in bi-alphabetical word perception. *Journal of Memory and Language, 28,* 237-254.

Mattingly, I. G. (1972). Reading, the linguistic process, and linguistic awareness. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationships between speech and reading* (pp. 133-147). Cambridge, MA: MIT.

McCusker, L. X., Hillinger, M. L., & Bias, R. G. (1981). Phonological recoding and reading. *Psychological Bulletin, 89,* 217-245.

McCutchen, D., Bell, L., France, I. M., & Perfetti, C. A. (1991). Phoneme-specific interference in reading: The visual tongue-twister effect revisited. *Reading Research Quarterly, 26,* 87-103.

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.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of graphemic and phonemic codes in visual word recognition. *Memory & Cognition, 2,* 309-321.

Modern Chinese word frequency dictionary (1986). Beijing: Language Institute Press.

Naish, P. (1980). The effects of graphemic and phonemic similarity between targets and masks in a backward visual masking paradigm. *Quarterly Journal of Experimental Psychology, 32,* 57-68.

Navon, D., & Shimron, J. (1981). Does word naming involve grapheme-to-phoneme translation? Evidence from Hebrew. *Journal of Verbal Learning and Verbal Behavior, 20,* 97-109.

Paap, K. R., & Noel, R. W. (1991). Dual-route models of print and sound: Still a good horse race. *Psychological Research, 53,* 13-24.

Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and masked priming. *Journal of Memory and Language, 30,* 473-485.

Perfetti, C. A., Bell, L., & Delaney, S. (1988). Automatic phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language, 27,* 59-70.

Perfetti, C. A., & McCutchen, D. (1982). Speech processes in reading. In N. Lass (Ed.), *Speech and language: Advances in basic research and practice* (pp. 237-269). New York: Academic Press.

Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese words. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17,* 633-643.

Rumelhart, D. E., & McClelland, J. L. (1981). Interactive processing through spreading activation. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 37-60). Hillsdale, NJ: Erlbaum.

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., & McClelland, J. L. (1990). More words but still no lexicon: Reply to Besner et al. (1990). *Psychological Review, 97(3),* 447-452.

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.

Slowiaczek, M. L., & Clifton, C. (1980). Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior, 19,* 573-582.

Smith, F. (1985). *Reading without nonsense* (2nd. ed.). New York: Teachers College Press.

Turvey, M. T., Feldman, L. B., & Lukatela, G. (1984). The Serbo-Croatian orthography constrains the reader to a phonologically analytic strategy. In L. Henderson (Ed.), *Orthographies and reading: Perspectives from cognitive psychology, neuropsychology, and linguistics* (pp. 81-89). Hillsdale, NJ: Erlbaum.

Tzeng, O. J. L., & Hung, D. L. (1980). Reading in a nonalphabetic writing system: Some experimental studies. In J. F. Kavanagh & R. L Venezky (Eds.), *Orthography, reading and dyslexia.* Baltimore: University Park Press.

Tzeng, O. J. L, Hung, D. L., & Wang, W. S-Y. (1977). Speech recoding in reading Chinese characters. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 621-630.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition, 15,* 181-198.

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 371-385.

Van Orden, G. C., Pennington, B., & Stone, G. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97,* 488-522.

Xinhua Dictionary (1971). (*A dictionary of Chinese characters*). Beijing: Shang Wu Press.

Zhou, Y. G. (1978). To what degree are the "phonetics" of present-day Chinese characters still phonetic? *Zhongguo Yuwen, 146,* 172-177.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 14

"Assembled" Phonology and Reading: A Case Study in How Theoretical Perspective Shapes Empirical Investigation

Guy C. Van Orden
Gregory O. Stone
Karen L. Garlington
Lori R. Markson
Greta Sue Pinnt
Cynthia M. Simonfy
Tony Brichetto
Arizona State University, Tempe

What underlies the experience of reading a newspaper? Within contemporary psycholinguistics, the symbolic, human-information- processing framework assumes this experience to be the outcome of *recognition, retrieval,* and *validation* processes. Printed letter strings are recognized (classified) as tokens of familiar orthographic types (perhaps graphemes, morphemes, or words). These orthographic types are stored explicitly in memory. Recognition of orthographic types allows retrieval of corresponding meanings from lexical memory (and syntactic values). As a sentence's individual word-meanings accumulate, a validation process establishes *truth conditions* on whether proposition-size combinations of meanings refer to "true" conditions in a real world (Johnson-Laird, 1988). Valid combinations underlie typical, meaningful experience of text. Our present interest in this metaphor turns on recognition and retrieval processes. We claim a role for printed words' phonology in retrieving meanings, and we demonstrate that even a novel letter-string's (e.g., SLEAT's) phonology can be counted on to retrieve meaning from lexical memory.

Phonology's role in reading is controversial. However, at least, everyone agrees that a role for phonology is consistent with the symbolic approach. Of course, phonology's role is expressed in terms of symbolic processes of recognition, retrieval, and validation. Recognition of orthographic types retrieves phonologic types, and phonologic types may, in turn, retrieve corresponding meanings. Accommodation within the symbolic framework has shaped the search for phonology effects in reading performance. Methods of experiments always derive from an underlying theoretical metaphor (sometimes implicitly). In the present case, this is crucial because the symbolic framework failed to elicit phonology's role in reading.

We draw our perspective from an alternative metaphor in which representations are not stored explicitly in memory, and meaning is created on the fly (Stone & Van Orden, 1989; Van Orden, Pennington, & Stone, 1990, 1992; and compare Gerrig, 1989; Gerrig &

Murphy, 1991). Our perspective sometimes reveals novel methods for seeking phonology effects, and new methods are warranted.

Traditional methods are often focused too narrowly within the symbolic framework and often miss phonology effects. Subsequently, failures to observe phonology effects are taken as positive support for recognition processes that do not make reference to phonology. Such "nonphonologic" processes dominate most reading theories, even though their sole support comes from null effects. No positive empirical support exists, except positive interaction effects in which one condition shows a phonology effect and another does not (but cf. Paap & Noel, 1991).

Of course, recurring failures of a hypothesis, which show themselves in null findings, erode belief. Here, though, we also confront "nonphonologic" hypotheses leaning too heavily on null findings—to the extent that null findings must carry as much weight as opposite positive findings in discussions of phonology and reading. A reasonable line of attack then is to question the methods and perspective that yield these null effects.

Similarity, monotonicity, and reading performance. Most tests for phonology effects in reading derive from a central axiom of symbolic analysis. This *similarity axiom* concerns the similarity between stimuli and memory representations. A stimulus letter-string must be recognized as a familiar orthographic type prior to retrieval of meaning. Successful "bottom-up" recognition of a correctly spelled stimulus word (or any other familiar stimulus) relies on surface similarity between the stimulus and a memory representation—in this case, the orthographic type (Becker, 1976; Forster, 1979; Morton, 1969). Likewise, false positive recognition is a function of similarity. For example, false positive recognition might occur because SLEAT is highly similar to an orthographic type (SLEET) in lexical memory.

Similarity between stimuli and memory representations is assumed to produce monotonic effects on meaning-based performance in simple reading tasks (and other information-processing tasks). For example, when retrieval of meaning benefits performance, then increasing stimulus-memory similarity should benefit further. Of course, tasks could be constructed in which retrieval of meaning hurts performance. In this case, increasing similarity should increase the hurt. The *monotonicity hypothesis* does not set conditions on the direction of effects, it merely requires they continue in the same direction if similarity is increased. This hypothesis is certain for symbolic theories because recognition (classification) succeeds or fails as a function of similarity, and the retrieval of meaning requires successful recognition.

The monotonicity hypothesis provides a basis for testing whether phonology plays a role in reading. Manipulations of surface similarity between stimuli and lexical items should produce monotonic effects in simple reading tasks. For example, nonword stimuli such as SLEAT or SLERT can be constructed to vary in how similar they are to actual words. SLEAT and SLERT are equally similar in spelling to SLEET, but SLEAT is also identical in phonology. Subsequently, performance to SLEAT and SLERT can be compared in simple reading tasks. A phonology effect would be indicated if the increase in phonologic similarity causes a reliable monotonic change in some performance variable like accuracy or response time (RT). Especially, we are interested in effects consequent on SLEAT retrieving the meaning of SLEET. A few studies have reported such effects, but others have turned up null findings.

"Assembled" phonology and reading. This chapter will not review all the null phonology effects motivating nonphonologic reading processes (but see Van Orden et al., 1990). Rather, we confront a particular ongoing example of reasoning from null findings—in this case, null (or small) phonology effects to novel letter strings such as SLEAT. We confront the work of Veronika Coltheart and her colleagues who claim "assembled" phonology has small-to-null effects on reading comprehension (Coltheart, Avons, Masterson, & Laxon, 1991; Coltheart, Avons, & Trollope, 1990; Coltheart, Laxon, Rickard, & Elton, 1988).

"Assembled" phonology refers to transient phonology, computed (assembled) from a transient representation of the spelling features of a stimulus letter-string. In a traditional logic, assembled phonology is associated primarily with nonword letter-strings such as SLEAT, and is contrasted with actual words' "addressed" phonology stored explicitly in memory (Patterson & Coltheart, 1987). *Assembled* is quoted here and in our title because our perspective on reading makes no distinction between addressed and assembled phonology. We believe a massively interactive, adaptive, dynamic system codes word and nonword phonology (Van Orden, 1987; Van Orden et al., 1990, 1992), with multiple sources of constraint (Kawamoto & Kitzis, 1991; Kawamoto & Zemblidge, in press) which may discriminate between words and nonwords. But here, our purpose is merely to demonstrate reliable assembled (nonword) phonology effects on reading performance.

Coltheart and colleagues' null effects of assembled phonology are nested within a laudable body of positive findings. They first demonstrated that "exception word" phonology affects reading performance. Exception words are words like BEAR in which the pronunciation of _EA_ is an exception to the rule governing its pronunciation (as in BEAT). Previously, it was widely assumed exception words were read without reference to phonology, but Coltheart and her colleagues showed otherwise (Coltheart et al., 1988; 1991). Clearly then, these scientists have produced important positive findings. But their larger theoretical position restricts the role of assembled phonology, and this restriction derives exclusively from null findings. They are not alone, but they are prominent in that position and they spell out clearly the logic that lead them there.

The role of novel letter-strings' assembled phonology is controversial, in part, because contrasts between familiar versus novel stimuli confront key questions concerning the substrate of meaningful experience. For example, "Do familiar stimuli differ by degree or in kind from novel stimuli?" or "Do the same or different processes underlie (apparently) holistic recognition of familiar stimuli versus (apparently) analytic recognition of novel stimuli?" These questions have parallels in most areas of cognitive psychology, so it is not surprising positions on these questions run very deep.

For example, theories drawn on the symbolic metaphor must assume familiar stimuli differ in kind from novel stimuli. Familiar stimuli are represented explicitly in memory. In at least one reading task, we found word and nonword phonology affect performance equally (Van Orden et al., 1988). As noted, we believe this equal performance arises in a common dynamic process. We also believe meaning is created on the fly in the interaction of contextual and stimulus constraints. If meaning is created, not retrieved, then familiar stimuli need not be represented explicitly in memory. But, we do not focus here on the equality of word and nonword phonology. Rather, again, we merely seek reliable effects of nonword phonology to counter Coltheart et al.'s (1988, 1990, 1991) failures to demonstrate assembled phonology's role in skilled reading.

We report experiments that used three laboratory reading tasks. All produced large reliable effects of nonword phonology. If performance in laboratory tasks pertains to typical meaningful experience of text, then it should not be peculiar to a single task.

Ideally, qualities of performance may correlate across tasks to the extent tasks engage a common basis. Other investigators failed to observe correlated effects because they applied the similarity axiom too narrowly, and because the monotonicity hypothesis is sometimes false within that narrow focus. (The latter phenomenon was observed in all three proofreading experiments reported here.) Our chapter is organized by sections named for the reading tasks under discussion—beginning with the sentence verification task. The null phonology effects we confront came from experiments using this task (see Coltheart et al., 1988, 1990, 1991), and we describe those experiments in the next section. Later sections report new lexical decision, semantic categorization, and proofreading experiments.

Sentence verification experiments

Coltheart and her colleagues have found small-to-null effects of nonword phonology in several sentence verification experiments. Subjects were presented with individual sentences and instructed to respond "yes" if they were "appropriately spelled correct English sentences that made sense" (p. 390, Coltheart et al., 1988; p. 389, Coltheart et al., 1991), and "no" otherwise. The logic of these experiments derives simply and elegantly from the similarity axiom and the monotonicity hypothesis.

Concerning assembled phonology, Coltheart and colleagues' key contrast was between stimuli such as HER BLOO DRESS WAS NEW versus THE SKY IS BLOE TODAY. BLOO and BLOE are both similar in spelling to <u>BLUE</u>, but BLOE is not identical to <u>BLUE</u> in phonology. If collective similarity effects are monotonic, and if phonologic identity affects performance over-and-above mere phonologic similarity, then subjects should be more likely to falsely recognize BLOO as <u>BLUE</u> than BLOE as <u>BLUE</u>. False recognition will cause false positive retrieval of a meaning appropriate to <u>BLUE</u>. In turn, false retrieval of <u>BLUE</u>'s meaning will cause false positive semantic validation of HER BLOO DRESS WAS NEW, resulting in a false positive sentence verification error. Thus, if assembled phonology affects reading, then subjects should make more false positive sentence verification errors to HER BLOO DRESS WAS NEW than to THE SKY IS BLOE TODAY. Coltheart et al. (1988) observed a small effect of BLOO's phonology, but Coltheart et al. (1990; 1991) observed null effects.

As with all laboratory tasks, there are advantages and disadvantages to the sentence verification task. On the plus side, judging whether a sentence makes sense may fully engage recognition, retrieval, and validation processes because most stimuli are actual sentences. But we cannot be assured of "normal" processing on trials containing nonword stimuli.

For example, a correct "no" response to the stimulus HER BLOO DRESS WAS NEW might arise in several ways: Recognition could identify orthographic types appropriate to the words HER, DRESS, WAS, and NEW, but recognition could successfully reject BLOO because it does not match the familiar orthographic type <u>BLUE</u>. Rejection of BLOO may occur prior to retrieval of meaning and other lexical information, on the basis of familiarity alone (compare Balota & Chumbley, 1984), or because BLOO is simply not similar enough to trigger false recognition of <u>BLUE</u>. However, it is no less likely BLOO is first recognized as <u>BLUE</u>, on the basis of phonology, but an evaluation of stimulus spelling familiarity or specific knowledge of <u>BLUE</u>'s spelling detects the impostor. A meaning of <u>BLUE</u> could have been retrieved from lexical memory along with other <u>BLUE</u> knowledge, but subsequent processing saves performance from a false positive sentence verification error. In this latter case, the subject would not produce a false positive error, even though BLOO's assembled phonology caused false positive recognition of BLUE in the course of processing.

Coltheart et al. (1988; 1990; 1991) accept an outcome of processing—a correct "no" response in a sentence verification task—as a transparent indicator of activation prior to that outcome. Consequently, when the false positive error rate to BLOO-type foils is no greater than the error rate to BLOE-type foils, they conclude BLOO did not retrieve a meaning of BLUE. Further, they extrapolate from this null effect and conclude assembled phonology does not affect reading comprehension generally. But, the outcome "correct rejection" at the end of a sentence verification trial is silent as to whether a meaning of BLUE was retrieved in the course of rejecting the stimulus BLOO. Retrieval of meaning is the core issue here. Typical reading is not centered on discriminating familiar or correctly spelled sentences from foils, it centers on meaningful experience of text. But in laboratory reading tasks, processing past the point of recognition and retrieval can be affected as much by the artificial requirements of the task, as by the typical requirements of reading comprehension (compare Balota, 1990).

We are not persuaded by null effects in the sentence verification task that assembled phonology does not affect reading comprehension. The null evidence is especially suspect because it comes from a "lengthy stimulus-presentation-time" reading task allowing plenty of time for strategic processing. Also, in contradiction to Coltheart et al.'s (1988; 1991) null findings, experiments using a semantic categorization task reliably yield equal positive effects of word (addressed) and nonword (assembled) phonology (Van Orden et al., 1988; Coltheart et al., 1991). The semantic categorization task may be less subject to strategies, if only because target letter strings are presented individually, and subjects respond very quickly. (We will discuss these results in the section concerning the semantic categorization experiment.)

A lexical decision experiment

Subjects in a lexical decision experiment judge whether individually presented letter strings are words. They respond "word" to letter strings that are words, and "nonword" otherwise. We used the lexical decision task to test whether nonword stimuli falsely retrieve lexical memories of words similar in orthography and phonology. Our method detects whether correct rejection of a nonword foil like SLEAT included retrieval of lexical memory for SLEET. Specifically, this method tests for a frequency effect—a signature of lexical retrieval (Forster & Chambers, 1973)—upon the pattern of correct rejections to nonword letter strings.

For example, GREAN is similar to a high-frequency "base word" GREEN, and SLEAT is equally similar to a low-frequency base word SLEET. If more correct rejections are made to GREAN than to SLEAT, then familiarity with the base words GREEN and SLEET affects performance. But the base words are never presented in the experiment. A manipulation of base-word familiarity affecting performance to unfamiliar nonword stimulus foils dissociates base-word familiarity (memory) from stimulus familiarity (compare McCann & Besner, 1987).

A *base word frequency effect* to a nonword foil indicates false positive retrieval of base word lexical knowledge. This effect may be found even though "false retrieval" ends in a correct response on most trials. Consequently, a base word frequency effect indicates processing included retrieval of base word lexical knowledge, and the outcome "correct response" was based on this knowledge. Because the locus of the base word frequency effect must be lexical memory, we may use the magnitude of this effect to compare the capacities of different nonwords to retrieve lexical knowledge (compare Sánchez-Casas, García-Albea, & Bradley, 1991; Van Orden, 1991).

For the lexical decision experiment, we constructed pseudohomophone foils such as GREAN and SLEAT identical in phonology to high frequency or low frequency base words such as GREEN and SLEET, respectively, and also similar in spelling. Likewise, spelling control foils such as GREWN and SLERT were constructed to be as similar in spelling to base words as the previous yoked pseudohomophones, but not identical in phonology. We also included bigram control foils such as SHEAL and FLESS that did not resemble base words, but were composed of bigrams (letter pairs) as frequent across words in general as the bigrams composing pseudohomophone and spelling control foils. Bigram control foils bear a superficial, statistical resemblance to words' orthographic types and control for this aspect of other foils, but bigram controls are not similar to base words in spelling or phonology. Each resulting yoked foil triplet comprises a rank order of similarity to a base word. Pseudohomophones are more similar to base words than spelling controls, and spelling control foils are more similar to base words than bigram control foils.

We can extend the similarity and monotonicity hypotheses to the base word frequency effect. If phonologic identity affects recognition over phonologic similarity, then pseudohomophones should have enhanced capacity for causing a base word frequency effect. In this case, the base word frequency effect equals accuracy (percent correct rejections) to foils such as GREAN minus accuracy to foils such as SLEAT. Of course, spelling controls may also produce a base word frequency effect because they are also similar to base words. However, spelling controls are less similar in phonology so their capacity to produce this effect may be less than the capacity of pseudohomophones. Bigram controls should not produce a base word frequency effect because they bear no systematic relationship to base words.

We may also apply the similarity axiom and monotonicity hypothesis in the traditional fashion. They predict false-positive word recognition to the extent foils resemble words in lexical memory. If phonologic identity affects performance over similarity in phonology and spelling, then GREAN and SLEAT should produce fewest correct rejections because they are most likely to be falsely recognized as base words. GREWN and SLERT should produce more correct rejections, and the best performance should be seen to bigram controls such as SHEAL and FLESS. Of course, if assembled phonology does not affect reading, then spelling controls and pseudohomophones should affect performance equally.

Method

Subjects. The subjects were 19 undergraduates attending Arizona State University who received course credit for their participation. All were native English speakers.

Procedure. Subjects were instructed in the lexical decision task prior to 20 practice trials. No pseudohomophones were presented in the practice trials. After the block of practice trials subjects asked any remaining questions of the experimenter before beginning a block of 180 experimental trials. Each trial began with a fixation stimulus "+" visible for 500 ms, and replaced subsequently by a target stimulus such as SLEAT or FARM. Target stimuli remained visible until the subject responded. The next trial began with the fixation stimulus after an intertrial interval of 500 ms. Trials did not include feedback. Stimuli were centered on a standard monochrome monitor. They appeared in white capital letters on a black background. Responses were collected from a standard IBM keyboard. Subjects responded "word" by pressing the lower right "/" key on the keyboard. Subjects responded "nonword" by pressing the "z" key.

Stimuli. For each of 20 base words, we constructed three yoked stimulus foils: (1) a pseudohomophone foil (SLEAT) identical in phonology to the base word, (2) a spelling control foil (SLERT) as similar in spelling to the base word as its yoked pseudohomophone, and (3) a bigram control foil (FLESS) constructed from letter-pairs of approximately the same frequency as those of yoked pseudohomophone and spelling control foils. These stimuli are listed in Appendix A.

Base-word frequency was manipulated using the Kučera and Francis (1967) word frequency count. Half (10) of the base words had high frequency counts ranging from 27 to 348 per million (M=127, SD=96), and half had low frequency counts ranging from 1 to 16 per million (M=7, SD=4). Each pseudohomophone foil constructed from a low-frequency base word was yoked to a pseudohomophone foil constructed from a high-frequency base word. All ten yoked pairs were formed by spelling changes preserving identical local phonology, and eight pairs were formed by the same spelling change in their respective base words. For example, SLEAT is constructed from low frequency SLEET by substituting an A for an E; likewise, GREAN is constructed from high frequency GREEN by substituting an A for an E, preserving the same long-E phonology. We constructed stimuli using a variety of focal orthographic-phonologic correspondences.

To insure the homophony of the pseudohomophones, we used the method of Van Orden et al. (1988). Ten "judges" (different students from our subject population) read aloud quickly a list of pronounceable, nonhomophonic nonwords into which we had inserted randomly candidate pseudohomophones. We use a ratio of about seven-to-one to reduce the chance subjects would seek word names for the nonwords. To be a pseudohomophone, a foil must have been pronounced to match its base word by at least nine judges and any "mispronunciations" must have included misperception of the letters in the pseudohomophone (e.g., mispronouncing SLEAT as SLEEP).

Spelling similarity between pseudohomophone foils and base words, and between spelling controls and base words, was equated using OS, an estimate of orthographic similarity. OS ranges from 0 to 1, where 1 indicates identical spellings. OS was adapted from an estimate in Weber (1970) and is described in Van Orden (1987) and Van Orden et al. (1988). Essentially, OS is computed from the number of letters and letter pairs two letter strings share, giving special emphasis to beginning and end letters. Mean OS between low-frequency base words and corresponding spelling controls was 0.77, SD=.05; between low-frequency base words and corresponding pseudohomophone foils M=0.75, SD=.08; between high-frequency base words and corresponding spelling controls M=0.77, SD=.04, and between high-frequency base words and corresponding pseudohomophone foils M=0.75, SD=.06. Means for OS computed between bigram control foils and base words show these foils are low in spelling similarity to base words ("high-frequency" bigram foils M=.13, SD=.15; "low-frequency" bigram foils M=.14, SD=.10).

Please note, our estimate of orthographic similarity does not isolate this dimension. Stimuli similar in spelling are also similar in phonology, so spelling controls confound phonologic and orthographic similarity to base words. This does not qualify their control function against the effect of pseudohomophones' phonologic identity to base words, but it rules out detecting a "pure" effect of orthographic similarity.

Bigram frequency was controlled using the bigram frequency count of Massaro, Taylor, Venezky, Jastrzembski, and Lucas (1980). The mean log frequency of bigrams composing pseudohomophones was 2.561, SD=.81, spelling controls M=2.707, SD=.563, and bigram controls M=2.858, SD=.5574. A final control took into account N, the number of words formed from a letter-string by changing one letter. Mean N for "high-frequency" pseudohomophones was 2.9, SD=2.9; for low-frequency pseudohomophones M=3.1, SD=2.1; for high-frequency spelling controls M=5.4, SD=4.6; and for low-frequency spelling controls M=4.5, SD=3.3. Bigram controls had the lowest average N: high-frequency bigram controls M=1.7, SD=1.5; low-frequency bigram controls M=1.9, SD=1.7. Notice OS, bigram frequency, and N, are all slightly higher for spelling controls than for pseudohomophones; this insures these factors work against pseudohomophones' phonology effect (and against the similarity rank order).

In the above method, each high-frequency pseudohomophone is yoked to a low-frequency pseudohomophone. Every pseudohomophone is yoked, in turn, to a spelling control and bigram control foil yielding a yoked-stimulus-sextet item sampling unit.

Results and discussion

The results of the lexical decision experiment are presented in Table 1. The primary dependent variable was the percentage of correct rejections, the secondary variable was correct "no" response times (RTs). The two independent variables in the omnibus ANOVAs were foil type (pseudohomophones versus spelling controls versus bigram controls) and base word frequency (high versus low). ANOVAs were computed considering both subjects and items as random factors. The item sampling unit for omnibus analyses was a foil-sextet of two yoked pseudohomophones and their corresponding yoked control foils. However, tests for traditional effects used individual items as in previous traditional analyses. All descriptive statistics presented in tables come from subject analyses; item data are presented in Appendix A. The alpha level for all experiments was $p<.05$.

Table 1

Percentage of correct rejections (correct "no" responses) and correct "no" RTs (in parentheses) to pseudohomophone, spelling control, and bigram control foils in the lexical decision experiment.

	Pseudohomophones	Spelling Controls	Bigram Controls
High Frequency Base Words	78.7 (675)	77.4 (672)	92.1 (629)
Low Frequency Base Words	61.8 (690)	80.3 (655)	91.1 (643)
Mean	70.3 (682)	78.8 (664)	91.6 (636)
Difference (Base Word Frequency Effect)	16.9	-2.9	1.0

Traditional analysis. The highest rate of correct rejections was to bigram controls (91.6%, SE=1.5), followed by spelling control foils (78.8%, SE=2.5), and accuracy was lowest to pseudohomophones (70.3%, SE=3.1). This main effect of foil type on correct "no" responding was significant in both subject and item ANOVAs ($F(2,36)=33.76$ for subjects, $F(2,18)=9.02$ for items). The rank order of mean, correct, "no," response times corresponding to the main effect of foil type ($F(2,36)=4.55$ for subjects, $F(2,18)=7.24$ for items) also respects the rank order of similarity. The mean, correct, "no" RT to pseudohomophones was 682 ms (SE=20), for spelling controls M=664 (SE=17), and for bigram controls M=636 (SE=13).

Planned traditional comparisons were conducted on the column means of Table 1. The difference between pseudohomophones and spelling controls (8.6%) in percent correct rejections was significant by subjects ($t(18)=2.98$), but not by items ($t(19)=1.60$, $p=.13$). The difference between pseudohomophones and spelling controls (19 ms) in correct "no" RT was not significant in either analysis ($t(18)=1.26$, $p=.22$ for subjects, $t(19)=1.50$, $p=.15$ for items). By itself, this combination of results suggests nonword phonology has a small-to-null effect on word recognition. But, we will soon see the real failure here is one of method. The traditional method is merely focused too narrowly, and misses effects outside its sphere.

The traditional method does, however, detect the large difference in similarity between spelling controls and bigram controls. The effect of this difference on accuracy (12.8%) was significant by subjects ($t(18)=6.23$) and by items ($t(19)=2.72$), and the effect on correct "no" RTs approached significance by subjects ($t(18)=2.06$, $p<.06$) but not by items ($t(19)=1.41$, $p=.17$). (No other effects were found in RT data, all *F*s were less than 1.) It may be tempting here to attribute the performance differences between spelling controls and bigram controls to orthographic similarity alone. But, this inference cannot be trusted because spelling controls confound phonologic and orthographic similarity to base words. The literature concerning phonology in reading has many examples in which authors have failed to note this confound or forgotten it when it came time to interpret their data. Spelling controls typically confound phonologic and orthographic similarity and yet contrasts between spelling controls and other foils are typically and unjustifiably interpreted as orthographic effects.

Base word frequency effects. The omnibus ANOVAs for accuracy data revealed a significant interaction between foil type and base word frequency by subjects ($F(2,36)=8.50$) and this effect approached significance in the item analysis ($F(2,18)=2.76$, $p<.09$). Planned analyses of simple main effects reveal the locus of this interaction. More correct rejections were made to pseudohomophones (GREAN) constructed from high-frequency base words (78.7%, SE=3.8) than to pseudohomophones (SLEAT) constructed from low-frequency base words (61.8%, SE=4.3). This base word frequency effect (16.9%) was significant by subjects ($t(18)=3.83$) and by items ($t(9)=3.36$). In contrast, accuracy to spelling controls constructed from high-frequency base words (77.4, SE=2.8) did not differ significantly from accuracy to those constructed from low-frequency base words (80.3%, SE=4.1; $t<1$ for subjects and items). And, as expected, accuracy to bigram controls corresponding to high frequency base words (92.1%, SE=2.2) was essentially the same as accuracy to bigram controls corresponding to low frequency base words (91.1%, SE=2.1; $t<1$ for subjects and items).

Apparently, at least two sources of constraint affect accuracy to foils in the lexical decision task. One source reduces accuracy to spelling control foils relative to bigram control foils. But this source does not also produce a base word frequency effect to spelling control foils. So, we lack evidence processing of spelling controls included retrieval of base word lexical knowledge. Another possibility is that phonologic and orthographic similarity of spelling controls to words in general is the source of errors to these foils—not their similarity to particular base words. We computed correlations between accuracy and correct RTs to spelling control foils and (a) mean bigram frequency—an estimate of general similarity, (b) N=the number of words that differ by one letter from a spelling control—another estimate of general similarity, and (c) OS—an estimate of specific similarity to base words. These analyses failed to support a "general similarity hypothesis."

Neither bigram frequency nor N were correlated significantly with lexical decision performance to spelling controls. This was true for accuracy and correct "no" RTs (all ts<1, except the correlation between bigram frequency and "no" RTs where r=.24, $t(18)$=1.07, p=.30). These null findings should be interpreted cautiously, but a similar analysis found a significant correlation between OS (orthographic and phonologic similarity to base words) and accuracy (r=.61, $t(18)$=3.28), and between OS and "no" RTs (r=.50, $t(18)$=2.43). Apparently, recognition is affected by the phonologic and orthographic similarity of spelling controls to base words. However, lacking a base word frequency effect we do not know whether this effect includes retrieval of base word lexical knowledge.

Only pseudohomophones produce a base word frequency effect. We believe this effect emerges in a verification process (Van Orden, 1987, 1991). The most general form of the verification hypothesis assumes expectations of a stimulus surface form are verified in the process of recognition. We have added specific assumptions to explain the base word frequency effect. Surface expectations of familiar words are more readily and completely available than expectations of less familiar words (Van Orden, 1987, 1991). Consequently, pseudohomophones (GREAN) constructed from familiar words, are more readily rejected than those (SLEAT) constructed from less familiar words. But, for present purposes, our explanation of base word frequency effects is less pertinent than another conclusion: Pseudohomophones' phonology reliably retrieves knowledge of base words from lexical memory, and this knowledge affects accuracy in the lexical decision task.

Our previous discussion of spelling control data requires a caveat. We did not guarantee "high-frequency" and "low-frequency" spelling controls were equal in phonologic similarity. This lack of control over phonologic similarity could contaminate any analysis contrasting these two classes of spelling controls. For example, spelling control foils may fail to produce a base word frequency effect because "high-frequency" spelling controls are more similar in phonology to base words, and this phonologic similarity effect counters the effect of base word frequency. But, if this were true, it would merely strengthen our contention that assembled phonology affects reading performance. Consequently, although a potential confound between phonologic similarity and other variables does qualify the conclusions that can be drawn from spelling control data, it does not qualify our demonstration of an assembled phonology effect.

A categorization experiment

Subjects in the lexical decision task judged whether individually presented letter strings were words. Ideally, the lexical decision task isolates word recognition due to the isolated presentation of letter strings, and because words merely need be recognized for correct performance. But not surprisingly, perhaps, the lexical decision task falls short of this ideal. It requires that words be discriminated from nonwords, not just recognized. Thus, effects thought to originate in recognition and retrieval are confounded with effects of (possibly) task-specific processes of discrimination (Balota, 1990).

Worse yet, even in the ideal, the lexical decision task does not force strong conclusions concerning retrieval of meaning. Lexical decisions do not require words' meanings. Of course, meaning can affect lexical decision performance—we know this from semantic priming, where prior presentation of DOCTOR facilitates lexical decision performance to NURSE. Also, of course, retrieval of meaning is a sufficient basis for lexical decisions, because words have conventional meanings and nonwords do not. But, retrieval of meaning may not be a necessary basis for lexical decisions.

Although the base word frequency effect indicates lexical knowledge affects performance, the knowledge underlying this effect need not be correlated with retrieval of meaning. For example, suppose this effect arises in processes specific to word-nonword discrimination, and these processes rely on spelling knowledge, exclusively. One explanation of Coltheart et al.'s null effects of BLOO's phonology included retrieval of base word BLUE's spelling for use in rejecting BLOO's spelling. It is possible verification of BLOO's spelling serves word-nonword discrimination, but not word identification for reading.

We do not espouse this possibility. We believe verification is fundamental to recognition (and it is controlled strategically by parametric changes in the relative force of top-down versus bottom-up constraints, compare Stone & Van Orden, 1989, 1992). But, we cannot yet determine whether verification is specific to reading tasks requiring target/foil discriminations (Jared & Seidenberg, 1991; Van Orden, 1987). Likewise, we cannot determine whether pseudohomophones in the lexical decision task caused base words' spellings to be retrieved, but not their meanings.

We would be better assured that assembled phonology retrieves meaning if pseudohomophones produced a base word frequency effect in a task requiring meaning. We answer this concern with a semantic categorization experiment. In the semantic categorization task, subjects are presented with a category name such as WEATHER followed by a target such as RAIN or CHAIR. Subjects respond "yes" if the target is an exemplar of the category and "no" otherwise. Of course, this task comes with its own discrimination requirements. But the task emphasizes a semantic discrimination between exemplars and nonexemplars, and this discrimination is based on meaning.

Van Orden et al. (1988) confirmed assembled phonology's effect on reading in a categorization experiment. They presented high school student subjects with foils such as SLEAT and SLERT (for the category WEATHER). Correct rejection rates to pseudohomophone foils (SLEAT, 78.7%, SE=2.7) were much lower than to spelling controls (SLERT, 97%, SE=1.3, from Van Orden et al., 1988, Experiment 1). Coltheart et al. (1991) replicated the difference in accuracy between pseudohomophones (88.6%) and spelling controls (98.1%), but their college student subjects made fewer errors overall (see also Jared & Seidenberg, 1991). Within the traditional logic, then, categorization performance to pseudohomophones clearly demonstrates a reliable, positive, assembled phonology effect.

Van Orden et al. (1988) also observed equal effects of word homophone and pseudohomophone phonology in two experiments. They chose yoked pairs of word homophones (BEATS for the category VEGETABLE) and pseudohomophones (SLEAT for the category WEATHER) to be equal on dimensions affecting categorization performance. In two experiments, using different foils and subjects, yoked homophonic foils produced virtually identical rates of correct rejections (for Experiment 1, BEATS=78.2%, SE=2.6, SLEAT=78.7%, SE=2.7; for Experiment 2, BEATS=67.2%, SE=3.0, SLEAT=67.5%, SE=3.3). Since then, we have replicated this equality with three other groups of subjects including adult developmental dyslexics (means are reported in Van Orden et al., 1990). Coltheart et al. (1991) also replicated this virtual equality (BEATS=9.5%, SLEAT=11.4%, and see Jared & Seidenberg, 1991).

We are still slightly amazed by this unexpected finding. The possibility that a key cognitive system treats familiar and novel stimuli equally in a task assuring meaning evaluation chips at the foundation of the symbolic framework. Familiar stimuli should differ in kind from unfamiliar stimuli because familiar stimuli are stored explicitly in memory. Processing should always respect this difference. Any equal effects must be tracked to artifacts that trade-off with familiarity (an unlikely possibility in this case, see Van Orden et al., 1988). Or, in the case of word identification, equal effects must be explained in a symbolic rule-governed process, *blind to the familiarity of word-size letter combinations* (also unlikely, see Van Orden et al., 1990). This rule governed process must compute phonologic codes of both words and nonwords, and these phonologic codes must be the sole basis for retrieval of meaning.

Here though, we are less concerned with the specific character of phonologic coding than with an additional demonstration of assembled phonology effects. In the present categorization experiment, we used the same pseudohomophone, spelling control, and bigram control foils from the lexical decision task. This allowed us to contrast lexical decision and categorization performance. Of course, it also allowed us to examine performance for base word frequency effects as well as traditional effects upon overall error rates. Van Orden et al. (1988) did not manipulate base word frequency. If we find a base word frequency effect to pseudohomophone stimuli in the categorization task, a task emphasizing meaning evaluation, then we are better assured the base word frequency effect in lexical decision is correlated with retrieval of meaning.

Method

Subjects. The subjects were a new group of 19 undergraduates from the same population as the previous experiment.

Procedure. Subjects were instructed in the semantic categorization task prior to 40 practice trials. The practice trials did not include nonword foils. After the block of practice trials, subjects asked remaining questions of the experimenter before beginning a block of 200 experimental trials. Each trial began with the presentation of a category name (WEATHER) for 1500 ms, followed by a fixation stimulus "+" visible for 500 ms, and replaced immediately by a target stimulus (e.g., RAIN, SLEAT, or CHAIR). The target stimulus remained visible until the subject responded. After 500 ms, the next trial began with presentation of another category name. Trials did not included feedback.

Stimuli were centered on a standard monochrome monitor. They appeared in white letters on a black background. Category names were presented in lower case and targets were presented in capital letters. Responses were collected from a standard IBM keyboard. Subjects responded "yes, exemplar" by pressing the "/" key on the lower right of the keyboard. Subjects responded "no, nonexemplar" by pressing the "z" key on the lower left of the keyboard.

Stimuli. The 20 pseudohomophone, 20 spelling control, and 20 bigram controls foils were the same as those used in the lexical decision task. Category names were chosen appropriate to the base words (e.g., WEATHER for SLEET, or COLOR for GREEN), and filler categories, targets, and foils were chosen for 20 practice and 100 filler "yes" trials (exemplar targets) and 20 practice and 40 filler "no" trials (word nonexemplar foils).

Results and discussion

The results of the semantic categorization experiment are presented in Table 2; item data are presented in Appendix A. Again, the primary dependent measure was the percentage of correct rejections, and the secondary measure was correct "no" RTs. The two independent variables were foil type (pseudohomophones versus spelling controls versus bigram controls) and base word frequency (low versus high).

Traditional analysis. As in the lexical decision experiment, accuracy was greatest to bigram controls (97.8%, SE=0.8), followed by spelling controls (90.5%, SE=1.6), and pseudohomophones produced the lowest rate of correct rejections (64.2%, SE=4.4). This main effect of foil type was significant in both subject and item ANOVAs ($F(2,36)=48.79$ for subjects; $F(2,18)=26.73$ for items), as was the concomitant RT effect (bigram controls M=740 ms, SE=24, spelling controls M=788 ms, pseudohomophones M=802 ms, SE=25, SE=24, $F(2,36)=8.48$ for subjects; $F(2,18)=7.56$ for items). Once again, the rank orders of mean performance respect foils' similarity to base words, confirming the similarity axiom and the monotonicity hypothesis.

The planned comparisons of column means in Table 2, found a large difference (26.3%) between accuracy to pseudohomophone versus spelling control foils ($t(18)=6.29$ for subjects; $t(19)=4.63$ for items). However, the difference (14 ms) between correct "no" RTs to pseudohomophones versus spelling controls was not reliable ($t<1$ in both analyses). The difference between accuracy to Table 2 spelling controls versus bigram controls (7.2%) was also reliable ($t(18)=3.78$ for subjects, $t(19)=3.50$ for items), as was the corresponding difference (48 ms) in correct "no" RTs ($t(18)=3.79$ for subjects, $t(19)=2.65$ for items).

Table 2

Percentage of correct rejections and correct "no" response times (in parentheses) to pseudohomophone, spelling control, and bigram control foils in the semantic categorization task.

	Pseudohomophones	Spelling Controls	Bigram Controls
High Frequency Base Words	81.1 (791)	93.7 (768)	98.4 (731)
Low Frequency Base Words	47.4 (812)	87.4 (808)	97.1 (749)
Mean	64.2 (802)	90.5 (788)	97.8 (740)
Difference (Base Word Frequency Effect)	33.7	6.3	1.3

This pattern parallels results from the lexical decision task except for the large difference in categorization accuracy between pseudohomophones versus spelling controls. The corresponding difference from the lexical decision task (Table 1) was small and not quite reliable. Following the traditional logic, we might conclude word identification in the semantic categorization task is constrained by stimulus phonology, but much less so in the lexical decision task. However, it is also possible lexical decision errors to pseudohomophones included retrieval of base word meanings, while errors to spelling controls did not. Subsequently, in the categorization task, spelling control foils produce a higher rate of correct rejections because they did not retrieve base word meanings necessary to cause semantic categorization errors. This latter possibility agrees with the pattern of base word frequency effects.

Base word frequency effects. The omnibus ANOVAs on accuracy data revealed a significant interaction between foil type and base word frequency ($F(2,36)=25.03$ for subjects, $F(2,18)=24.60$ for items). Analyses of simple main effects confirmed pseudohomophones' larger capacity for base word lexical retrieval. The base word frequency effect on accuracy to pseudohomophone foils (33.7%) was significant by subjects ($t(18)=6.48$) and by items ($t(9)=6.39$). But the corresponding base word frequency effect on correct "no" RTs (21 ms advantage for "high-frequency" pseudohomophones) was not significant ($t<1$ for subjects; $t(9)=1.10$, $p=.30$ for items). The base word frequency effect to spelling controls (6.3%) was significant by subjects ($t(18)=3.08$), but not by items ($t(9)=1.41$, $p=.19$). Similarly, in RT analyses, the 40 ms advantage for "high frequency" spelling controls approached significance for subjects ($t(18)=2.08$, $p<.06$), but not for items ($t(9)=1.43$, $p>.18$). Spelling controls did not show a base word frequency effect in the lexical decision task. So, the trend toward a base word frequency effect in the categorization task may stem from an interaction between a few items' phonologic and orthographic similarity to base words, and the context provided by a related category name (compare Van Orden, 1991).

Even in a strong biasing context, spelling controls do not show a reliable capacity for retrieving lexical knowledge of base words. But once again pseudohomophones produce a large reliable base word frequency effect, this time in a task that emphasizes meaning evaluation. Because pseudohomophones produce base word frequency effects in both lexical decision and semantic categorization tasks, we claim pseudohomophones retrieve the meanings of base words in both tasks, biasing context or not.

Lexical decision versus categorization. We may contrast the data from the previous two experiments because the same foils were used in both. This contrast focuses on accuracy data to pseudohomophone and spelling control foils. The only effect of task on bigram foils is a larger correct rejection rate in the categorization task (97.8% versus 91.6%, $F(1,36)=9.00$ for subjects, $F(1,9)=8.47$ for items). No interaction was found between task or base word frequency for bigram foils (all Fs<1).

As for "no" RT data, the only reliable effects were main effects of task and foil type. Correct "no" RTs in the categorization task (777 ms, SE=14) were slower than in the lexical decision task (661 ms, SE=10; $F(1,36)=9.16$ for subjects; $F(1,9)=143.86$ for items). Also correct "no" RTs respected the rank order of stimulus similarity to base words in both tasks with fastest RTs to bigram controls (688 ms, SE=15), next fastest to spelling controls (726 ms, SE=16), and slowest to pseudohomophones (742 ms, SE=17; $F(2,72)=12.62$ for subjects; $F(2,18)=9.53$ for items), but, as in the separate task-analyses, the foil effect is only reliable between "no" RTs to spelling controls versus bigram controls.

For accuracy data, the three-way interaction of task (lexical decision versus categorization) × foil type (pseudohomophones versus spelling controls) x base word frequency (high versus low) was not significant (subject $F < 1$; item F=1.04) so we may focus on two-way interactions and main effects. Of course, foil type and base word frequency interacted significantly ($F(1,36)$=36.15 for subjects, $F(1,9)$=10.90 for items). We expected this effect because the previous analyses found substantially larger base word frequency effects to pseudohomophone foils.

Task interacted significantly with foil type ($F(1,36)$=12.25 for subjects; $F(1,9)$=13.85 for items). This confirms the difference between accuracy to pseudohomophones and spelling controls (8.5%) in the lexical decision task is smaller than the corresponding difference (26.3%) in the categorization task. Task also interacted significantly with base word frequency ($F(1,36)$=10.59 for subjects; $F(1,9)$=15.74 for items). We can better understand these interaction effects by examining the effects of task x frequency, taking pseudohomophone and spelling control data separately.

For spelling controls, the interaction between task and frequency is significant by subjects ($F(1,36)$=5.17), but not by items ($F(1,9)$=2.77, p>.13). This confirms the previous pattern of no base word frequency effect in the lexical decision task and an effect by subjects only in the categorization task. A main effect of interest is the significant rise in accuracy to spelling controls from the lexical decision task (78.8%) to the categorization task (90.5%). It can be difficult to interpret main effects across tasks due to differing task demands—perhaps the categorization task is merely "easier." However, in the present case, this main effect is part of a larger interaction effect, so we may interpret it in relation to the opposite trend to pseudohomophones.

The analysis of pseudohomophones confirmed the main effect of base word frequency found previously ($F(1,36)$=55.08 for subjects; $F(1,9)$=30.45 for items). The main effect of task was not significant ($F(1,36)$=1.24, p>.27 for subjects; $F(1,9)$=2.43, p>.15 for items). However, the interaction between task and base word frequency is significant ($F(1,36)$=6.12 for subjects; $F(1,9)$=12.89 for items). Accuracy to pseudohomophones of high frequency base words was about the same in the two tasks, but accuracy to pseudohomophones of low-frequency base words fell from 61.8% in the lexical decision task to 47.4% in the categorization task.

Across tasks, accuracy to spelling controls goes up in the categorization task while accuracy to "low-frequency" pseudohomophones goes down! This pattern is puzzling. Ideally, lexical decisions are based on the output of word recognition processes, and semantic categorization is based on the output of word identification—including stimulus recognition and retrieval of meaning. So, to explain this result, we might appeal to differences between the tasks affecting recognition or retrieval processes, respectively.

Concerning the spelling control data, we assume false positive errors to spelling controls in the lexical decision task did not include false positive retrieval of base words' meanings. Consequently, because the categorization task bases performance on meaning evaluation, accuracy rises to spelling controls. But the spelling controls produced a small base word frequency effect in categorization performance (significant by subjects), a trend toward retrieval of base words' meanings. If category context is the source of this trend, then we confront apparently contradictory effects of context on accuracy to spelling controls. Context biases processing of spelling controls toward base word lexical knowledge—presumably, a bias toward retrieval of base words' meanings and decreased categorization accuracy—but accuracy increases in the categorization task. An explanation of these opposite effects within the symbolic framework could become uncomfortably ad hoc, especially if it also answered why accuracy falls to "low-frequency" pseudohomophones.

Concerning pseudohomophone data, suppose the changes in performance from the lexical decision task to the categorization task are due to the added context when category names precede targets. Context could affect performance in two ways. It could bias processing toward base word identities and increase the likelihood of false positives for all foils, generally reducing accuracy. But accuracy is only reduced to "low-frequency" pseudohomophones. Or, context might make targets and foils more distinct. In this case, accuracy would rise to all foils, or, at least, the direction of change between tasks would be the same for all foils. But, again, accuracy to some pseudohomophones goes down while accuracy to spelling controls goes up!

Perhaps matching (verification) processes are differentially affected by the change from lexical decision to semantic categorization. For example, consider broadly the processing metaphor underlying the interactive-activation model of letter and word perception (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). A word node such as SLEET matching the majority of stimulus (SLERT) letters is favored early-on over word nodes that do not. On some of our lexical decision trials, spelling controls (SLERT) may have activated word nodes (SLEET) early-on to the point of causing a false positive response, even though word identification did not eventually settle on SLEET and did not retrieve SLEET's meaning. Perhaps a capacity for early activation of word nodes correlates with an evaluation of overall stimulus familiarity (compare Balota & Chumbley, 1984; Besner, Davelaar, Alcott, & Parry, 1984).

Suppose we added phonology to this dynamic, increasing the false positive support for SLEET when the stimulus is the pseudohomophone SLEAT. Enough false positive support could allow relatively stable false-positive activation of SLEET as the outcome of word identification, and would include retrieval of SLEET's meaning. In such a dynamic system, implicit match and mismatch criteria, analogous to the thresholds in a random walk model, determine whether false positive word identification occurs (see Stone & Van Orden, 1992). False positive word identification is prevented through top-down verification of stimulus letters. SLEET can protect itself from false positive word identification through precise top-down activation of letter nodes. Correct top-down activation of SLEET's letter nodes would mismatch the letter nodes activated by the stimulus SLEAT and destabilize the false positive dynamic. We assume SLEET's capacity for precise top-down activation is correlated with word frequency, and is a product of covariant learning (Van Orden et al., 1990, 1992).

Finally, we add the biasing context of category names in the semantic categorization task. A context appropriate to SLEET increases activation of this node and pushes the system further toward false positive word identification. This push will be felt most by dynamics otherwise borderline in preventing false positive word identification. If SLEET only barely produces enough mismatch to escape false positive word identification in the lexical decision task, then the push of context may be enough to stabilize the dynamic in favor of false positive word identification. Borderline capacity is more likely for low-frequency base words because, as we noted, capacity for precise top-down activation ("mismatch") is correlated with word frequency.

Context also occasionally pushes coding of a few spelling control items into false positive word identification (again, as function of base word frequency). However, early activation of a base word node is no longer sufficient to produce a false positive error, as it was in the lexical decision task. The categorization task emphasizes meaning evaluation, and accuracy tracks foils' capacities to falsely retrieve base word meanings. We expect such an account could be implemented in a fully distributed dynamic system in which patterns of activation replace word nodes (compare Stone & Van Orden, 1989, 1992; Van Orden et al., 1992).

Proofreading experiments

Both experiments to this point have presented key stimulus items in relative isolation. In contrast, a proofreading task presents subjects with connected text and may better approximate typical reading. In the proofreading task, subjects mark misspellings detected while reading a stimulus passage. Daneman and Stainton (1991) claimed proofreading has greater ecological validity than lexical decision and semantic categorization tasks. They confirmed this claim in a correlation between proofreading performance and a measure of reading comprehension. Of course, measures of reading skill correlate with word identification as well (Perfetti, 1985). And, of course, we seek common effects across reading tasks, so we do not require any single task to be ideal. However, our experience with different reading tasks leads us to agree with Daneman and Stainton. We believe proofreading better reflects typical reading with less focus on individual word identification and more focus on integrating words' meanings into ongoing comprehension.

The proofreading task has been used previously in tests for phonology effects. These analyses typically derive from the similarity axiom and the monotonicity hypothesis. For example, Banks, Oka, and Shugarman (1981) substituted homophone (SLEAT) and spelling control (SLERT) foils into text to create misspellings of base words (SLEET), and then compared detection rates for these misspellings. They reasoned: If (a) false-positive word identification is a monotonic function of similarity, and (b) false-positive word identification underlies miss-errors in the proofreading task, and (c) phonologic similarity affects word identification, then (d) homophone misspellings should be more difficult to detect than spelling control misspellings. In Banks et al.'s Experiment 7a, subjects detected 54% of the homophone foils versus 66.4% of the spelling controls. This phonology effect was significant by subjects, but not in an F' analysis considering both subjects and items.

Van Orden (1991) also conducted proofreading experiments within the traditional logic. In his Experiment 1, subjects detected 64.6% (SE=3.5) of pseudohomophone foils and 68.8% (SE=2.8) of spelling controls. This effect was not significant (both subject and item $Fs<1$). However, in Experiment 2, subjects detected 57.5% (SE=3.3) of pseudohomophone foils versus 69.2% (SE=2.4) of spelling control foils, a difference significant in both subject and item analyses.

In a similar experiment, Daneman and Stainton (1991, Experiment 1) found equal detection rates (58%) for homophone and control foils. But, their key data came from a new proofreading task. Subjects first read a passage containing correct spellings of base words, then proofread the same passage with homophone and control foils substituted for base words. In the second reading, subjects detected 63% of homophone foils versus 75% of controls. This difference was reliable in both subject and item analyses.

In contrast to the previous mixed results from traditional contrasts, Van Orden (1991) found large and reliable base word frequency effects in detection rates to pseudohomophone foils (26.9% in Experiment 1 and 30.4% in Experiment 2). The base word frequency effects to spelling controls were smaller (9.3% in Experiment 1 and 10% in Experiment 2), and significant only in subject analyses. Van Orden (1991) claimed the base word frequency effect to spelling controls was due to an interaction between context and a few items (just as we argued for spelling controls in the categorization task), but base word frequency effects to pseudohomophone misspellings can be attributed to stimulus phonology.

The next experiments test the previous claims in a manipulation of context. Context is manipulated by presenting foils in contexts *appropriate* to base words (e.g., ...THE WET, FREEZING SLEAT...) versus presenting foils in contexts *inappropriate* to base words (e.g., ACROSS THE DOUBLE SLEAT LINES...). If base word frequency effects to pseudohomophone foils originate in "bottom-up" constraints of stimulus phonology, then we may find a base word frequency effect, even in the inappropriate context condition. In contrast, if the base word frequency effect to spelling controls is due to context, then this effect may disappear in the inappropriate context condition.

We also included a bigram control foil condition. Van Orden (1991) argued sources of error other than false-positive word identification affect proofreading performance to spelling control foils. For example, subjects may fail to mark a spelling control or a bigram control because they are perceived as novel words with unfamiliar meanings, not misspellings of familiar words. These sources of error may be insensitive to similarity between foils and base words and, consequently, these sources may obscure phonology effects in proofreading. The detection rates to bigram control foils estimate sources of error in proofreading performance other than false-positive word identification.

The first proofreading experiment is concerned primarily with context effects on detection of foils. The two additional proofreading experiments also test for sources of miss-errors in proofreading other than false-positive word identification. The second proofreading experiment asks subjects directly whether they perceive the foils as novel words, and the third proofreading experiment asks this question less directly. Other than differences in subjects' instructions concerning how to respond, the three proofreading experiments all incorporate the same method.

General method

Subjects. The subjects were new groups of 120, 100, and 100 undergraduates from the same population as the previous experiments who participated in the first, second and third proofreading experiments, respectively.

Stimuli. The two short stories used in the experiment— *Lucille* and *The Hour of Least Resistance* (see Appendices C and D, respectively)—were written by Guy Van Orden and John Graham Holden. These stories were written around 20 base words from which experimental foils were constructed. The constraints on composition were that each of the 20 base words was used only once in its respective story (each story contained 10 base words), and they were spaced roughly, evenly across the story. The stories were formatted using *Word Star* word processing software and printed in upper-case.

We constructed stimulus foils from base words as in the lexical decision and semantic categorization experiments. For each of the base words we chose a pseudohomophone, a spelling control, and a bigram control foil. These foils appear in Appendix B.

Base word frequency was manipulated as before using the Kučera and Francis (1967) word frequency count. Half (10) of the base words had high-frequency counts ranging from 55 to 216 per million (M=125, SD=48), and half had low-frequency counts ranging from 1 to 10 per million (M=5, SD=3). Each pseudohomophone foil constructed from a low-frequency base word was yoked to a pseudohomophone foil constructed from a high-frequency base word. All pairs were formed by spelling changes preserving identical local phonology, and all pairs were formed by the same spelling change in their respective base words. We chose stimuli to include a variety of focal orthographic-phonologic correspondences. The homophony of the pseudohomophones was insured using the method described in the previous lexical decision task and Van Orden et al. (1988).

Degree of spelling similarity between pseudohomophones and base words, and spelling controls and base words, was controlled as before using OS. Mean OS computed between low-frequency base words and corresponding spelling controls was 0.79, SD=.04, and between low-frequency base words and corresponding pseudohomophone foils it was 0.78, SD=.04; between high-frequency base words and corresponding spelling controls it was 0.78, SD=.04, and between high-frequency base words and corresponding pseudohomophone foils it was 0.78, SD=.04. OS computed between bigram control foils and base words shows that these foils are very low in spelling similarity to base words ("high-frequency" bigram controls M=.08, SD=.04, "low-frequency" bigram controls M=.17, SD=.09).

Control for bigram frequency was accomplished using the bigram frequency count of Massaro et al. (1980). The mean log frequency of bigrams that compose pseudohomophones was 2.303, SD=.972, spelling controls M=2.457, SD=.810, and bigram controls M=2.443, SD=.827. Mean N for high-frequency pseudohomophones was 3.1, SD=2.8; for low-frequency pseudohomophones M=2.2, SD=1.9; for high-frequency spelling controls M=2.3, SD=1.0; for low- frequency spelling controls M=2.1, SD=1.0; for high-frequency bigram controls M=2.1, SD=3.4; and for low-frequency bigram controls M=.5, SD=1.0.

Procedure. Subjects were assigned to conditions randomly and run in groups of 20-40. One half of the subjects read stories into which foils (SLEAT, SLERT, FLESS) had been substituted into contexts appropriate to corresponding base words (SLEET). The remaining subjects read stories in which foils appeared in contexts inappropriate to corresponding base words. We used ten random arrangements. Each arrangement was used equally often with the three kinds of foils so yoked pseudohomophone, spelling control, and bigram control foils appeared in the same inappropriate contexts, across subjects. Within the *appropriate context* condition and the *inappropriate context* condition, one third (20) of the subjects in each condition read stories containing pseudohomophone foils (SLEAT), another third read stories containing spelling controls (SLERT), and the remaining third read stories containing bigram controls (FLESS). (The second and third proofreading experiments did not include the inappropriate context condition for bigram control foils because it is virtually identical to the appropriate context condition with respect to bigram controls.)

In a typical session, each subject received a booklet containing an instruction cover-sheet and two short stories. The ordering of the two stories in the booklets was counter-balanced across all other conditions. At the beginning of the session, the experimenter read aloud the instructions while the subjects read them silently. For the first proofreading experiment subjects were instructed to (a) Read straight through the stories. Don't back-track to look for misspellings, but (b), If you happen to come across any misspellings put an X through them. And (c) Try to ignore the fact that this is a psychology experiment and read the stories as you might read them for your own pleasure.

For the second experiment subjects were instructed as in (a) and (b) previously, then (c) And, if you read a word and don't know its meaning, put a ? on it. (d) As much as is possible... read [the stories] as you might read them for your own pleasure. Allowing "?", a "novel word" response, tests whether subjects sometimes read "misspelled" stimuli in a proofreading task as novel words for which they do not know the meaning. Proofreading tasks that do not allow a novel-word response may overestimate miss-errors. Worse yet, differences in rates of novel-word responding to homophone and spelling control foils may be confounded with differences in miss error rates (see also Van Orden, 1991). For the third experiment subjects were instructed as in (a) previously, then (b) If you happen to notice a misspelled word, put an X through it, and then write the correct spelling nearby in the margin. (c) If the misspelling has changed the word so that it is unrecognizable, then put an X through the misspelling and a ? in the margin. (d) As much as possible, try to ignore the fact that this is a psychology experiment and read the stories as you might read them for your own pleasure. (No doubt, this last instruction became increasingly difficult to obey as each experiment required more elaborate forms of responding.) This third version of responding tests whether subjects are reluctant to mark foils as unfamiliar words. Marking foils as unfamiliar may entail a stigma associated with "not knowing." We hoped marking a foil as unrecognizable due to our having changed it would avoid such a stigma.

After the instructions, the subjects opened their booklets and proceeded to read through both stories.

Results and discussion: Appropriate context

The following planned analyses examine first the data from the appropriate context conditions of the three proofreading experiments, and then like data from the inappropriate context conditions. Item data from all these conditions are presented in Appendix B. We begin with the results from the appropriate context condition of the first proofreading experiment and then move to results from the appropriate context conditions of the second and third proofreading experiments. The percentage of "misspelled" foils detected by subjects in the appropriate context condition of the first proofreading experiment is presented in Table 3. Overall, in this experiment, three independent variables were manipulated: foil type (pseudohomophones versus spelling controls versus bigram controls), base word frequency (high versus low), and context (appropriate versus inappropriate—compare Tables 3 and 6). The three way interaction between these variables approached significance by subjects ($F(2,114)=2.87$, $p<.07$), but not by items ($F(2,18)=1.64$, $p>.20$).

Traditional analysis. Mean accuracy across the three foil conditions contradicts the similarity axiom and monotonicity hypothesis. As expected, overall accuracy was greater to spelling controls (80%, SE=2.6) than to pseudohomophones (67%, SE=3.9). However, although we expected subjects to be most accurate to bigram control foils, these foils produced the lowest rate of correct detections (46%, SE=4.9). The similarity axiom and monotonicity hypotheses must predict the rank order of accuracy to be (1) bigram controls, (2) spelling controls, (3) pseudohomophones, as found in both previous experiments, not the rank order of (1) spelling controls, (2) pseudohomophones, (3) bigram controls found here.

Table 3

Percentage of pseudohomophone, spelling control, and bigram control foils marked as misspellings in the appropriate context condition of the first proofreading experiment.

	APPROPRIATE CONTEXT		
	Pseudohomophones	Spelling Controls	Bigram Controls
High Frequency Base Words	80.0	87.0	46.5
Low Frequency Base Words	54.0	73.0	45.5
Mean	67.0	80.0	46.0
Difference (Base Word Frequency Effect)	26.0	14.0	1.0

The difference in accuracy (21%) between bigram controls and pseudohomophones was significant ($t(38)=2.64$ for subjects; $t(19)=3.63$ for items), as was the difference in accuracy (13%) between pseudohomophones and spelling controls ($t(38)=2.56$ for subjects; $t(19)=2.85$ for items). By the traditional logic then, items least similar in surface characteristics to base words are most likely to falsely retrieve base word lexical knowledge!

The previous conclusion makes no sense. But, of course, the bizarre results behind this bizarre conclusion may be due to some artifact, or some instability in the proofreading method. For example, perhaps subjects generate two hypotheses (at least) when presented with a nonword foil in the proofreading task. The foil could be a misspelling, consistent with the explicit task demands of proofreading, but it could also be an unfamiliar correctly spelled word. The second hypothesis is reasonable because all readers are likely to have come across unfamiliar words in connected text, so it is consistent with the implicit task demands of typical reading.

The method of the second proofreading experiment tests this possibility directly. We instructed subjects in two modes of marking nonword foils. They could mark a nonword foil as a misspelling, or they could mark it as a novel word. If the extra-poor performance to bigram controls in the first proofreading experiment comes from subjects perceiving them as novel words, then the novel-word response may pick up this effect. Perhaps then, an overall measure of accuracy combining the percentage of foils marked as misspellings and the percentage of foils marked as novel words, *in one score,* may regain compliance with the similarity axiom and the monotonicity hypothesis.

The results from the appropriate context condition of the second proofreading experiment are presented in Table 4. Again, even summing the percentage of foils marked as misspellings with the percentage of foils marked as novel words, mean accuracy across the three foil conditions contradicts the similarity axiom and monotonicity hypothesis.

Table 4

Percentage of pseudohomophone, spelling control, and bigram control foils marked as misspellings or novel words in the appropriate context condition of the second proofreading experiment. The number in parenthesis is the percentage of foils in each cell marked as novel words.

	APPROPRIATE CONTEXT		
	Pseudohomophones	Spelling Controls	Bigram Controls
High Frequency Base Words	74.0 (0)	79.0 (11.0)	64.5 (50.5)
Low Frequency Base Words	52.5 (1.5)	72.5 (15.0)	71.5 (65.0)
Mean	63.2	75.8	68.0
Difference (Base Word Frequency Effect)	21.5	6.5	-7.0

However, this time the rank order of accuracy was spelling controls (75.8%, SE=3.1), followed by bigram control foils (68.0%, SE=4.1), and then pseudohomophones (63.2%, SE=4.5).

The difference in accuracy (7.8%) between spelling controls and bigram controls was not significant ($t(38)=1.14$, $p=.26$ for subjects; $t(19)=1.69$, $p>.10$ for items), and the difference in accuracy (4.8%) between bigram controls and pseudohomophones also failed to reach significance (both ts<1). However, accuracy to spelling controls was significantly higher than to pseudohomophones in the item analysis ($t(19)=3.13$) and approached significance in the subject analysis ($t(38)=1.84$, $p<.08$).

We may make sense of the difference between accuracy to pseudohomophones and accuracy to spelling controls, but how should we interpret the intermediate effect of bigram controls? The overall pattern of accuracy across spelling controls, bigram controls, and pseudohomophones again contradicts the similarity axiom and the monotonicity hypothesis. However, perhaps subjects were sometimes reluctant to mark foils as novel words. This response may entail a stigma associated with "not knowing." The third proofreading experiment tested for this possibility.

In the third experiment, subjects were told spellings of some words had been changed only a little, but other words had been changed so much they are unrecognizable. We hoped marking a foil as unrecognizable due to our having changed it would avoid any stigma associated with "not knowing." We expected this change would bring our results in line with the similarity axiom and monotonicity hypothesis. Subjects were instructed first to mark misspellings in the text, then write the correct spelling in the margin or put a "?" in the margin for unrecognizable misspellings. The results from the appropriate context condition of the third proofreading experiment are presented in Table 5.

Table 5

Percentage of pseudohomophone, spelling control, and bigram control foils marked as misspellings in the appropriate context condition of the third proofreading experiment. The percentage of foils marked as misspellings of unknown words ("?" in margin) are presented in parentheses. The percentage of foils marked as misspellings and identified as misspellings of base words (base word in margin) is presented in brackets.

APPROPRIATE CONTEXT

	Pseudohomophones			Spelling Controls			Bigram Controls		
High Frequency Base Words	78.0	(0.0)	[78.0]	84.0	(6.5)	[76.0]	67.0	(55.5)	[1.5]
Low Frequency Base Words	50.0	(0.5)	[49.5]	79.5	(10.0)	[68.0]	62.0	(55.5)	[1.0]
Mean	64.0			81.8			64.5		
Difference (Base Word Frequency Effect)	28.0			4.5			5.0		

For a third time, mean accuracy across the three foil conditions contradicts the similarity axiom and monotonicity hypothesis! As in the second proofreading experiment, the rank order of accuracy was spelling controls (81.8%, SE=3.2), followed by bigram control foils (64.5%, SE=4.0), and then pseudohomophones (64.0%, SE=3.9).

The difference in accuracy (17.8%) between spelling controls and pseudohomophones was significant ($t(38)=2.98$ for subjects; $t(19)=3.04$ for items) as was the difference (17.3%) between spelling controls and bigram controls ($t(38)=2.55$ for subjects; $t(19)=3.04$ for items). Accuracy to pseudohomophone and bigram foils was virtually the same (both ts<1). Within the traditional logic then, we would conclude the phonologic identity of pseudohomophones to base words causes a decrease in proofreading accuracy due to retrieval of base word lexical memories (relative to spelling controls). But, the lack of similarity between bigram controls and base words (and words in general) yields equal retrieval of base word lexical memories, reflected in an equal decrease in proofreading accuracy (relative to spelling controls).

The similarity axiom and monotonicity hypothesis have failed to predict overall accuracy in three proofreading experiments. This failure motivates careful scrutiny of theories and methods that must assume monotonic effects of stimulus-memory similarity. We may no longer assume *a priori* that "meaningful performance" is a monotonic function of the similarity between a stimulus and an explicit, permanent memory structure. Additionally, because this failure is not due to any obvious artifact, it must be accommodated by an adequate theory of proofreading.

Meaning creation and proofreading. Suppose meaning is not stored explicitly in memory, and ongoing comprehension is not an evaluation of fixed word meanings based on a "real" world. Rather, meaning is continuous in a semantic space (Stone & Van Orden, 1989). "Fixing" meaning at a point in space corresponds to creating meaning in ongoing comprehension. A word's conventional meanings reflect converging, average, experiential constraints (compare "image schemas" in Johnson, 1987; Lakoff, 1987; J. Mandler, in press). These converging, average, experiential constraints (attractors) are a natural consequence of covariant learning in a dynamic system (compare Van Orden et al., 1990, 1992). In reading isolated words, the "fixed points" in semantic space are some function of these conventional meanings. In reading text, "fixed points" balance contextual, experiential, and stimulus constraints.

These assumptions allow speculation as to why more proofreading miss-errors are made to bigram control foils than to spelling control foils. A spelling control is very similar to the surface form of a previously experienced word, but bigram controls are not. Perhaps a spelling control's surface resemblance to a base word causes it to be coded close to an "attractor point" in semantic space. But, a bigram control is coded well between attractor points because its resemblance to words is merely a statistical resemblance to all words. Assume (a) somewhat coherent activation close to a viable attractor point destabilizes ongoing creation of meaning more than does less coherent activation well between viable attractor points, and (b) detection of misspelled foils depends on their capacity to destabilize ongoing meaning creation (compare *cognitive interruption* in G. Mandler, 1984). Given these assumptions, spelling controls will be detected more often than bigram controls. In other words, detection of foils during ongoing comprehension is more than mere failure of word recognition. Detection requires that the "failure" of word recognition interrupts ongoing meaning creation.

On the other hand, the phonology of pseudohomophones provides strong stimulus-driven support for a contextually viable attractor point. Consequently, detection only occurs when "mismatch" within word identification disrupts ongoing meaning creation. In this case, "mismatch" refers to mismatch of top-down and bottom-up spelling activation (as we described previously for lexical decision and semantic categorization performance).

Base word frequency effects. In contrast to the previous traditional analyses, analyses of base word frequency effects in the three appropriate context conditions all confirm the similarity axiom and monotonicity hypothesis. In the first proofreading study (see Table 3), ANOVAs showed a significant interaction effect between foil type and base word frequency ($F(2,57)=10.28$ for subjects; $F(2,18)=4.61$ for items). Planned analyses of simple main effects found a large, significant base word frequency effect (26%) to pseudohomophone foils ($t(19)=6.20$ for subjects; $t(9)=3.27$ for items), and a smaller but significant base word frequency effect (14%) to spelling controls ($t(19)=3.76$ for subjects; $t(9)=2.41$ for items).

ANOVAs considering only the data from pseudohomophones and spelling controls confirmed the larger base word frequency effect to pseudohomophones in the two-way interaction between base word frequency and foil type ($F(1,38)=4.58$ for subjects; $F(1,9)=5.16$ for items). This pattern of base word frequency effects regains compliance with the similarity axiom and monotonicity hypothesis. Pseudohomophone stimuli generate the largest base word frequency effect.

The second proofreading experiment (see Table 4) also found a significant interaction effect between foil type and base word frequency in the appropriate context condition ($F(2,57)=13.69$ for subjects; $F(2,18)=8.92$ for items). Planned analyses of simple main effects found a large, significant base word frequency effect (21.5%) to pseudohomophone foils ($t(19)=5.39$ for subjects; $t(9)=3.63$ for items), but the base word frequency effect (6.5%) to spelling controls was not significant ($t(19)=1.58, p>.13$ for subjects; $t(9)=1.40, p>.19$ for items).

The failure to find a base word frequency effect to spelling controls in the second proofreading experiment may be due to a tendency for more low-frequency spelling controls to be marked as novel words, and this partly counters the opposite tendency for more high-frequency spelling controls to be marked as misspellings. If we looked exclusively at items marked as misspellings, the size of the resulting base word frequency effect to spelling controls (10.5%) would be about the same as the effect observed in the appropriate context condition of the first proofreading experiment (14%). To trust this isolated effect we would need to know items marked as novel words in the second proofreading experiment would not have been marked as misspellings in the first proofreading experiment. But, our argument for phonology effects does not rest on an absence of base word frequency effects to spelling controls. We merely require smaller base word frequency effects to spelling controls than to pseudohomophones, as we have found in every condition of every experiment.

In the third proofreading experiment (see Table 5), the significant interaction effect between foil type and base word frequency ($F(2,57)=15.43$ for subjects; $F(2,36)=3.86$ for items) confirms yet again the three types of foils have different capacities for retrieving base word lexical knowledge. Planned analyses of simple main effects found a large, significant base word frequency effect (28%) to pseudohomophone foils ($t(19)=10.10$ for subjects; $t(9)=3.01$ for items), but the base word frequency effect (4.5%) to spelling controls was not significant ($t(19)=1.23, p>.23$ for subjects; $t<1$ for items).

Of course, the previous analyses of base word frequency effects on proofreading performance all come from appropriate context conditions—so context is confounded with stimulus similarity. The inappropriate context condition better isolates effects of stimulus similarity.

Results and discussion: Inappropriate context

Traditional analysis. The results from the inappropriate context condition of the first proofreading experiment are presented in Table 6. As in all previous traditional analyses, mean accuracy across the three foil conditions contradicts the similarity axiom and monotonicity hypothesis. Accuracy was greatest to spelling controls (64.2%, SE=4.0), followed by pseudohomophones (55.5%, SE=4.2), and lowest to bigram control foils (43%, SE=3.9).

The difference in accuracy (12.5%) between bigram controls and pseudohomophones approached significance ($t(38)=1.81$, $p<.08$ for subjects; $t(19)=2.05, p<.06$ for items), but the difference (8.8%) in accuracy between pseudohomophones and spelling controls was not significant ($t(38)=1.26, p>.21$ for subjects; $t(19)=1.63, p>.11$ for items). Following the traditional logic then, we would again conclude the items least similar in surface characteristics to base words are most likely to be falsely recognized as base words, and we would concede phonology fails to affect proofreading performance.

Table 6

Percentage of pseudohomophone, spelling control, and bigram control foils marked as misspellings in the inappropriate context condition of the first proofreading experiment.

INAPPROPRIATE CONTEXT			
	Pseudohomophones	Spelling Controls	Bigram Controls
High Frequency Base Words	69.5	63.5	43.5
Low Frequency Base Words	41.5	65.0	42.5
Mean	55.5	64.2	43.0
Difference (Base Word Frequency Effect)	28.0	-1.5	1.0

The results from the inappropriate context condition of the second proofreading experiment are presented in Table 7. In this inappropriate context condition, we manipulated foil type (pseudohomophones versus spelling controls) and base word frequency (high versus low). (As noted previously, no data were collected for bigram control foils in this condition.)

Subjects were more likely to mark spelling controls as misspellings or novel words (75.8%) than pseudohomophones (63.2%). This main effect of foil type approached significance in the subject ANOVA ($F(1,38)=3.20$, $p<.09$) and was significant in the item ANOVA ($F(1,9)=11.48$).

The results from the inappropriate context condition of the third proofreading experiment are presented in Table 8. The dependent measure was the total percentage of foils marked by subjects as misspellings, but we again present the percentage of foils marked as unknown words (a "?" was written in the margin next to the foil), and the percentage of foils for which the base word was written in the margin next to the foil.

The difference in accuracy (11.4%) between pseudohomophones and spelling controls was not significant by subjects ($F(1,38)=2.55$, $p>.11$), but it was significant by items ($F(1,9)=8.80$). Once again, we have failed to find a reliable, traditional phonology effect. By the traditional logic, some subjects make use of assembled phonology, but others do not. Considering the reliable traditional phonology effects from the appropriate context conditions, we would conclude, overall, that subjects generally use assembled phonology when the context is biased toward base words, but only a few subjects use assembled phonology in all contexts. Of course, this pattern adds "null support" to the various "nonphonologic" hypotheses offered to explain the comings and goings of phonology effects across tasks and conditions.

Table 7

Percentage of pseudohomophone and spelling control foils marked as misspellings or novel words in the inappropriate context condition of the second proofreading experiment. The percentage of foils in each cell marked as novel words appears in parentheses.

	INAPPROPRIATE CONTEXT			
	Pseudohomophones		Spelling Controls	
High Frequency Base Words	72.5	(32.5)	78.0	(61.5)
Low Frequency Base Words	54.0	(29.0)	73.5	(61.0)
Mean	63.2		75.8	
Difference (Base Word Frequency Effect)	18.5		4.5	

Table 8

Overall percentage of pseudohomophones and spelling control foils marked as misspellings in the inappropriate context condition of the third proofreading experiment. The percentage of foils marked as misspellings of unknown words ("?" in margin) are presented in parentheses. The percentage of foils marked as misspellings and identified as misspellings of base words (base word in margin) is presented in brackets.

	INAPPROPRIATE CONTEXT					
	Pseudohomophones			Spelling Controls		
High Frequency Base Words	77.0	(29.5)	[32.0]	77.5	(45.0)	[5.0]
Low Frequency Base Words	54.5	(25.0)	[19.5]	77.0	(48.0)	[3.0]
Mean	65.8			77.2		
Difference (Base Word Frequency Effect)	22.5			0.5		

However, different data in Table 8 reveal a large phonology effect: 25.8% of pseudohomophones were marked as misspellings and then identified as a misspelling of their base word. This, even though they appeared in a context inappropriate to that base word. Only 4% of spelling controls were identified as misspellings of base words.

This difference was significant despite more opportunities for spelling controls to be identified in the margin as base words because they were marked as misspellings 11.4% more often than pseudohomophones ($F(1,38)=20.3$ for subjects; $F(1,9)=73.96$ for items). (We cannot trust the interaction or the respective base word frequency effects on the production of base words in the margin because they may merely coincide with the rate of opportunities for this response—fewer foils marked as misspellings means fewer opportunities to produce the base word in the margin.)

Consistent with the previous finding, 46.5% of spelling control foils attracted the "?" response versus 27.2% of pseudohomophones. This difference was significant ($F(1,38)=5.28$ for subjects; $F(1,9)=36.60$ for items). (No other variables affected the rate of "?" responding.) Pseudohomophones and spelling controls are equally similar to their base words except pseudohomophones are phonologically identical to base words. The phonologic identity of pseudohomophones and base words reduces subjects' tendency to mark foils as misspellings of unknown words, even in a context inappropriate to the base word.

Base word frequency effects. Only pseudohomophones produced base word frequency effects in inappropriate context conditions. In the first proofreading experiment (see Table 6), ANOVAs found a significant interaction effect between foil type and base word frequency ($F(2,57)=15.92$ for subjects; $F(2,18)=6.00$ for items). Planneu analyses of simple main effects found a large, significant base word frequency effect (28%) to pseudohomophone foils ($t(19)=5.37$ for subjects; $t(9)=3.41$ for items), but no base word frequency effect was found to spelling controls (both ts<1).

In the second proofreading experiment (see Table 7), ANOVAs found a second significant interaction effect between foil type and base word frequency ($F(1,38)=5.75$ for subjects; $F(1,9)=6.10$ for items). Analyses of simple main effects found a large significant base word frequency effect (18.5%) to pseudohomophones ($t(19)=3.79$ for subjects; $t(9)=3.70$ for items), but no significant effect to spelling controls ($t(19)=1.41, p>.17$ for subjects; $t(9)=1.03, p>.33$ for items).

In the third proofreading experiment (see Table 8), ANOVAs found a third significant interaction between foil type and base word frequency ($F(1,38)=14.99$ for subjects; $F(1,9)=6.74$ for items). This interaction indicates for the tenth data set that pseudohomophones and spelling controls have different capacities for retrieving base word lexical knowledge. Planned analyses of simple main effects again found a large, significant base word frequency effect (22.5%) to pseudohomophone foils ($t(19)=5.25$ for subjects; $t(9)=2.56$ for items), but the base word frequency effect (0.5%) to spelling controls was not significant (both ts<1).

Spelling controls do not produce base word frequency effects except in contexts appropriate to corresponding base words. Proofreading performance to spelling controls in the appropriate versus inappropriate context conditions parallels in quality our previous contrast between spelling control data in the lexical decision task (no base word context) versus the semantic categorization task (context appropriate to base words). There too we found an small base word frequency effect to spelling controls in an appropriate context condition (the semantic categorization task), but not otherwise. Apparently, spelling control items cannot retrieve base word lexical knowledge except in prejudicial contexts. Of course, these are converging null results and the usual caveats apply.

Pseudohomophones produce reliable base word frequency effects in all proofreading conditions. We may attribute pseudohomophones' capacity for retrieving lexical knowledge to their phonologic identity with base words. Once again then, we have demonstrated a reliable assembled (nonword) phonology effect, this time in a task presenting subjects with extensive passages of connected text.

Appropriate versus inappropriate context: Novel-word response

These final analyses examine "novel-word response" data from the appropriate and inappropriate context conditions of the second proofreading experiment (see Tables 4 and 7, respectively). Spelling controls attract significantly more novel word responses in the inappropriate context condition (61.2%) than in the appropriate context condition (13%, $t(38)=7.21$ for subjects; $t(19)=14.16$ for items). In fact, the overall percentage of novel word responses to spelling controls in the inappropriate context condition (61.2%, see Table 7) is about the same (both ts<1) as the percentage of novel word responses to bigram controls (57.8%, see Table 4). It appears spelling controls are usually processed like bigram controls when they appear in contexts inappropriate to base words. They are perceived as novel words on the majority of trials, and they do not produce a base word frequency effect.

Thirteen percent of spelling controls are marked as novel words in contexts appropriate to base words, but less than 1% of pseudohomophones are marked as novel words (see Table 4). (A parallel effect can be seen in Table 5 from the third proofreading experiment.) This may explain why traditional analyses so often turn out null phonology effects in the proofreading task. "All other things are not equal" between pseudohomophones and spelling controls (compare Van Orden, 1991). Sources of proofreading miss-errors to spelling controls exist that do not always affect performance to pseudohomophones. Consequently, spelling controls' function as baseline controls is compromised.

The pseudohomophones were marked as novel words on an unexpectedly large number of trials (30.8%) in the inappropriate context condition (see Table 7, and like data in Table 8). But, while subjects mark some pseudohomophones as novel, they still produce a base word frequency effect (18.5%) of about the same magnitude as the base word frequency effect in the appropriate context condition (both Fs<1). The undiminished base word frequency effect in the inappropriate context condition is superficially inconsistent with the exaggerated percentage of novel word responses. Perhaps pseudohomophones are marked as novel words because subjects notice the disagreement between context and base word identities, and they sometimes resolve this disagreement by marking the pseudohomophone as a novel word.

One last result is worth mentioning. We found an unexpected base word frequency effect (50.5 - 65 = -14.5%, see Table 4) on the percentage of novel word responses to bigram controls in the appropriate context condition ($t(19)=3.68$ for subjects; $t(9)=2.12$, $p<.07$ for items). If this result were reliable, it would indicate a difference in the sentence contexts in which "low frequency" and "high frequency" foils are presented, or a difference between the two groups of bigram controls (they differ in N, for example). However, this difference did not replicate for bigram control foils in any condition of any other proofreading experiment.

General discussion

Three laboratory reading tasks all produced base word frequency effects to pseudohomophones (SLEAT). Base word frequency effects indicate pseudohomophones retrieve lexical knowledge of base words (SLEET), and this knowledge affects whether pseudohomophones are falsely identified as base words. Performance to spelling control foils (SLERT), is also, occasionally, affected by base word frequency—but only when these controls appear in contexts appropriate to the base words. Thus, our method isolates the phonologic identity of pseudohomophones to base words as a general and reliable source of base word frequency effects.

These results are conclusive. Pseudohomophones' "assembled" phonology reliably constrains word identification toward phonologically identical lexical items. We claim a common mechanism underlies phonologic coding of both words and nonwords (Van Orden et al., 1988; Van Orden et al., 1990). The present results strengthen this claim because novel letter strings, equal to familiar words in phonology, activate lexical knowledge as reliably as the words themselves. This possibility is handled nicely if we assume novel and familiar letter strings are coded in a common mechanism.

Phonology and reading. The present findings are consistent with our strong position concerning the role of phonology in reading. Reading ability depends on printed word identification (Perfetti, 1985), and printed word identification depends on coherent phonology. For example, the primary correlate of reading disability is a concomitant failure to develop adequate phonologic coding skills (Pennington, 1991; Pennington, Van Orden, Smith, Green, & Haith, 1990). Every instance of printed word identification is constrained fundamentally by the dynamic between phonologic and orthographic codes. Elsewhere, we describe a phonologic coherence hypothesis that motivates our claim computationally (Van Orden, 1991; Van Orden et al., 1990).

In skilled reading, a printed word activates phonology in the earliest moments of word perception (Lesch, 1990; Lesch & Pollatsek, 1992; Perfetti & Bell, 1991). This early phonologic code is relatively coherent and it survives brief presentations followed by pattern masking (Perfetti, Bell, & Delaney, 1988; Van Orden, 1987). As a consequence, the dynamic between phonologic and orthographic codes is a basis for early holism in word perception (as revealed by methods associated with word- superiority effects, Chastain, 1981, 1984; Hawkins, Reicher, Rogers, & Peterson, 1976).

Phonologic activation is automatic (Perfetti & Bell, 1991), and phonologic codes reliably constrain other lexical codes—even when computed from unfamiliar letter strings such as pseudohomophones. Phonologic constraints cannot be suppressed (Jared & Seidenberg, 1991; Van Orden, 1984), although strategic compensation other than suppression is possible when stimulus phonology works against optimal performance in laboratory reading tasks (Stone & Van Orden, in 1992). One such strategy may distinguish between pseudohomophones and word homophones (Coltheart et al., 1991; Jared & Seidenberg, 1991), although their phonology is equal (Lukatela & Turvey, in press; Van Orden et al., 1988) and arises from a common mechanism of phonologic coding (Van Orden et al., 1990).

This brief review summarizes our present beliefs concerning phonology and reading. Given the growing body of positive phonology effects in recent years, one might wonder how "nonphonologic" hypotheses survive.

Previous failures to observe phonology effects. Traditional methods failed to observe reliable phonology effects because they were narrowly focused within a symbolic, human-information-processing metaphor. The logic of those methods derived from the similarity axiom coupled with the monotonicity hypothesis: Letter-strings should affect reading performance monotonically as a function of their similarity to lexical representations. Consequently, pseudohomophones' phonologic identity to base words should always produce effects over-and-above the effects of spelling controls.

Despite the elegance of this logic, it cannot always be trusted. For example, the traditional logic required that spelling control foils provide a stable baseline from which to evaluate phonology effects. However, in our experiments, overall performance to spelling control foils varied dramatically with task demands and context conditions. Worse yet, proofreading performance cannot be trusted to respect the similarity and monotonicity hypotheses. In contrast to the unstable effects of spelling and bigram controls, pseudohomophones produced similar false positive error rates and reliable base word frequency effects, in every condition, across all three reading tasks. If we had remained focused narrowly within the traditional logic, we would have missed pseudohomophone phonology's reliable and pervasive effect on reading performance.

Sometimes assumptions of the symbolic framework have more subtly influenced the outcome of experiments. One subtle influence caused a prominent failure to find base word frequency effects. McCann, Besner, and Davelaar (1987) and Coltheart et al. (1988, 1991) all failed to observe base word frequency effects to pseudohomophone foils (and word homophone foils). Coltheart et al. (1988, 1991) observed null base word frequency effects on sentence verification performance. McCann et al. (1987) observed null base word frequency effects on lexical decision performance to pseudohomophones (in planned correlational analyses). McCann et al.'s failure is of greater concern because Coltheart and her colleagues only sought the effect in post hoc analyses, and, as they point out, the range of base word frequency was truncated, reducing the chance of detecting this effect.

Turning to McCann et al. (1987) then, we propose they failed to find a base word frequency effect because they assumed explicit rules govern phonologic coding. Of course, the explicit rules axiom is central to the symbolic framework. Belief in explicit rules caused McCann et al. to misconstruct a portion of their pseudohomophones. We do not suggest they were careless in their method; they were meticulous. However, their operational definition of *pseudohomophone* was based on a questionable theory of printed word identification.

How should we operationalize pseudohomophony? Traditional symbolic analysis has relied on a theory-dependent operational definition of *pseudohomophone*. For example, dual process theory (M. Coltheart, 1978), the most widely accepted symbolic theory of printed word identification, assumes the phonology of novel letter strings is assembled by explicit rules. If orthographic-phonologic correspondence is rule governed, then experimenters may deduce the phonology of letter strings. Once you know the rules, you can generate the phonology of any letter string (but who could claim to know the rules?). This operational definition requires the explicit rule hypothesis be true, but the explicit rule hypothesis has been challenged (e.g., see Van Orden et al., 1990).

The challenge to explicit rules lead us to seek a theory-neutral operational definition of *pseudohomophone*. If all potential subjects pronounce a "pseudohomophone" to sound like the same word, then it is operationally defined as a *pseudohomophone*. In practice, of course, one cannot ask every potential subject to pronounce a candidate pseudohomophone; a subset of subjects must suffice. In our method, ten "judges" read aloud candidate pseudohomophones embedded in long lists of nonhomophonic pseudowords (to minimize the likelihood of a strategy by which judges try to puzzle out which words the candidates "sound like"). To be a *pseudohomophone* at least nine of ten judges must produce the homophonic pronunciation, and any disagreement must come from a misperception of the letters (e.g., SLEAT pronounced as SLEEP—see Van Orden et al., 1988, and the method of our lexical decision task). This definition is theory neutral and it better guarantees pseudohomophones will activate homophonic phonology.

To our knowledge, all methods in failures to find pseudohomophone effects also relied on experimenters' rule-intuitions concerning pseudohomophony. For example, McCann et al. (1988) relied on experimenter intuition and they failed to find a base word frequency effect using 80 pseudohomophones in a lexical decision task. These same pseudohomophones appeared previously in the naming experiments of McCann and Besner (1987), and naming error data were included in an appendix. Fully half (80) of the items in McCann and Besner's (1987) naming task were the "pseudohomophones" in question, and subjects were "informed that some of the letter strings, when pronounced, would sound like English words." Still, subjects produced plenty of nonhomophonic pronunciations. Approximately 14% of their "pseudohomophones" were given nonhomophonic pronunciations by over 10% of subjects, and an additional 19% (approximately) were given nonhomophonic pronunciations by exactly 10% of subjects. Because McCann and Besner's method favors homophonic pronunciations, an even greater portion of their items would likely have failed our criteria.

Notice, our operational definition of *pseudohomophone* would almost always produce pseudohomophones consistent with experimenters' rule-intuitions. But, pseudohomophones deduced from rule-intuitions are much less likely to satisfy our criteria. Subjects in our norming studies often produce pronunciations inconsistent with our intuitions. The previous asymmetry is correlated with observed phonology effects. We find them, they sometimes fail to find them.

Statistical regularity. It is not always possible to choose theory-neutral operational definitions. For example, we believe a statistical relation holds between orthography and phonology (Seidenberg & McClelland, 1989; Van Orden, 1987; Van Orden et al., 1990, 1992), not the rule-governed relation assumed in symbolic analysis. We also expect statistical regularity effects in verification. Covariant learning of statistical relations shapes top-down expectations as well as bottom-up codes (Van Orden et al., 1992).

In the present case, the vitality of this *statistical regularity hypothesis* lead us to control local properties of orthographic-phonologic correspondence between pairs of pseudohomophones. If the relation between spelling and phonology is statistical, then local correspondences, within a letter string, could differentially affect reading performance. For example, we matched local properties of orthographic-phonologic and phonologic-orthographic correspondence between "high-frequency" pseudohomophones (GREAN) and their yoked "low-frequency" pseudohomophones (SLEAT). Apart from yoked local correspondence, these items differed in meaning (for all tasks), in category membership (for the categorization task), in sentence context (for the proofreading task), and, of course, in base word frequency. They are only alike in this tenuous local thread of common phonology and a common change (usually) in base word spelling.

Still, item yoking carried significant variance for correct rejection rates in the lexical decision task (r=.74, $t(8)$=3.07, p<.05, but not for "no" RTs, t<1). And this correlation replicated in the semantic categorization task for both correct rejection rates (r=.66,

$t(8)=2.47, p<.05$), and correct "no" RTs ($r=.64$, $t(8)=2.33, p<.05$). Proofreading accuracy in the three appropriate context conditions was correlated in the same direction ($r=.44$, $r=.35$, $r=.27$, respectively, but these correlations were not significant). However, this relation broke down in the inappropriate context conditions ($r=.43$, $r=-.11$, $r=-.40$, respectively). No other stimulus foils showed like correlations under any conditions. Of course, a more careful analysis is needed to establish whether this relation is truly predictive, and whether its origin is in orthographic-to-phonologic or phonologic-to-orthographic correspondence, or both.

Other correlations also confirm the value of control for statistical regularity. We computed correlations between accuracy to various foils and N (the number of words formed by changing a single letter in a foil). Stimulus N did not predict performance to bigram controls or spelling control foils. However, a negative correlation between pseudohomophones' N and lexical decision accuracy was significant ($r=-.47$, $t(18)=2.28$), and the parallel correlation between semantic categorization accuracy approached significance ($r=-.39$, $t(18)=1.81, p<.09$). (No correlations between N and proofreading performance approached significance.) McCann et al. (1988) observed parallel, significant correlations between pseudohomophones' N and lexical decision performance. We believe N estimates underlying orthographic-linguistic correlations, especially orthographic-phonologic correlations, and covariant learning approximates this structure (compare Andrews, 1992; Seidenberg & McClelland, 1989), although N may be a fairly coarse-grain estimate (compare Treiman, Goswami, & Bruck, 1990).

These analyses confirm the value of control for statistical regularity. Accuracy to pseudohomophone items varied widely and a large portion of this variance travels with statistical regularity. We could very well have missed pseudohomophones' base word frequency effect in the lexical decision and semantic categorization tasks had we neglected this control (proofreading accuracy is less clearly correlated with regularity). Once again, our method reflects our theoretical perspective, and our method better reveals phonology effects.

We hesitate to suggest a rigid form for control of statistical regularity. Accumulating empirical work confirms a variety of correlated dimensions (M. Coltheart, Davelaar, Jonasson, & Besner, 1977; Glushko, 1979; Jared, McRae, & Seidenberg, 1990; Laxon, Coltheart, & Keating, 1988; Rosson, 1985; Taraban & McClelland, 1987; Treiman et al., 1990), and it is difficult to include every dimension as a control variable in every experiment.

Accepting the null hypothesis. Failures of phonologic hypotheses predicated on a symbolic human-information-processing metaphor eroded phonology's role in theories of printed word identification. These failures were also taken as converging evidence supporting "nonphonologic" hypotheses, "the sole alternative" to phonologic hypotheses. Of course, psychology's conventions of hypothesis testing disallow accepting the null hypothesis. In practice, however, highly visible null findings often "confirm" positive hypotheses. The tendency to accept null support is strongest when it confirms obvious hypotheses. Nonphonologic hypotheses are obvious because they are necessary to explain explicit rules' failures to assemble phonology—they are "the sole alternative" to rule-governed phonologic solutions. But, they are "the sole alternative" only if we restrict hypotheses to those within the symbolic framework. Likewise, they only appear obvious from the restricted perspective of the symbolic metaphor and its explicit rule axiom.

The explicit rule axiom was never truly axiomatic. It was always as potentially arbitrary, narrow, and misleading as any other hypothesis—misleading to the point of blindness to "assembled" phonology's pervasive effect on reading performance. Previous failures to observe "assembled" phonology effects are now revealed as failures to penetrate the dependence of hypothesis and method on a questionable underlying metaphor. The absence of positive support for "sole alternative" nonphonologic hypotheses, coupled with repeated failures of phonologic hypotheses, should have alerted theorists to inherent fallacies in the symbolic analysis. Instead, theorists accepted the null hypothesis and reduced or eliminated phonology's role in reading.

Acknowledgment
Preparation of the manuscript was supported by a National Institute of Health FIRST award (CMS 5 R29 NS26247-02) to Guy Van Orden. The lexical decision and categorization experiments originally composed an independent project conducted by Greta Pinnt, the first proofreading experiment was an independent project conducted by Lori Markson, and the second proofreading experiment was part of an independent project conducted by Karen Garlington. We thank Jennifer Morrison and Eric Voldeng for preparing stimulus booklets, running subjects, and coding data for the third proofreading experiment.

References

Andrews, S. (1992). *Frequency and neighbourhood effects on lexical access: Lexical similarity or orthographic redundancy?* (Manuscript submitted for publication).

Balota, D. A. (1990). The role of meaning word recognition. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading.* 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.

Banks, W. P., Oka, E., & Shugarman, S. (1981). Recoding of printed words to internal speech: Does recoding come before lexical access? In O. J. L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology* (pp. 137-170). Hillsdale, NJ: Erlbaum.

Becker, C. A. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance, 2,* 556-566.

Besner, D., Davelaar, E., Alcott, D., & Parry, P. (1984). Wholistic reading of alphabetic print: Evidence from FDM and the FBI. In L. Henderson (Ed.), *Orthographies and reading: Perspectives from cognitive psychology, neuropsychology, and linguistics* (pp. 121-135). Hillsdale, NJ: Erlbaum.

Chastain, G. (1981). Phonological and orthographic factors in the word-superiority effect. *Memory & Cognition, 9,* 389-397.

Chastain, G. (1984). Phonological access to the mental lexicon in a target discrimination task. *Journal of General Psychology, 110,* 53-60.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). London: Academic Press.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535-555). New York: Academic Press.

Coltheart, V., Avons, S. E., Masterson, J., & Laxon, V. (1991). The role of assembled phonology in reading comprehension. *Memory & Cognition, 19,* 387-400.

Coltheart, V., Avons, S. E., & Trollope, J. (1990). Articulatory suppression and phonological codes in reading for meaning. *Quarterly Journal of Experimental Psychology, 42A,* 375-399.

Coltheart, V., Laxon, V., Rickard, M., & Elton, C. (1988). Phonological recoding in reading for meaning by adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 387-397.

Daneman, M., & Stainton, M. (1991). Phonological recoding in silent reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 618-632.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 27-85). Hillsdale, NJ: Erlbaum.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12,* 627-635.

Gerrig, R. J. (1989). The time course of sense creation. *Memory & Cognition, 17,* 194-207.

Gerrig, R. J., & Murphy, G. L. (1991, November). Contextual influences on the comprehension of complex concepts. Poster presented at the meeting of the Psychonomic Society, San Francisco.

Glushko, R. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 361-379.

Hawkins, H. L., Reicher, G. M., Rogers, M., & Peterson, L. (1976). Flexible coding in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 380-385.

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., & Seidenberg, M. S. (1991). Does word identification in reading proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General, 120,* 358-394.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reasoning.* Chicago: University of Chicago Press.

Johnson-Laird, P. N. (1988). How is meaning mentally represented? In U. Eco, M. Santambrogio, & P. Violi (Eds.), *Meaning and mental representations* (pp. 99-118). Bloomington, IN: Indiana University Press.

Kawamoto, A. H., & Kitzis, S. N. (1991). Time course of regular and irregular pronunciations. *Connection Science, 3,* 207-217.

Kawamoto, A. H., & Zemblidge, J. (in press). Pronunciation of homographs. *Journal of Memory & Language.*

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

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago, IL: University of Chicago Press.

Laxon, V. J., Coltheart, V., & Keating, C. (1988). Children find friendly words friendly too: Words with many orthographic neighbours are easier to read and spell. *British Journal of Educational Psychology, 58,* 103-119.

Lesch, M. F. (1990). *Differential effects of exposure duration on semantic priming from homophones: Evidence for Van Orden's verification model.* Unpublished master's thesis, University of Massachusetts, Amherst.

Lesch, M. F., & Pollatsek, A. (1992). *Automatic access of semantic information by phonological codes in visual word recognition.* Manuscript submitted for publication.

Lukatela, G., & Turvey, M. T. (in press). Similar attentional, frequency, and associative effects for pseudohomophones and words. *Journal of Experimental Psychology: Human Perception and Performance.*

Mandler, G. (1984). *Mind and body: Psychology of emotion and stress.* New York, NY: W. W. Norton & Company.

Mandler, J. M. (in press). How to build a baby: II. Conceptual primitives. *Psychological Review.*

Massaro, D. W., Taylor, G. A., Venezky, R. L., Jastrzembski, J. E., & Lucas, P. A. (1980). *Letter and word perception.* Amsterdam: North-Holland.

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 14-24.

McCann, R. S., & Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do word-frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance, 14,* 693-706.

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). The interaction of information in word recognition. *Psychological Review, 76,* 165-178.

Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research, 53,* 13-24.

Patterson, K., & Coltheart, V. (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (421-447). Hillsdale, NJ: Erlbaum.

Pennington, B. F. (1991). *Diagnosing learning disorders.* New York: Guilford Press.

Pennington, B. F., Van Orden, G. C., Smith, S. D., Green, P. A., & Haith, M. M. (1990). Phonological processing skills and deficits in adult dyslexics. *Child Development, 61,* 1753-1778.

Perfetti, C. A. (1985). *Reading ability.* New York: Oxford.

Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and masked priming. *Journal of Memory and Language, 30,* 473-485.

Perfetti, C. A., Bell, L., & Delaney, S. (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language, 27,* 59-70.

Rosson, M. B. (1985). The interaction of pronunciation rules and lexical representations in reading aloud. *Memory & Cognition, 13,* 90-99.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 10,* 645-657.

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-Casas, R. M., García-Albea, J. E., & Bradley, D. C. (1991). On access representation in visual word recognition: The temporal separation technique. *Psychological Research, 53,* 53-61.

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

Stone, G. O., & Van Orden, G. C. (1989). Are words represented by nodes? *Memory & Cognition, 17,* 511-524.

Stone, G. O., & Van Orden, G. C. (in press). Resolving empirical inconsistencies concerning priming frequency, and nonword foils in lexical frequency, and nonword foils in lexical decision. *Journal of Experimental Psychology: Human Perception and Performance.*

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26,* 608- 631.

Treiman, R., Goswami, U., & Bruck, M. (1990). Not all nonwords are alike: Implications for reading development and theory. *Memory & Cognition, 18,* 559-567.

Van Orden, G. C. (1984). *A ROWS is a /ROZ/ is a ROSE: Visual, phonological, developmental and strategic influences upon lexical access.* Unpublished doctoral dissertation, University of California, San Diego.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition, 15,* 181-198.

Van Orden, G. C. (1991). Phonologic mediation is fundamental to reading. In D. Besner & G. Humphreys (Eds.), *Basic processing in reading: Visual word recognition* (pp. 77-103). Hillsdale, NJ: Erlbaum.

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 371-386.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97,* 488-522.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1992). *What do double dissociations prove? Inductive methods and isolable systems.* (Unpublished manuscript.)

Weber, R. M. (1970). A linguistic analysis of first-grade reading errors. *Reading Research Quarterly, 5,* 427-451.

Appendix A

Item data (percent correct rejections) from the lexical decision task (LDT) and semantic categorization task (SCT) to pseudohomophone, spelling control, and bigram control foils constructed from high and low frequency base words.

High-Frequency Base Words			Low-Frequency Base Words		
Pseudohomophones					
	LDT	SCT		LDT	SCT
YELLOE	94.7	84.2	ELBOE	100.0	73.7
SCAIL	73.7	63.2	CRAIN	36.8	36.8
MUNTH	100.0	100.0	MUNKEY	94.7	84.2
WHEAL	73.7	78.9	SLEAT	57.9	57.9
MUZIC	84.2	100.0	PANZY	68.4	47.4
CHERCH	78.9	84.2	DURBY	63.2	52.6
MARKIT	78.9	89.5	CRICKIT	31.6	26.3
SHURT	89.5	78.9	BURCH	73.7	31.6
HEET	63.2	84.2	JEAP	42.1	52.6
FEER	57.9	47.4	SPEER	57.9	10.5
Spelling Controls					
YELLOT	89.5	100.0	ELBOT	94.7	89.5
SCANE	52.6	84.2	CRAFE	78.9	94.7
MONCH	100.0	100.0	MOCKEY	68.4	84.2
WHELL	89.5	89.5	SLERT	84.2	94.7
MUDIC	100.0	100.0	PANDY	68.4	68.4
CHUNCH	47.4	73.7	DERPY	89.5	78.9
MASKET	47.4	100.0	CRISKET	52.6	78.9
SHART	84.2	89.5	BINCH	89.5	100.0
HELT	78.9	100.0	JELP	89.5	89.5
FLAR	94.7	100.0	SPLAR	94.7	94.7
Bigram Controls					
BARREF	100.0	100.0	ADEAT	89.5	100.0
BRANA	89.5	94.7	ROICK	100.0	94.7
SWICK	78.9	100.0	GRABLY	100.0	100.0
WOULL	94.7	100.0	TRUDS	78.9	100.0
BUGID	100.0	100.0	JARAW	100.0	97.4
BEEDAS	94.7	94.7	OPECT	78.9	94.7
BEALIC	89.5	100.0	DINDERS	94.7	94.7
WOIRL	100.0	100.0	BEERT	100.0	94.7
FRON	89.5	100.0	VOMP	84.2	100.0
SACH	89.5	94.7	CHONT	89.5	94.7

Appendix B

Item data (percent correct detections) from the appropriate and inappropriate context conditions of the three proofreading experiments (PRF1, PRF2, and PRF3, respectively) to pseudohomophone, spelling control, and bigram control foils constructed from high and low frequency base words.

Appropriate Context

High Frequency Base Words				Low Frequency Base Words			
Pseudohomophones							
	PRF1	PRF2	PRF3		PRF1	PRF2	PRF3
YELLOE	75	70	85	ELBOE	85	65	75
MUNTH	100	90	90	MUNKEY	90	85	100
SPEEK	65	65	55	SNEEK	40	40	45
DOKTOR	90	95	95	NEKTAR	75	70	75
GREAN	35	50	35	SLEAT	30	40	30
MARKIT	90	60	90	CRICKIT	45	50	40
MUZIC	95	90	90	PANZY	25	30	10
SKORE	90	80	80	SKALP	65	75	65
PURSON	85	75	95	DURBY	25	30	25
FEER	75	65	65	SPEER	60	40	35
Spelling Controls							
YELLOT	65	60	75	ELBOT	80	75	85
MINTH	100	95	100	MOFKEY	100	80	90
SPELK	100	90	95	SNELK	100	90	90
DOSTOR	95	90	95	NESTAR	65	65	65
GRELN	75	75	70	SPEET	75	70	75
MACKET	90	70	90	CRINKET	80	85	90
MUDIC	90	85	85	PAKSY	45	55	60
SCOME	90	75	85	SWALP	75	70	95
PEKSON	80	75	55	DERGY	50	65	75
FLAR	85	75	90	SPLAR	60	70	70
Bigram Controls							
PHOLOP	55	80	65	OCCEM	60	90	90
RURLY	70	80	85	FUNKEL	60	60	45
LIMEW	70	95	90	PUNEW	50	80	75
GEPTOL	60	85	90	DUZTAS	15	75	60
SHEAL	5	20	30	FLESS	50	55	40
SHOWAY	45	70	75	FRECKOM	60	75	75
PIZIL	50	85	80	WANUR	50	75	45
OLMES	25	20	45	JEKEM	50	75	75
MERLLY	45	50	55	KIMPE	25	60	50
SORK	40	60	55	LARLD	35	70	65

Inappropriate Context

High Frequency Base Words				Low Frequency Base Words			
Pseudohomophones							
	PRF1	PRF2	PRF3		PRF1	PRF2	PRF3
YELLOE	80	85	90	ELBOE	60	55	65
MUNTH	70	85	85	MUNKEY	80	85	70
SPEEK	55	60	55	SNEEK	45	50	60
DOKTOR	85	80	80	NEKTAR	45	60	65
GREAN	45	50	60	SLEAT	35	45	55
MARKIT	50	70	70	CRICKIT	45	50	65
MUZIC	85	85	95	PANZY	15	30	10
SKORE	80	80	75	SKALP	45	70	75
PURSON	80	80	85	DURBY	15	55	30
FEER	65	50	75	SPEER	30	40	50
Spelling Controls							
YELLOT	70	80	75	ELBOT	50	70	65
MINTH	70	80	85	MOFKEY	70	85	80
SPELK	55	70	75	SNELK	80	80	80
DOSTOR	60	85	80	NESTAR	55	70	65
GRELN	70	85	65	SPEET	80	65	85
MACKET	50	75	60	CRINKET	70	80	75
MUDIC	70	80	90	PAKSY	65	60	80
SCOME	70	85	85	SWALP	75	90	85
PEKSON	55	55	75	DERGY	65	70	85
FLAR	65	85	85	SPLAR	40	65	70
Bigram Controls							
PHOLOP	30			OCCEM	60		
RURLY	45			FUNKEL	25		
LIMEW	45			PUNEW	40		
GEPTOL	35			DUZTAS	25		
SHEAL	45			FLESS	40		
SHOWAY	45			FRECKOM	50		
PIZIL	30			WANUR	55		
OLMES	35			EKEM	40		
MERLLY	75			KIMPE	40		
SORK	50			LARLD	50		

Appendix C

LUCILLE

I ROUND THE CORNER FROM 42ND STREET TO 8TH AVENUE AND THE FULL BLAST OF THE WIND MAKES TEARS FORM IN MY EYES. THE SPACES BETWEEN THESE DAMN SKYSCRAPERS BLOW LIKE WIND TUNNELS, NEARLY RIPS YOUR EYEBROWS OFF. THIS IS MY FIRST TIME IN NEW YORK, BUT THEN I'VE SEEN A LOT OF FIRST TIMES IN THE LAST TEN YEARS. MOVING IS WHAT I'M DOING, MOVING AND NEVER STAYING.

TOO MANY PEOPLE—NEW YORK IS TOO MANY PEOPLE. SURE, THE ENERGY IS HERE, BUT THAT ENERGY IS FUELED BY PEOPLE ON PEOPLE'S NERVES AND PEOPLE'S NERVES ON PEOPLE'S NERVES. TENSION FEEDS ANXIETY; ANXIETY COMPLETES THE LOOP BY FEEDING TENSION AND THE MENTAL ILLNESS IS SELF PERPETUATING. IF THE NUTS GET TOO CLOSE HERE, THE ILLNESS ABSORBS YOU. MY BEST DEFENSE IS MY RIGHT ELBOE. JABS CREATE SPACE AND SPACE PRESERVES SANITY.

I TRAVEL NOW, BUT I USED TO LIVE SOMEWHERE, WITH LUCILLE. I LOVED HER AND I LOVE HER AND TOMORROW IT WILL BE THE SAME. I IMAGINE THAT SHE LOVED ME WITH THE SAME HEAT, BUT HOW CAN YOU KNOW FEELINGS WHEN ALL YOU GET IS TALK. WE SPENT A LOT OF TIME AT THE ZOO. WHO KNOWS WHY—HELL, WHY ANYTHING? SHE FIRST SAID SHE LOVED ME AS WE WALKED FROM THE PRIMATE HOUSE. SHE ALWAYS STOPPED THERE TO VISIT A FAVORITE MUNKEY NICKNAMED MR. CALABASH. THE NAME WAS ONE SHE'D HEARD ON TV. THE BEAST WAS MERELY PITIFUL AND LONELY, NEVER ABLE TO FIT IN WITH THE REST OF THE TROOP THAT SHARED THE CAGE. MAYBE THAT WAS WHY IT WAS HER FAVORITE, MISFITS ATTRACT EACH OTHER. ANYWAY, MISFITTING NO LONGER MATTERS.

I SPOT A LIKELY BAR AND HEAD FOR ITS ENTRANCE. AS I DUCK THROUGH THE DOOR, I LOOK BACK OVER MY SHOULDER. STUPID, AS THOUGH SOMEONE WOULD RECOGNIZE ME ENTERING A NEW YORK BAR, OR AS THOUGH IT MATTERS. I DON'T KNOW HOW, BUT THE FEELING STILL HOLDS THAT I SHOULD SNEEK INTO BARS—THE RESIDUE OF EARLY GUILT, THE LEGACY OF MY PASTOR FATHER.

I SLIP INTO THE KHAKI GREAN VINYL SEAT OF A BOOTH NEAR THE JUKEBOX. BUT I GET BACK UP TO FEED SILVER INTO THE CIGARETTE MACHINE. I CALL OUT MY ORDER AS I RETURN TO MY SEAT—SCOTCH NEAT, WATER BACK. MY DRINK IS DELIVERED; SWEET SCOTCH, BOTH MY GOD AND THE NEKTAR OF MY GOD. WORSHIP IS SHORT, MY DRINK IS GONE, AND I ORDER ANOTHER. AFTER A HALF DOZEN SACRAMENTS I CAN FACE MY WORLD, MY NEW BIGGEST CITY WORLD.

THE BRIGHT LIGHT OF THE AFTERNOON SUN GIVES ME SECOND THOUGHTS AND I RE-ENTER THE BAR. AS I ORDER ANOTHER DRINK, SOMEONE ASKS ME FOR THE TIME. I LOOK FOR A CLOCK, BUT DON'T FIND ONE. I TELL THE GUY I DON'T HAVE THE TIME AND I JOKE WITH THE TRUTH, THAT I DON'T EVEN KNOW WHAT MUNTH IT IS. ONCE I RECLAIM MY BOOTH, I FEEL AT HOME AGAIN; AND, RELATIVE TO ALL OTHER PLACES IN THIS CITY, I

AM AT HOME. TO CELEBRATE THIS INSIGHT I TOSS BACK MY DRINK AND ORDER ANOTHER. THIS SACRED CELEBRATION CONTINUES AS MY FOCUS NARROWS AND MY WORLD CONVERGES: HOME, CHURCH, FAMILY AND GOVERNMENT ARE NOW ONE WITH ME IN MY NEW BOOTH, WITHIN MY NEW BAR. I FEEL THE PEACE OF KNOWING MY PLACE AND HOLDING MY POSITION. ALL IS AS IT SHOULD BE.

•••••

I MUST HAVE FALLEN TO THE FLOOR AND NOW A CIRCLE OF STRANGE FACES IS STARING DOWN AT ME, BUT I CAN'T FIGURE OUT WHAT FLOOR I'VE FALLEN TO. SOMEONE ABOVE ME ASKS IF I CAN SPEEK AND I'M NOT SURE IF I CAN. I TRY TO ANSWER AND "YES" COMES OUT OF MY MOUTH— I GUESS I CAN TALK. THEY HELP ME TO A BOOTH AND I NOTICE THE EMPTY GLASS AND THE WATER BACK. ONCE I'M STEADY IN THE BOOTH, I'M LEFT ALONE. I HEAR ONE MAN ASK ANOTHER IF THEY SHOULD CALL A DOKTOR, AS THEY WALK BACK TO THE BAR. THE OTHER GUY SAYS THEY SHOULDN'T BOTHER.

MY PRESENT EXPERIENCE OF SITTING IN THE BOOTH INCLUDES NO KNOWLEDGE OF MY PAST. THE OBLIVION THAT PRECEDED THE CIRCLE OF FACES HAS GIVEN WAY TO A LACK OF MEMORY. BUT THE MOMENT OF AMNESIA TOO SOON REFORMS INTO REMEMBERING AND I RE-ENTER MY HELL OF SELF LOATHING AND RECRIMINATION. IN MY GUILT-EDGED MEMORIES, MY CAR SLIPS OUT OF CONTROL, ACROSS THE DOUBLE YELLOE LINES, AND HEAD-ON INTO THE TRUCK. THE TRUCK'S FORCE SHEARS OFF THE FRONT HALF OF THE CAR AT AN ANGLE THAT INCLUDES THE PASSENGER SEAT, AN ANGLE THAT INCLUDES LUCILLE.

THE LUXURIOUS, ELABORATE, CONSUMING GUILT PROBABLY COMES FROM MY DESIRE TO CONTROL. THE NEED THAT DRIVES SELF- DETERMINATION NOW DRIVES MY CLAIM, MY ADDICTION, TO GUILT FOR LUCILLE'S DEATH. MY FRIENDS CONSOLED ME; THEY BLAMED THE ACCIDENT ON THE DARK NIGHT, THE WET, FREEZING SLEAT, AND THE GLASSY ROADS. BUT I KNOW THE TRUTH. I REMEMBER LAUGHING AT LUCILLE'S FEAR OF DRIVING THAT NIGHT. I REMEMBER HER WARM, QUIET VOICE ASKING IF WE MIGHT TRAVEL THE NEXT DAY. I REMEMBER MY ANGER AT HER STUPID PREMONITIONS. I ORDER A DOUBLE SCOTCH AND ANOTHER WATER BACK.

Appendix D

THE HOUR OF LEAST RESISTANCE

IN THE GARDEN, ALL IS QUIET. I MOVE NOT AT ALL, LIKE A HUNTED FAWN, THOUGH I AM THE HUNTER. WHEN I FINALLY ADVANCE, MY MOTION IS CALCULATED AND PRECISE. ADRENALIN SCREAMS THROUGH MY VEINS AS I WATCH AND LISTEN FROM MY VANTAGE POINT. I AM PREPARED FOR MY DESTINY, I AM FILLED WITH PURPOSE, I AM POISED, CALMLY ELECTRIC, BEFORE GOD'S ENEMY.

THE NIGHT AIR CARRIES DISTANT STREET SOUNDS AND SHATTERED BITS OF CONVERSATION FROM THE OPEN WINDOWS OF THE HOUSE. A <u>CRICKIT</u> CHIRPS OFF TO MY LEFT AS I CREEP TOWARD A SIDE DOOR. MY TARGET IS STILL FAR AWAY IF DISTANCE IS MEASURED BY MY CAREFUL PROGRESS, BUT I DO NOT RUSH MY MOVEMENTS. THE EVENT AT HAND NEED NOT COME QUICKLY. I AM FILLED WITH THE WAITING TIME AS I WILL BE FILLED WITH THE ACTS TO FOLLOW.

AN UNFAMILIAR MELODY BEGINS AS A SILHOUETTE ENTERS AND THEN IS FRAMED IN THE WINDOW BEFORE ME. THE <u>MUZIC</u> COVERS THE NOISE OF MY MOVEMENT SO I START ONCE AGAIN TOWARD THE DOOR. INCH BY INCH, NO NEED TO RUSH AND EVERY NEED FOR CAUTION.

A MAN COMES TO THE DOOR AND LOOKS OUT; IT IS MARCOLA. FOR YEARS HE WAS SUPREME RULER—HIS COUNTRY'S POLITICAL GODFATHER. DEATH SQUADS ARMED HIS AUTHORITY. BUT, NOW, HE LIVES IN NEW JERSEY, IN MORRISTOWN, AN ORDINARY <u>PURSON</u> SHEPARDING THE FORTUNE THAT WAS WRUNG FROM HIS HUNGRY PEOPLE. MARCOLA MOVES BACK INSIDE THE HOUSE AND I BEGIN AGAIN MY CAREFUL PROGRESS. I AM GOD'S INSTRUMENT; MARCOLA AND HIS FAMILY MUST DIE.

I FEEL THE GARDEN'S SOFT WARM EARTH BENEATH MY BODY AND I SMELL THE FRAGRANCE THAT WAFTS THROUGH THE ROWS OF FLOWERS THAT NOW SURROUND ME. I PUT MY HAND FORWARD AND, INADVERTENTLY, RIP THE BLOSSOM FROM A <u>PANZY</u>—MY FIRST VICTIM. SUDDENLY I AM OVERWHELMED BY DEJA VU. IT'S NOT POSSIBLE, BUT I FEEL THE GLOW OF THE FAMILIAR, AS THOUGH I HAVE PREVIOUSLY WORMED THROUGH THIS GARDEN. NONSENSE, IT MUST BE AN ILLUSION, AN ILLUSION FUELED BY MY ERODING VIGILANCE, A VIGILANCE THAT I MUST MAINTAIN DESPITE NIGHTS WITHOUT SLEEP AND DAYS WITHOUT FOOD. THAT'S IT, AND THAT'S THE SOURCE OF THE OCCASIONAL LIGHT-HEADEDNESS, AND THAT'S WHAT CLIMBS UP MY SPINE AND STANDS MY HAIR ON END WITH A BURROWING, ITCHING, QUIVER ACROSS MY <u>SKALP</u>. BUT, I CAN'T BE ABSORBED BY THIS PERTURBED EXPERIENCE; MY MISSION REQUIRES ABSOLUTE CLARITY OF MIND. GOD HAS CONDEMNED THIS HOUSEHOLD AND I AM GOD'S EXECUTOR. THE FORCE OF THIS VISION CLEARS MY MIND.

I CLOSE IN ON THE DOOR. THE WINDOWS ARE DARK ACROSS MOST OF THE HOUSE. THE MARCOLA FAMILY HAS RETIRED. I MOVE CAREFULLY UP THE FEW STEPS AND PAUSE BEFORE THIS SACRIFICAL ALTAR TO GATHER MY RESERVES OF BODILY STRENGTH AND

SPIRITUAL ENERGY. I WILL COLLECT PAYMENT UPON MARCOLA'S DEBT OF HORROR, SETTLE THE SKORE FOR AN ENTIRE COUNTRY.

JUST INSIDE THE DOOR, I NOTICE A HAT RACK. MARCOLA'S ENGLISH DURBY HANGS AMONG AN ASSORTMENT OF CAPS. FUNNY HOW, PRIOR TO THE COUP, THIS SILLY HAT BECAME SUCH A HATED SYMBOL OF MARCOLA'S WESTERN WAYS. AS I SCAN THE ROOM, I NOTICE THE MANY AFRICAN ARTIFACTS. THE SUBTLE IRONY OF THIS COLLECTION YIELDS A BITTER IMAGE; THE LEADER OF ONE DESPERATELY POOR COUNTRY PROCURING THE COSTLY HISTORICAL REMAINS OF OTHERS. CARVED IMAGES DECORATE ALL THE WALLS AND ABOVE THE MANTLE HANGS A HUGE SHIELD AND A SPEER THAT IS AT LEAST TEN FEET LONG.

THE STAIRS THAT LEAD TO THE BEDROOMS ARE TO MY RIGHT. MY CAREFULLY CHOSEN SLIPPERS MAKE NO SOUND AS I ASCEND TO THE SECOND FLOOR.

————SOMETHING CUTS INTO MY THROAT. MY BREATH CLOSES OFF. MY HEAD JERKS BACK. ALL IS TERROR. ALL IS FEER. I NEED AIR. I NEED LIFE. SPONGY BLACK OBLIVION SOAKS UP THOUGHT, BIT BY BIT, OUTSIDE IN. I SEE PULSE. I HEAR BLOOD—MY GOD! OH, MY GOD! OH, MOTHER! HELP ME————

MARCOLA'S BODYGUARD IS PLEASED THAT THE STRUGGLE HAS NOT DISTURBED THE HOUSEHOLD. AS ALWAYS, THE PIANO-WIRE GARROTE KILLS DISCREETLY. THE BODYGUARD KNOWS THAT HIS EXCELLENCY MUST RISE EARLY THIS NEXT MORNING FOR HIS APPOINTMENT WITH HIS BROKER IN THE CITY. THEY PLAN TO REVIEW RECENT TRENDS IN THE STOCK MARKIT, TOWARD UPDATING THEIR INVESTMENT STRATEGIES. THE BODYGUARD TAKES CARE TO DISPOSE QUIETLY OF THE ASSASSIN'S CORPSE.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 15

Dual-route Models of Print to Sound: Red Herrings and Real Horses

Kenneth R. Paap
Ronald W. Noel
Linda S. Johansen
New Mexico State University

The intent of this chapter is to provide a defense of dual-route models for reading aloud. It has two major themes. One of these is our view of the essential characteristics of a viable dual- route theory. The thesis advanced is that a model should be considered dual-route if there are functionally independent routines that operate in parallel to generate both addressed phonology based on whole-word units and assembled phonology based on sub-word units. Dual route theory should be viewed as a horse race between two different routines that can be shown to differ in interesting ways.

The second counterpoint is the examination of what have become red herrings in the current evaluation of dual-route theory. These are issues related to dual-route theory, but not essential to dual-route theory. Some of them, such as the role of grapheme phoneme correspondence (GPC) rules, have strong historical links to early motivations for dual-route theory. We argue that the merit of dual-route theory should be judged not in terms of what these models were, but what they are now, and even what they should be. In labeling these issues red herrings we do not wish to imply that other researchers have been deliberately dragging them across the trail in order to distract us from the real merits of dual-route theory, nor do we imply that they fail to understand and anticipate the modifications to dual-route theory that we promote. Nonetheless, we fear that new hounds drawn to the scent of a most interesting topic are being misled. Accordingly, this chapter stands more as a tutorial to newcomers than an advanced lecture to the colleagues with whom we engage in continued debate. Should this endeavor bring the knowledgeable reader to a finer appreciation of the contribution of dual-route theory to our understanding of reading aloud, so much the better.

The essentials of dual-route theory

Figure 1 depicts a generic dual-route model for reading aloud. The generic model makes no strong commitment to the procedures used to either address or assemble phonology. However, a commitment is made to the assumption that there are two routines. One routine involves accessing an addressed phonology based on the spelling of a whole word.

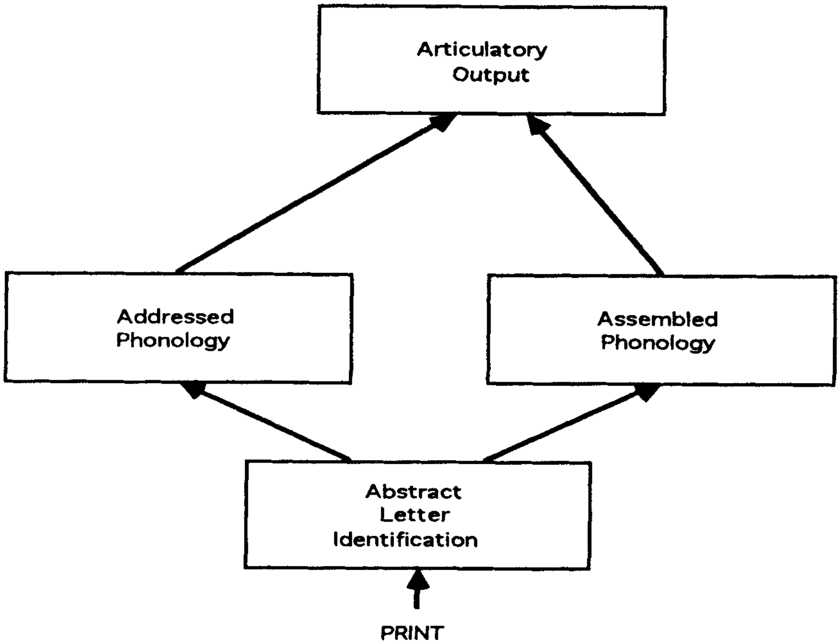

Figure 1. **A generic dual-route theory for reading aloud. Two different procedures race to generate an articulatory code.**

Note that the sequence of abstract letter identities corresponding to a whole word is used to access addressed phonology and that it is incorrect to assume that addressed phonology always relies on a holistic cue like word shape. The other routine generates an assembled phonology that involves the mapping of sub-word units of orthography to corresponding units of phonology. In most cases the two routines operate in parallel, taking as input the abstract letter identities corresponding to the printed word. Under these circumstances reading aloud can be viewed as a horse race between the two routines with the timing and choice of the response determined by the outcome of the race. The two routines must display some degree of functional independence, although their results interact when they compete for the determination of the articulatory output.

If the two routines are to constitute a useful distinction, then they must differ in interesting ways. A procedure that operates on sub-word units must be different from one that considers the whole word. The sub-word routine, however it may be implemented, will need to parse the input string into potential units and will eventually have to assemble or synthesize a whole utterance from some parts. These differences have led dual-route theorists to suggest that the two routines are differentially sensitive to different properties of stimulus structure and that the two routines differ in their attentional requirements. These issues will be taken up in detail in a later section on the independent-processes hypothesis. For now the important point is that there are qualitative differences between the two routines that offer a good heuristic for packaging what we know about reading aloud and for generating new predictions.

Dual process theory can be directed toward the goal of determining the pronunciation of a letter string or the meaning of a printed word. In discussing dual process theory it is important to keep the specific goal in mind since the issues and evidence can be quite different. When researchers contemplate the plausibility of direct access to meaning or the necessity for phonological mediation in lexical access they are focusing on the latter goal. In contrast, this chapter focuses primarily on the task of reading aloud and, accordingly, Figure 1 ignores meaning and is concerned only with the two routines that permit generation of an articulatory output.

Red herring

The GPC and consistency red herring

Early versions of dual-route theory (Coltheart, 1978) proposed that assembled phonology was governed by the application of grapheme-phoneme correspondence (GPC) rules. As we have acknowledged earlier (Paap & Noel, 1991), we assume that a viable dual-route theory based on rules can not be limited to the minimal units that define GPC. In fact a viable rule-based system may require units of various grain sizes, rules of varying strength, and rules that generate conflicting phonological hypotheses. The routine for assembling phonology from this more flexible rulebase will be referred to with the more neutral phrase orthographic-to-phonological conversion (OPC). An OPC rule-based implementation of the generic model is sketched in Figure 2.

In a recent review that critiqued dual-route theory Van Orden, Pennington, & Stone (1990) juxtapose the following two statements. First, in reference to an OPC routine that includes rules that vary in grain size or strength they assert that "....such distinctions merely postpone the demise of dual-process theory, and do not qualify the outcome of our critique" (p. 489). Later on the same page they claim that "Glushko (1979) supplied devastating evidence against the GPC hypothesis. He showed that the consistency of orthographic-phonologic correspondence is a more potent variable in naming performance than GPC regularity." The proximity of these assertions seems to imply that devastating evidence against the GPC hypothesis constitutes equivalent difficulty for an OPC version of dual route theory. Such an implication seems a red herring to us and we will air some doubts regarding first the quality of the empirical evidence and then its true implications for dual-route models. It should also be noted that this red herring runs in a large school and echoes many earlier voices. To add just one more quote, Taraban and McClelland (1987), in referring to the results of Glushko's Experiment 3 showing a disadvantage for regular inconsistent words compared to regular consistent words, state that "Conspiracy models predict a disadvantage for these 'regular' words with inconsistent neighbors, while dual-route models do not, and these results provide one of our main sources of support for such models" (p. 610).

A simple GPC routine assumes that there are all-or-none rules for mapping graphemes to phonemes. All-or-none rules are all of the same strength and are never in conflict with one another, e.g., the grapheme I always corresponds to the phoneme /I/. Furthermore, the rule that maps I to /I/ would have the same strength as the rule that maps B to /b/. If these are the only rules instantiated in a dual-route model, then any given word either conforms to the rules or fails to conform to at least one rule. Thus, the former constitute a set of regular words that can be processed with either routine while the latter constitute the set of exception words that can be correctly named only through addressed phonology.

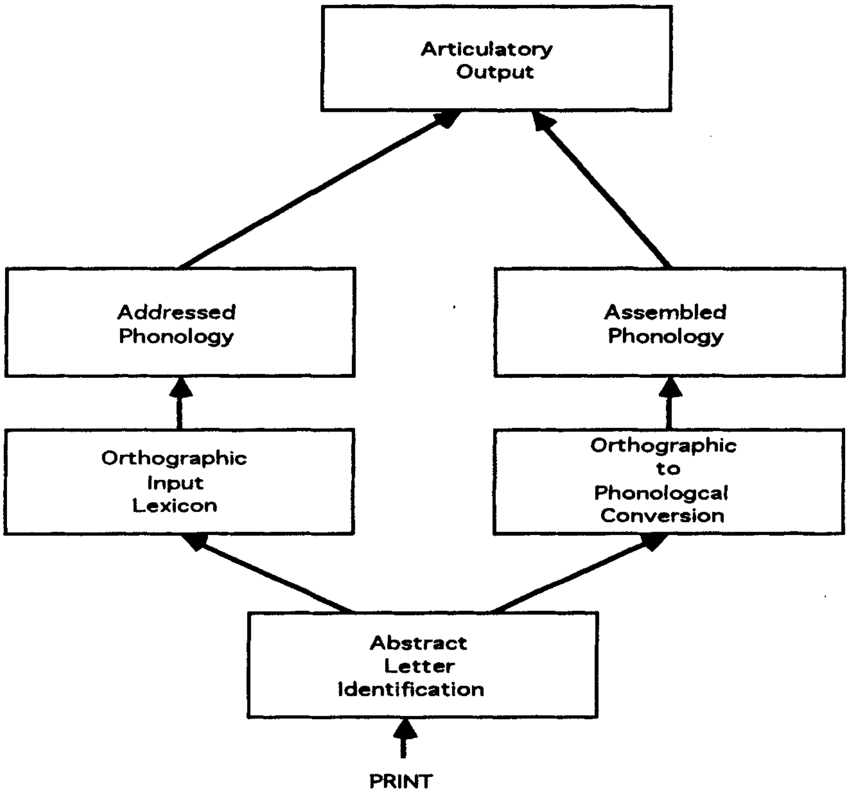

Figure 2. **A rule-based dual-route model in which OPC rules vary in size, strength, and may conflict with one another.**

Thus, PINT is an exception word because the I is pronounced /aI/ and fails to conform to the rule that I corresponds to /I/. Words that conform to the rules constitute the set of regular words, which can be named correctly either by accessing their addressed phonology or by assembling their phonology using the all-or-none rules, while exception words can be named correctly only by accessing their addressed phonology.

Under the assumption that there is some overlap in the distribution of finishing times for the two routines (the time-course section below describes these distributions in more detail), a simple GPC routine provides an account of the regularity effect, i.e., regular words are sometimes named faster than exception words. It also implies that a pseudoword can only be named along the GPC route according to the all-or-none rules that apply. Thus, if the rules fit and pseudowords with similar syllabic structure are considered, there is no basis for expecting differences in naming latency. For example, there would be no reason to predict differences based on the size or consistency of the neighborhood. [However, Coltheart, Curtis, Arkins, & Schreter (1991) have recently described a dual-route computational model based on simple GPC rules that overcomes some of the earlier weaknesses by assuming that partial activation of word units in the lexicon can influence nonword pronunciations that would otherwise be determined exclusively by the rules.]

As noted above, a simple GPC process is incapable of explaining the consistency effects first reported by Glushko. However, it would be prudent to carefully consider the evidence supporting the hypothesis that consistency is a more potent variable in naming performance than GPC regularity, before modifying the simple GPC-based, dual-route model. Consistency refers to the degree to which all the words in a neighborhood rhyme. A standard method of quantifying consistency is to parse a word into its head and body where the head consists of the initial consonant or consonant cluster. The body is then used to search an on-line lexicon to determine all words that have the same body. Consistency for any given word is then the proportion of words in the body's neighborhood that rhyme. Some regular words like HUNT are completely consistent because all the words in the neighborhood rhyme, e.g., BRUNT, BUNT, PUNT, RUNT, and SHUNT. Other regular words like HINT are inconsistent because, although they conform to the dominant pronunciation (e.g., LINT, MINT, and TINT), there is at least one word that is pronounced differently, viz. PINT.

Glushko tested regular consistent words like PUNT, regular inconsistent words like HINT, and exception words like PINT. The naming latencies did show a significant consistency effect with consistent words (529 ms) faster than inconsistent words (546 ms). Furthermore, exception words (550 ms) were not any slower to name than the inconsistent words. However, there was a significant regularity effect in the error data with exception words (8.3%) more error prone than inconsistent words (2.9%).

Since Glushko's seminal work there have been many studies on the regularity and consistency effect and, as usual, researchers discovered that the initial phenomenon was more complicated than it first appeared. For example, the inferiority of a regular inconsistent word (e.g., HINT) may be restricted to words of low-frequency and to conditions where there has been a recent occasion to pronounce its irregular neighbor (e.g., PINT; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). The opportunity for intralist priming by an irregular neighbor was present in Glushko's original experiment. Other studies have failed to use separate controls for both the regular inconsistent and exception words. For example, Paap, Chen, & Noel (1986) used completely consistent controls (e.g, PUNT) that matched the initial phoneme of the exception words (e.g., PINT), but these may have provided a poor baseline for the regular inconsistent words (e.g., HINT) that began with different phonemes. Finally, before on-line dictionaries or word count corpora were commonplace, the selection of completely consistent stimuli was difficult and many lists (e.g., Andrews, 1982) of "consistent" words contained a fair number of inconsistent words that failed to take into account inconsistent neighbors that begin with clusters of two or three consonants.

The culmination of this learning experience was Taraban and McClelland's (1987) careful examination of the consistency effect. Their experiments are the only studies we are aware of that did not allow for the opportunity of intralist priming and had separate controls for both regular consistents and regular inconsistents. The consistency effect evaporates! The difference between regular inconsistents (e.g., TINT) and their matched consistent controls (e.g, TAPS) was +10 ms in Experiment 1A, -4 ms in the intact condition of Experiment 1B, -2 ms in the degraded condition of Experiment 1B, and +4 ms in the unprimed condition of Experiment 2. In contrast, all of the obtained exception effects (+32, +34, and +38 ms) were significant.

The Taraban and McClelland results just reviewed justify the conclusion that consistency is *not* a more potent variable than regularity when reading words aloud. Naming times for exception words always differed significantly from their consistent controls, whereas those for inconsistent words never did. However, three other methods of testing for consistency effects have yielded significant results. One method involves immediate priming of the exceptional pronunciation, a second examines latency differences in the naming of exception words as a function of the number of regular enemies, and the last compares inconsistent pseudowords to consistent controls. We turn to those demonstrations next.

The degree of inconsistency for many of Taraban and McClelland's inconsistent words was quite low. For example, among the 24 low-frequency regular-inconsistents, the neighborhoods of 17 contained only a single exception word. Under the conditions reviewed above this weak manipulation was not able to generate a significant consistency effect. However, when the inconsistent words (e.g., TINT) immediately follow the naming of an irregular neighbor (e.g., PINT), the inconsistent words are now named 16 ms slower than their completely consistent controls. Thus, an exception word prime can interfere with the naming of a regular-inconsistent target. Taraban and McClelland conclude that: "The significant interference effects in this experiment suggest that 'exception' words and 'regular' words in a lexicon are not isolated from each other and incapable of influencing one another" (p. 618).

This red herring pragmatically implies that dual-route theories must hold the products of exception and regular word processing in splendid isolation. This may be true for a dual-route model that uses all-or-none GPC rules, but would not hold for models that embrace rules of varying strength. If the only rule for the grapheme I in a GPC routine was the correspondence to /I/, then PINT would have preactivated this regular correspondence and no interference would be expected. Alternatively, if there are multiple rules of varying strength, then the rule that maps INT to rhyme with PINT may be considerably weaker than the rule that maps _INT to rhyme with TINT, but if the rules are continually updated (strengthened) following every word encounter then the usually weak rule may be in an unusually high state of activation if the subject has just read PINT. This heightened activation should interfere with the normally rapid assembly of the correct pronunciation of TINT.

If the all-or-none regularity distinction is abandoned then the difference between regular-inconsistents and irregular-inconsistents (exception words) is simply the degree of inconsistency. From this framework the fact that Taraban and McClelland observe large exception effects and tiny consistency effects in their first experiment is nothing more than a contrast between high and low levels of consistency. To further explore this possibility, Taraban and McClelland performed a post-hoc analysis and found that exception words with high numbers of enemies were named slower than those with fewer enemies. They claim that this analysis "....is also contrary to the predictions that one would make using a dual-route model, in which exception word pronunciations are read out directly from a lexical entry for the word" (p. 616). The red herring here is the assumption that the addressed phonology of the exception word can be executed as a pronunciation with no influence from the other routine. Dual-route models such as ours (cf. Paap & Noel, 1991) assume that the two routines are in a horse race and that if two competing responses arrive within the photo finish interval, then the assembled phonology can interfere with the prompt execution of the addressed phonology. The magnitude of the interference is determined by the likelihood of a photo finish, and this, in turn, is influenced by rule strength. An exception word with many enemies is racing against a fast OPC horse riding strong rules. An exception word with only a few enemies is racing against a slower OPC horse backed by weaker rules.

Consistency effects are also observed in the naming of pseudowords. In his Experiment 2 Glushko showed that pseudowords formed from a neighborhood of completely consistent words (e.g., BINK) are named significantly faster and with fewer errors (+24 ms, +7.0%) than pseudowords formed from inconsistent neighborhoods (e.g., BINT). The pseudowords were matched in terms of initial phoneme and there was no opportunity for intralist priming. Competing rules of varying strength can also account for these consistency effects. Since the stimulus is a nonword its pronunciation must be assembled from the rule base. Other things being equal, rules for body-sized units that inhabit completely consistent neighborhoods will be stronger than those that inhabit inconsistent neighborhoods. Furthermore, a fairly strong rule that corresponds to the dominant pronunciation of a neighborhood will have to compete with the weaker rules for that neighborhood. Dual-route models with OPC rules that vary in strength and sometimes conflict with one another, like conspiracy models, would certainly predict that consistency effects would be stronger in the naming of pseudowords than words.

The alternative theoretical approach embraced by both Glushko and Taraban and McClelland is a single-route process that relies on lexical analogy. Glushko calls his version activation- synthesis, whereas Taraban and McClelland refer to theirs as a conspiracy model. The key ingredient for these researchers is that the relevant phenomena can be explained without having to appeal to explicitly represented rules that provide a basis for assembled phonology. In contrast, the pronunciation of pseudowords and oftentimes low-frequency words is assumed to reflect the activation and synthesis of a set of candidate words that are orthographically similar to the input string. These models reasonably assume that the higher the degree of consistency in the activated set of candidates the faster a pronunciation can be synthesized. The immunity of high-frequency words to consistency effects is assumed to occur because their pronunciations are derived from only the corresponding lexical entry and not from the set of orthographically similar entries.

Beneath the surface of lexical-analogy theory's apparent parsimony, Glushko spawns the following red herring regarding the proliferation of rules required by an OPC routine: "....it is necessary to postulate a separate specific rule for many spelling-to-sound correspondences embodied by only a single word. The explanation requires the proliferation of rules by the hundreds and perhaps thousands. While some theorists may call patterns of this level of generality 'rules', in doing so they have sacrificed the economy that motivated rules in the first place. In addition a workable system with such specific rules may be indistinguishable from an activation framework that in effect derives the relevant multiletter rules each time it is needed (p. 687)."

Is it unlikely that encounters with PINT will establish a weak rule by which _INT is mapped to /aI/, despite the fact that the rule could be applied to no other word? The fallacy in the framing of this question is the presupposition that it applies to no other word. The rule abstraction mechanism, assuming an OPC routine, should go to work as each exemplar is processed. If the learning mechanism ignored the first instance of a correspondence it would never get off the ground. After all, you never know when you will discover that the name for an OPC rule that applies to only one word is a BINT and that it rhymes with PINT! The rule abstracter should be prepared and create an entry for each novel correspondence. In fact, one might expect the learning rule to generate an exponential growth function such that the first one or two exemplars contribute more gains in strength than the 101st and 102nd.

Would such a rule abstraction mechanism be profligate? A skilled reader may know tens of thousands of words and there may be some economy in learning a thousand rules for generating their pronunciations. More important may be considerations of processing efficiency. Lexical analogy theories perform rule abstraction on the fly when they synthesize a single pronunciation from the set of activated candidates. It may be easier to look up a rule than to invent one.

Taraban and McClelland arrive at a conclusion similar to Glushko's in discussing their third experiment. In this experiment pseudoword targets (e.g., RINT) are primed by either exception words (e.g., PINT) or regular inconsistent words (e.g., TINT). One measure of interest was the latency differences between these two prime conditions when the subject, in fact, pronounced the pseudoword according to the major spelling-sound correspondence. This differences was +59 ms. In addition to this standard type of "rhyme" prime involving the body, other conditions assessed the contribution of the beginning consonant-vowel environment (e.g., PINT - PINF vs. PINK - PINF) or only the vowel (e.g., PINT - TISH vs. PINK - TISH). The latency differences for these two conditions were +36 and +12 ms, respectively.

It appears that you need a consonant context either preceding or following the vowel in order to get significant priming effects, and this has the following implication for Taraban and McClelland: "In order to account for conjunctive effects, one might imagine a dual-route model with very specific and detailed context for spelling-sound rules.... This sort of model might adequately account for the data, but it seems that the number of rules in such a model would need to be quite large.... In some cases activating these rules would be equivalent to activating the words that embody them.... A conspiracy model provides for word and subword information without replicating the information in distinct independent components, as a dual-route model does" (p. 625).

In further addressing this concern we would like to appeal to two principles espoused by Langacker (1987) and Ball (1991) regarding representations in language-processing systems. The first is termed the exclusionary fallacy and refers to a type of conceptual tunnel vision that would lead one to assume that when faced with two or more viable alternative levels of representation the language processing system will always select a single level. Although Langacker and Ball were both concerned with "schema" for representing sentences , the principle applies nicely to an OPC type of dual-route model. The odor of red herring may be in the air whenever there is the implication that OPC rules must be either of GPC size or larger size. A rule abstraction system is likely to start by acquiring specific schemes corresponding to larger units and to refine the rule base hierarchically by retaining commonalities and eliminating unimportant variation. Thus, knowledge of the correspondences for BUNT, HUNT, and PUNT will eventually lead to a rule for _UNT where "_" represents any consonant or consonant cluster. Additional knowledge of other neighborhoods like _UBS, _UCK, and _UMP will eventually lead to the more abstract rule for _U_.

The second principle, the specificity principle, governs the processing assumptions associated with activating schema or rules at multiple levels of representation. In short, one must assume only that the actual input activates consistent representations in proportion to the degree of specific matching. Thus, PUNT activates the lexical unit PUNT the most, larger specific rules like PUN_ and _UNT somewhat less, and the most abstract rules like _U_ the least. The small amount of priming reported by Taraban and McClelland in the vowel only condition (e.g., PINT - TISH vs PINK - TISH) would be consistent with a system governed by these two principles.

The bypass red herring

This brief consideration of how rules might be acquired also speaks to the bypass hypothesis, the second of the three red herrings we sniff in the Van Orden et al. critique. According to these authors "The bypass hypothesis assumes that beginning readers make nearly exclusive use of phonologic mediation but that phonologic mediation is eventually bypassed as direct associations develop between orthographic codes and lexical codes.... (p. 492)" We are not disputing their careful and telling analysis of the bypass hypothesis. In fact we agree with Van Orden et al. that a verification hypothesis provides a better account of the data on homophonic impostors than does the bypass hypothesis.

What we wish to dispel is the implication that the bypass hypothesis forms an essential aspect of dual-process theory. In contrast, in its strongest form it is the antithesis of dual-route. That is, if bypass is taken to mean that most words in the skilled reader's vocabulary are recognized via direct access and that there is no attempt at phonological mediation via assembled phonology, then we have a single-route model, not a dual-route model. Rather, we are mostly in agreement with Van Orden et al.'s conclusion that ".... the results reviewed here are consonant with a theory in which phonologic coding operates in every instance of word identification, irrespective of a reader's familiarity with the word being read" (p. 493). The hedge "mostly" in our statement of agreement is necessitated by our belief that the process that assembles phonology can be strategically turned off. We review the empirical evidence for this belief in the later section on the independent-processes hypothesis.

In their discussion of the covariant learning hypothesis, the heart of the theory preferred by Van Orden et al., they take the attack on the bypass hypothesis one step further and examine evidence that the bypass hypothesis actually has the ontogeny backwards. For example, the developmental study by Zinna, Liberman, & Shankweiler (1986) shows that first-graders display word frequency effects in reading aloud, but that consistency effects and a Consistency x Frequency interaction do not appear until third grade. This pattern is consistent with the assumption that first graders rely on word-specific coding and that lexically-based naming is augmented by rule (or rule-like) behavior only with increased reading skill. Although this pattern is consistent with Van Orden et al.'s framework, for much the same reason, it is also consistent with a rule learning and abstraction process that follows the specificity principle.

The time-course for each race horse

The red herring discussed in the next section, the delayed-phonology hypothesis, requires a fairly sophisticated understanding of the time-course assumptions for our activation-verification dual-route model (Paap, McDonald, Schvaneveldt, & Noel, 1987; Paap & Noel). We review here those assumptions relevant to an account of the Frequency x Regularity interaction. The explanation is readily understood by appealing to a common analogy that views dual-route theories of print to sound as a horse race between a lexical horse for addressed phonology and a non-lexical horse for assembled phonology. The lexical horse runs the lexical route. The pronunciation of a word on the lexical route is generated by first recognizing the word and then looking up its pronunciation, information that is presumably stored with each lexical entry. The non-lexical horse runs the orthographic to phonological conversion (OPC) track to assembled phonology. The letter string is parsed (probably in more than one way) into various sub-word units, corresponding phonological units are activated, and a phonological sequence is assembled that corresponds to the sequence with greatest support.

A clear winner on either track can trigger the naming response. If the OPC track produces a clear winner when an exception word has been presented, then the subject will err, e.g. pronouncing PINT to rhyme with MINT. However, we assume that the track stewards are quite conservative in their assessment of clear winners and usually declare a photo finish. If both horses are delivering the same pronunciation, as would be the case for a regular consistent word like PUNT, then a quick decision can be made. The first or consensus pronunciation is simply executed.

However, when the horses deliver competing pronunciations then the stewards require more time. They are biased in favor of the lexical horse since this horse is a "favorite" in the sense of reliably specifying the correct pronunciation of any word. Although they usually declare the lexical horse to be the winner, the sorting out of the competition takes longer compared to when both horses deliver congruent responses. Thus, interference due to competing responses is the mechanism by which dual-route theories account for the regularity effect.

Frequency effects in the activation-verification, dual-route theory influence only the lexical horse. Word recognition occurs sooner for high-frequency words because verification is a serial comparison operation that, in the absence of associative context, considers candidates in descending order of word frequency (Paap et al., 1982; Paap et al., 1987). Thus, high-frequency words send thoroughbreds down the lexical track while low-frequency words must be delivered by lexical nags.

In the absence of any OPC horses, it is obvious that the winning times posted by high-frequency words will be appreciably faster than those posted by low-frequency words. However, a low-frequency nag on the lexical track will have little adverse effect on naming times if there is a reasonably fast horse running the OPC track. This OPC horse will produce a clear winner (no photo finish) whose winning time will be nearly as fast as the lexical thoroughbreds. Thus, a good OPC horse can severely erode the magnitude of the word-frequency effect. The speed of an OPC horse should be determined by the consistency of the word's spelling-sound correspondence, e.g. words with completely consistent bodies should finish sooner than those that are inconsistent.

These assumptions give an account of the Frequency × Regularity interaction. Exception effects occur for low-frequency words because competing responses are generated in a photo-finish. High-frequency words escape the interference because the lexically-based pronunciation is a clear winner. Frequency effects are smaller for regular-consistent words than exception words because the OPC horse for a regular-consistent word is faster and will beat a slow lexical horse on a higher proportion of the trials.

Figure 3 shows hypothetical distributions of the times required to generate a pronunciation for two levels of frequency on the lexical pathway and two levels of consistency on the OPC route. The first distribution (labeled HF Words) represents the finishing times generated by the lexical route for a variety of high-frequency words. It has a mean of t1 and a variance that is determined by variability among items and in the process itself. The distribution of finishing times for low-frequency words (LF Words) has a mean of t3. The t3-t1 difference corresponds to the expected frequency effect if responses were determined only by the lexical route.

The distribution with a mean of t2, labeled Con Rules, represents the time required by the OPC route to generate a pronunciation for a letter string that activates completely consistent rules, e.g. PUNT. If t2, as indicated in Figure 3, is nearly as fast as t1 and significantly faster than t3; then regular consistent words like PUNT should show frequency effects much less than the t3-t1 difference. As described earlier, these conditions take the lexical nag (LF Words) out of the running in most races since the fast OPC horse (Con Rules) frequently will produce clear winners and the average winning time (t2) for the OPC horse is nearly as fast as the average winning time (t1) for the lexical thoroughbred (HF Words).

The distribution with a mean of t4, labeled Incon Rules, represents the time required by the OPC route to generate a pronunciation for a letter string that activates inconsistent rules, e.g PINT or MINT. The t4-t2 difference represents the cost associated with conflict resolution when the orthography specifies more than one phonology. The t4-t1 and t4-t3 differences influence the magnitude of the expected exception effect for high- and low-frequency words, respectively. For the situation depicted in Figure 3, high-frequency exceptions should be named just as fast as their regular counterparts. This follows since even the fastest times generated on the OPC route for inconsistent rules are considerably slower than the slowest times generated on the lexical route for high-frequency words. In contrast, the distribution for low-frequency words does overlap with that for inconsistent rules and it is quite likely that a regular and competing pronunciation will be output by the OPC route during the photo finish interval. Resolving that competition will produce a large exception effect for low-frequency words.

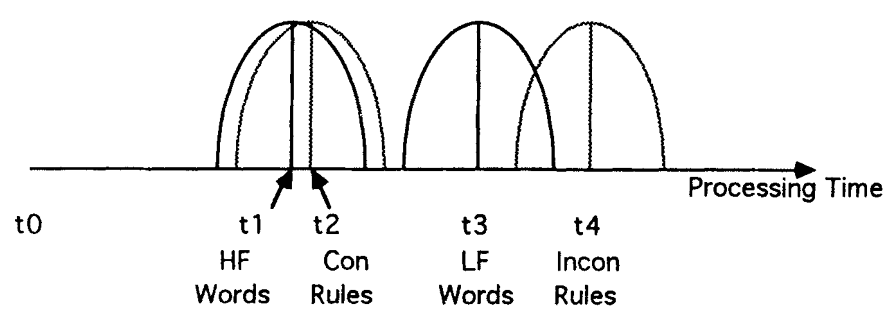

Figure 3. **Hypothetical distributions corresponding to the times required to generate a pronunciation on the lexical and OPC routes. Finishing times on the lexical route are faster for high-frequency words (t1) than low-frequency words (t3). Finishing times on the non-lexical route are faster for consistent rules (t2) than inconsistent rules (t4).**

The expected error rates and naming latencies are not completely determined by the means and variances associated with the four distributions shown in Figure 3, but also require a specification of the photo finish interval. The influence of the photo finish interval can readily be seen by considering the distribution of difference scores between specific pairs of distributions. Figure 4 shows the difference distribution associated with the LF-Words and Incon-Rules distributions. A regularization error occurs when a low-frequency exception word results in a clear OPC winner. The likelihood of this error is represented by the area under the difference distribution to the left of the photo finish interval. Increasing the size of the photo finish interval would further reduce the probability of a regularization error. The exception effect on latency occurs when competing responses are generated during the photo finish interval. A longer photo finish interval will increase the area within the photo finish interval and produce greater interference for low-frequency exception words.

Figure 5 shows the difference distribution associated with a low-frequency word like PUNT that involves the lexical distribution for low-frequency words and the OPC distribution for consistent rules. Because t3 (the mean of the lexical route) is substantially longer than t2 (the mean of the OPC route) the OPC route will produce many clear winners. Furthermore, the consistent pronunciation generated by the OPC route will always be correct for a consistent word. Thus, the area under the distribution to the right of the photo finish interval represents the proportion of trials in which the slow lexical horse (low-frequency words) benefits from a correct response generated by the fast OPC horse (consistent rules).

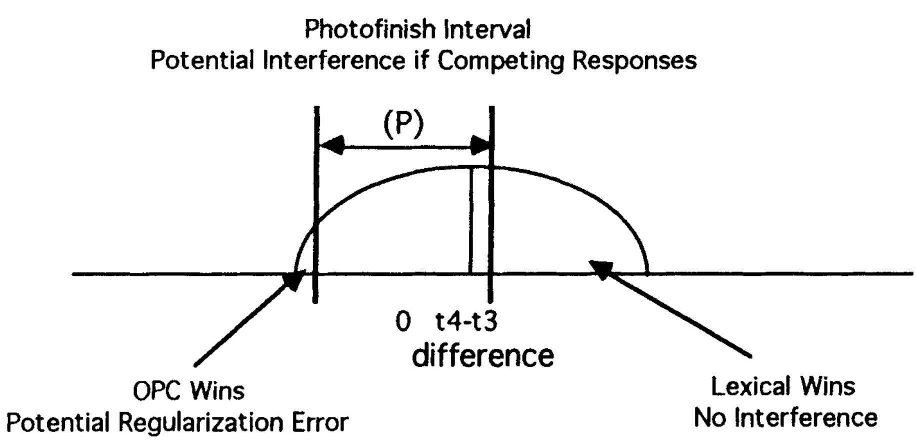

Figure 4. **The distribution of differences sampled from the LF Words and Incon Rules distributions shown in Figure 3.**

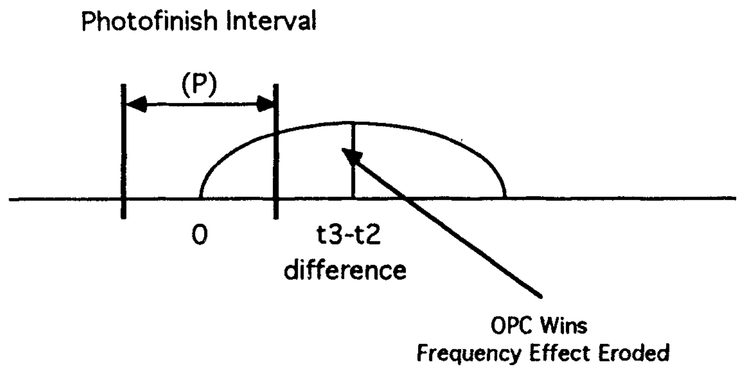

Figure 5. **The distribution of differences sampled from the LF-Words and Con-Rules distributions shown in Figure 3.**

The delayed-phonology red herring

The final red herring which comes from Van Orden et al. is the delayed-phonology hypothesis. This hypothesis assumes that "....phonologic codes are late sources of constraint in lexical coding relative to direct access from orthographic codes" (p. 490). Note that in this context the delayed-phonology hypothesis is cast primarily with regard to the goal of accessing the meaning of a word. Indeed, the discussion of empirical results is restricted to homophony effects in categorization tasks where, presumably, the meaning of a word must be determined in order to do the task. Figure 6 shows an expanded version of an OPC-based dual-route model that assumes that word meanings can be addressed (retrieved) from either orthography, addressed phonology, or assembled phonology.

There is nothing intrinsic to a model like that shown in Figure 6 that compels the assumption that semantic candidates nominated through phonological constraint should arrive much later than those nominated directly by orthography. The route through assembled phonology involves different processes from those involved in direct access and both routines could have similar time courses. As we have just seen, dual-route models like ours must predict that phonological candidates follow hot on the heels of orthographic ones.

Our horse race model was aimed at accounting for naming performance and, accordingly, no explicit assumptions were made regarding the time-course of semantic decisions. However, the horse race that occurs in dual-route theory is assumed to result in many photo finishes between addressed and assembled phonology. For many words assembled phonology occurs in the same time window as addressed phonology, and it follows that homophonic impostors activated via phonology will frequently be activated in time to compete with word codes activated via orthography.

Another interesting challenge to dual-route theory comes from Van Orden et al.'s claim that there is no compelling evidence for a direct route from print to meaning. Within the framework shown in Figure 6 this questions the existence of the pathway from Abstract Letter Identification to the Orthographic Input Lexicon to Addressed Semantics. Could all access to meaning be phonologically mediated? Broad principles of learning theory suggest that this is unlikely. Repeated presentations of the same word should lead to the acquisition of direct connections between orthography and meaning even if access were phonologically mediated at first. (Note that the acquisition of a direct route does not require the dismantling or disuse of the phonologic route, as suggested by the bypass hypothesis.) We resonate to the inevitability of this learning, but concede that others may not.

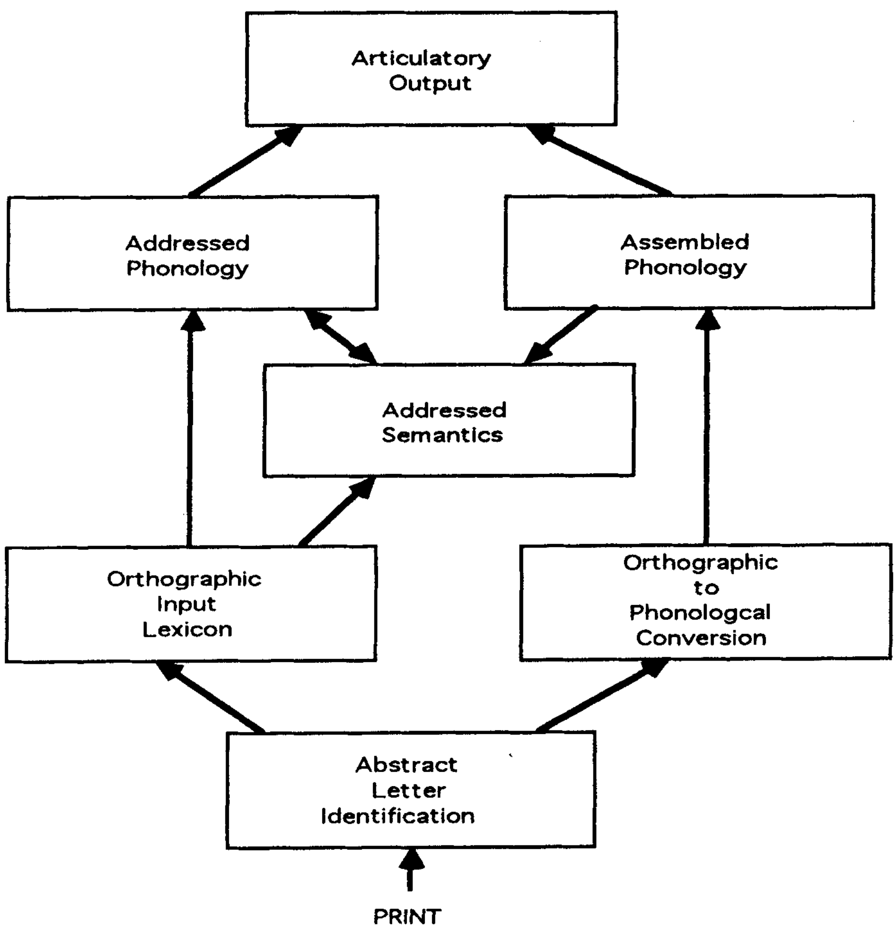

Figure 6. **A dual-route model for both reading aloud and making semantic decisions.**

Dual-route without rules

We have argued that dual-process theory need not embrace rules based on all-or-none GPC's and that a more flexible OPC rulebase is immune to many of the shortcomings of a GPC system. We are perfectly willing to take this a step further and contemplate dual-route models that have no rules at all! The analogy theories sketched in Glushko's activation-synthesis model and Taraban and McClelland's conspiracy theory provide intriguing possibilities. Either of these lexical-analogy theories are, in terms of our criteria, single route theories since there is no horse race. If a high-frequency word dominates the neighborhood, then the pronunciation will be determined solely by this word unit. In contrast, if the stimulus is a low-frequency word, then no word unit gains an early upper hand and the pronunciation will be the product of synthesis or interactive-activation.

As shown in Figure 7, analogy theory becomes dual-process theory if one simply assumes that the addressed phonology associated with the selection of a single candidate races against the assembled phonology synthesized from the set of orthographically similar candidates. The routine that synthesizes a pronunciation from several word candidates is tantamount to abstracting rules on the fly, based only on the relevant instances. For example, the presentation of BUNT might activate a neighborhood consisting of BUNK, BUNT, BUST, HUNT, PUNT, and RUNT. The correct pronunciation should be assembled from these candidates. At the same time that the assembled route is attempting to synthesize a pronunciation from multiple candidates, the addressed route is attempting to verify (cf. Paap, et al. 1982; Paap, et al, 1987) which of the candidates actually corresponds to the input. If and when the verification process finds an acceptable match between an activated candidate and the visual representation of the input, then the word is "recognized" and its addressed phonology can be retrieved.

The standard assumptions of a dual-route horse race can now be applied. If either routine (verification or synthesis) produces a clear winner it determines the response without influence from the other routine. If both routines generate outputs within the photo finish interval then consistent responses will be fluently executed while competing responses will interfere. Competition, for example, will occur if the verification of PINT occurs at about the same time that /pInt/ is synthesized from the neighborhood PINE, PINS, PING, PINK, HINT, LINT, MINT, and PINT. This lexically-based dual-route model, like rule-based dual-route models, assumes the correct pronunciation of exception words is always determined via addressed phonology (verification) whereas pseudowords are always pronounced via assembled phonology (synthesis).

We are developing a computational model of dual-route theory along the lines sketched above. Our examples are not intended to imply a commitment as to how the relevant neighborhood is defined or how the synthesis takes place. Instead of defining a single neighborhood in terms of Coltheart's N, the synthesis process may select candidates that match in terms of the body (e.g., _INT) or the head plus vowel (e.g., PI__). Such a scheme would recruit more relevant information concerning the vowel than the consonants, which may be desirable since the orthographic-to-phonological mapping tends to be less consistent for vowels. Furthermore, both head and body information may be needed to mimic a system of rules like Venezky's (1970), since "regularity" is sometimes conditional upon specific consonant environments such as a preceding W or a following R or L. Finally, if the beginning consonants are synthesized from the head candidates and the ending consonants from the body candidates, then much needless competition for consonant pronunciations can be eliminated. That is, if the initial phoneme for BUNT as input is determined by the candidates BUNK, BUST, and BUNT then there is consistent support for /b/, but if the entire neighborhood were used to assemble the pronunciation then the /b/ unit would have to compete with the initial phonemes of HUNT, PUNT, and RUNT.

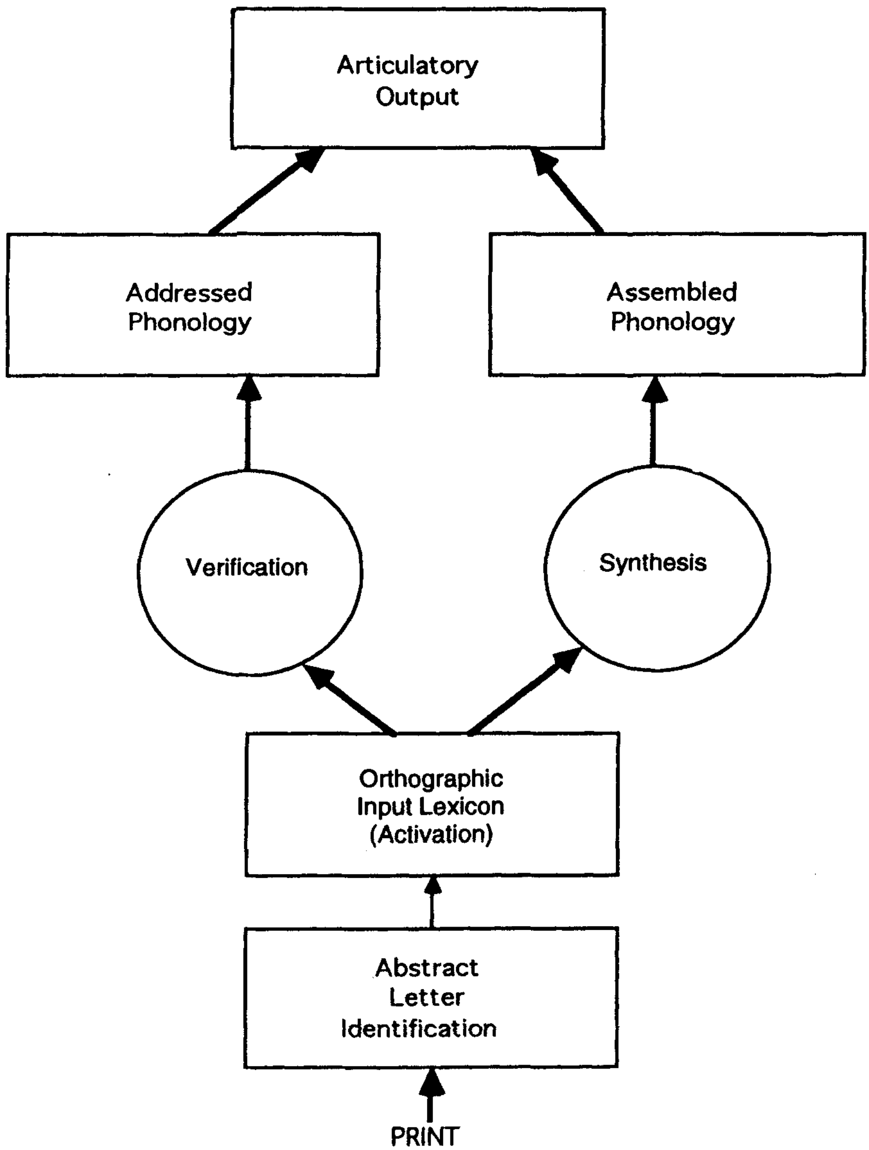

Figure 7. **A dual-route model that abstracts rules on the fly.**

Why complicate basic conspiracy theory by turning it into a horse race? First, the strategy and attentional effects reviewed in the last section of this chapter demonstrate that any model must have some mechanism for dramatically modifying the time-course of processing low-frequency exception words such as PINT. As we shall see, having strategic control over the two routines provides dual-route theory with such a mechanism. A second advantage for dual-route theory may lie in its ability to simulate performance on low-frequency consistent words. For words drawn from completely consistent neighborhoods, low-frequency words are sometimes responded to as rapidly or even more rapidly than high-frequency words. For example, the high-frequency advantage was -7, +13, +6, and -10 ms in three experiments reported by Taraban and McClelland and the low-load condition of Paap and Noel, respectively.

This pattern of fast responses to low-frequency consistents follows from the assumptions of dual-route theory. Completely consistent words have strong rules in an OPC rule-based system and activate completely consistent candidates in a synthesis version. These completely consistent conditions enable low-frequency words, ordinarily lexical nags, to generate a phonology as quickly or even more quickly than the lexical thoroughbreds, high-frequency words.

Not all models can accommodate these results. A conspiracy model based on the principles of interactive-activation may have some difficulty with the speed of low-frequency consistent words. In this model frequency effects occur because high-frequency words have higher resting levels of activation and can gain the upper hand very early in the activation process. They can strangle the competition of similar neighbors before these other candidates gather any momentum and the response is determined only by the addressed phonology associated with the input word. The pronunciation of low-frequency words, on the other hand, is assumed to require the synthesis of a response based on the activation of multiple candidates. Now, since Taraban and McClelland did not actually implement their conspiracy model in a simulation based on interactive-activation we must speculate some here. Suppose that the low-frequency consistent word BUNT activated BUNK, BUNT, BUST, HUNT, PUNT, and RUNT. Could the competition at the phoneme level (e.g., between /b/, /h/, /p/, and /r/ in the initial position) ever resolve fast enough to provide a response as fast as that generated by a high-frequency consistent word like BOOK? The definitive answer, of course, rests on carrying out this exercise in simulation modeling.

It will be difficult to test OPC rule-based theories like those shown in Figure 2 against the synthesis-based theory shown in Figure 7. However, the models may make differential predictions regarding the size of the neighborhood defined by a word's body, particularly for completely consistent neighborhoods. Consider the _ORE body which defines a dense neighborhood of 20 words versus the _OBE body which defines a sparse neighborhood of only four words. It seems clear that an abstraction process for learning OPC rules would acquire stronger rules for _ORE than _OBE. It seems less clear that a synthesis process would determine an output faster from 20 candidates than 4 candidates and one could advance the case that the opposite should be true. In a recent experiment conducted in our laboratory we found that pseudoword naming was more rapid when the body defined a smaller neighborhood than a larger neighborhood. However, the small neighborhood advantage reversed for a second group of subjects who named words as well as nonwords. Furthermore, Bowey & Shatte (1991) recently reported a large-neighborhood advantage when subjects were reading only nonwords. Thus, although the evidence appears quite contradictory at this time, neighborhood size may eventually emerge as an important tool for adjudicating between rule-based and synthesis-based models of print to sound.

Independent-processes hypothesis

Earlier we expressed the conviction that the independent-process hypothesis is the fundamental assumption of dual-process theory. We agree with Van Orden et al. when they state that the viability of dual-process theory depends on the demonstration of separate routines for addressing and assembling phonology.

Recently, we have shown that the two routes can be disassociated on the basis of their attentional requirements (Paap & Noel). The logic of the design of these experiments adopted the view that attention can be viewed as the selective allocation of a limited and central supply of processing capacity (Norman & Bobrow, 1975). Up to some resource limit, it is assumed that processes that require attention will execute faster and more accurately as more capacity is devoted to that process. Processes vary in terms of their resource requirements. Automatic processes are said to require little or no capacity for optimal performance, whereas controlled processes are more sensitive to the available resources. Controlled processes are also assumed to be under the intentional control of the subject, whereas automatic processes tend to be more obligatory. In terms of resource requirements, the routine for addressing phonology was assumed to be more automatic, while the routine for assembling phonology was assumed to be more controlled. The two routes were also assumed to differ on the intentional-obligatory dimension with assembled phonology more easily influenced by strategic factors.

In Experiment 1 the Frequency x Regularity interaction was examined under dual-task conditions. The primary task required subjects to retain in memory either five digits or only a single digit. The secondary task, naming, occurred under either high or low levels of memory load. The single-digit load was assumed to be trivial and, accordingly, the familiar pattern of interaction was expected and obtained: a strong exception effect only for low-frequency words and a strong frequency effect only for exception words.

Next, consider the effects of a high memory load. Retaining five randomly selected digits with near-perfect accuracy is assumed to be very capacity demanding and should severely reduce the resources available to the naming task. If the routine for assembled phonology is more resource demanding, then the reduction in available resources should handicap the assembled horse more than the addressed horse. The consequence of this for naming time depends on the type of word. Naming times for low-frequency exception words should be faster under high memory load. This follows from the assumption that the exception effect is caused by competing responses being generated within the window of a photo finish. If the assembled horse is slowed significantly more than the addressed horse, then the addressed horse will finish a clear winner and evade the competition. This long-shot prediction was confirmed as low-frequency exception words were actually named 39 ms faster under high memory load as compared to low load!

The differential interference caused by the high memory load also had special influence on the low-frequency consistent control words. As discussed earlier, these words are typically named almost as fast as those of higher frequency and, indeed, under low load they were actually named 10 ms faster. We attributed this to the frequent generation of fast responses on the assembled phonology route. However, if the assembled horse is very vulnerable to the effects of memory load, then naming time will have to rely on the slower low-frequency addressed horse. These words should show more secondary-task interference than high-frequency consistent words. Thus, a second major prediction was that a high memory load should generate a substantial frequency effect for the regular-consistent words, an effect that is usually small or absent. As predicted, a frequency effect emerged under high load with the low-frequency consistents (708 ms) taking longer to name than the high-frequency consistents (680 ms).

We also ran a second experiment to provide converging evidence for the assumptions that the assembled routine is less automatic and more controlled than the lexical routine. Specifically, Experiment 2 tested the hypothesis that the assembled routine is more resource-demanding and that its use is under the strategic control of the subject.

In the typical version of the naming task where exception words occur on no more than half the trials, we assume that both routes process the printed word on every trial. However, if a group of subjects encounter only exception words there is no advantage in running the assembled routine because it will never generate a correct response. If assembly is under the subject's control, she should shut it down under circumstances where assembled phonology always yields a competing and incorrect response. This should have two consequences. First, in comparison to a more typical group that experiences both regular and exception words, the naming times for an all-exception group should be faster than the naming times for exception words in a mixed group. Indeed, latencies to the exception words that occurred on critical trials were significantly longer in the mixed group (559 ms) than in the all-exception group (484 ms). This outcome supports the hypothesis that using the assembled phonology route is under the intentional control of the subject.

Second, the overall processing requirements of the naming task should be fewer for the all-exception group, where only the addressed routine is running, compared to those of the mixed group, which must feed both the addressed routine and the resource-hungry assembled routine. If this analysis is correct, then the naming task can be used as a primary task to predict differential amounts of secondary task interference. To this end, Experiment 2 combined the naming task with a tone probe task. On the critical trials a tone was presented 50 ms after the onset of the to-be-named word. Baselines for secondary-task performance were established for each group by including trials in which the probe was coincident with a visual rectangle instead of a printed word. The all-exception group generated +42 ms of secondary-task interference compared to +125 for the mixed group. This supports the hypothesis that the all-exception group was able to turn off the assembled routine and that the assembled routine is demanding of processing resources.

The two experiments reported by Paap and Noel supply impressive evidence for the independent-proccesses hypothesis. The results support the assumption of functionally different routines for assembled and addressed phonology, greater resource demands for assembled phonology, and response competition as the cause of the regularity effect. Replications and extensions of Experiment 1 have been attempted in other laboratories. The good news is that both Herdman (personal communication, November 23, 1990) and Bernstein and Carr (1991) have replicated the counter-intuitive result that high digit loads actually speed the naming of low-frequency exception words. However, when the memory load involves visual memory (Herdman; Bernstein &Carr) or pseudowords (Bernstein & Carr) the enhancement does not occur.

These new results suggest that our original predictions based on a limited and general supply of processing resources is suspect. Paap and Noel anticipated this possibility when they answered the question: What if resources are soup stones? They allow that the influence of one task on a concurrent task need not reflect changes in the available pool of processing resources. That is, dual-task effects can be caused by some form of outcome conflict (Navon, 1984, 1985). "For present purposes two points are highly relevant: (a) the dual-task effects reported in the present experiments could be due to outcome conflict rather than to resource competition, and (b) the new empirical support for the dual-route model remains, regardless of whether the dissociation is generated through resource competition or outcome conflict (Paap & Noel, 1991, p. 22)."

Recent experiments by Baluch and Besner (1991) and in our own laboratory provide additional support for a dual-process model in which functionally independent routines can be strategically disabled. Although we assume that the addressed route tends to be obligatory and that the assembled route tends to be under intentional control, these new results are consistent with the idea that under special circumstances the addressed routine can be turned-off.

Baluch and Besner took advantage of a special property of the Persian language. Some words, termed opaque, are typically written without any specification of the vowel, whereas others, termed transparent, have the vowels explicitly marked. Within a dual-route framework the pronunciation of opaque words should occur via addressed phonology, whereas the transparent words are excellent candidates for fast assembled phonology because the orthography is very shallow when vowel information is provided. Consistent with this hypothesis, a first experiment showed that opaque words were named 21 ms faster when preceded by a semantically-related prime compared to an unrelated prime. Semantic priming should occur when naming is governed by lexical access. In contrast, the transparent words showed a non-significant +2 ms priming effect. This is likely to occur if the response is almost always determined by a non-lexical route that generates assembled phonology. A key aspect of Experiment 1 was that subjects also named nonwords. Experiment 2 deleted the nonwords and significant priming effects of +31 and +21 ms were obtained for opaque and transparent words, respectively. In the absence of nonwords subjects could disable the routine for assembled phonology and rely exclusively on addressed phonology. Thus, transparent words were now processed lexically and produced significant priming effects. This type of strategic control over the routine for assembled phonology is similar to that obtained by Paap and Noel in their Experiment 2 for the group that saw all exception words.

The novel result reported by Baluch and Besner occurs in Experiments 3A and 3B. In these experiments word frequency was manipulated rather than semantic priming. When subjects named only transparent words they appeared to rely on the lexically-based addressed phonology, since high-frequency words were named 35 ms faster than low-frequency words. This is somewhat surprising since there were no opaque words to emphasize the addressed routine and pronunciation of transparent words is easily assembled. More surprising is the impact of adding nonwords to the list containing transparent words. This reduced the frequency effect to a non-significant +11 ms.

Within dual-route theory the elimination of the significant frequency effect by adding nonwords to the experiment could be explained two ways. First, the nonwords may trigger a greater allocation of processing resources to the routine for assembled phonology. If the route to assembled phonology is now riding a faster horse it may be producing a proportion of outright winners high enough to subvert the frequency effect. A second possibility is that the lexically-based route to addressed phonology had been simply disabled. This possibility has considerable attraction for us in light of a recent experiment from our laboratory.

If assembled phonology is more resource demanding than addressed phonology, then naming nonwords should be more susceptible to secondary-task interference than naming words. To test this hypothesis we conducted an experiment similar to Experiment 1 from Paap and Noel (1991) where the naming task was embedded in the retention interval of a memory task. However, instead of testing exception and regular-consistent words the new experiment tested pseudowords and regular-consistent words. Otherwise the apparatus, method, and procedure were the same.

Both the words and the nonwords named in this study were consistent in terms of their spelling-sound correspondence. Consistency was operationally defined as follows. The body (sometimes referred to as the rime) of any monosyllabic word is the letter sequence remaining after the initial consonant or consonant cluster is removed. Word sets that shared the same body were retrieved from an on-line lexicon. The lexicon was the union of all 3, 4, 5, and 6 letter monosyllabic words from the Thorndike-Lorge (1944) and Kučera & Francis (1967) word counts. Nonwords and words that serve only as proper nouns were removed. The number of words that share the same body will be referred to as the bodycount to avoid potential confusion with Coltheart's N. A bodycount neighborhood is completely consistent if all its members rhyme.

The 20 low-frequency words used all come from completely consistent bodycount neighborhoods. The median bodycount is 11.0 and the median word frequency is 5.85. Eighteen of the 20 high-frequency words are completely consistent, but FLAT and WALL each have one exception word (viz. WHAT and SHALL) in otherwise large rhyming bodycounts of 15 and 14, respectively. The median bodycount for the high-frequency words is 12.5 and the median word frequency is 214.0.

For 39 of the 40 cases nonwords were formed from each word by maintaining the same initial letter (and phoneme) and changing one letter from the body. In the remaining case both of the last two letters were changed. The intent was to form a nonword that was orthographically similar to the base word, began with the same phoneme, and had a body that was easy to pronounce by rule. Nineteen of the 20 nonwords formed from the low-frequency words were completely consistent, while 16 of the 20 nonwords formed from the high-frequency words were completely consistent. The median bodycounts were 5.0 for those formed from low-frequency words and 5.5 for those formed from the high-frequency words. Thus, the median bodycounts for the nonwords were only about half of those associated with the words. The 20 odd-numbered subjects received the odd numbered items under Load 1 and the even numbered items under Load 5. The reverse assignment was made for the 20 even numbered subjects.

Performance on the memory task was uniformly high. Subjects were 94% accurate when the memory load was one digit and 93% accurate when the memory load was five digits. Percent correct on the memory task was not affected by either the lexicality or frequency of the to-be-named stimulus.

Latencies in the naming task were submitted to an analysis of variance with lexicality (word versus nonword), frequency (low vs high), and memory load (one digit versus five digits) as within-subject factors. All of the main effects were significant, but most important was the significant Lexicality x Frequency x Load interaction, $F(1,39) = 5.17, p < .05$, shown in Figure 8.

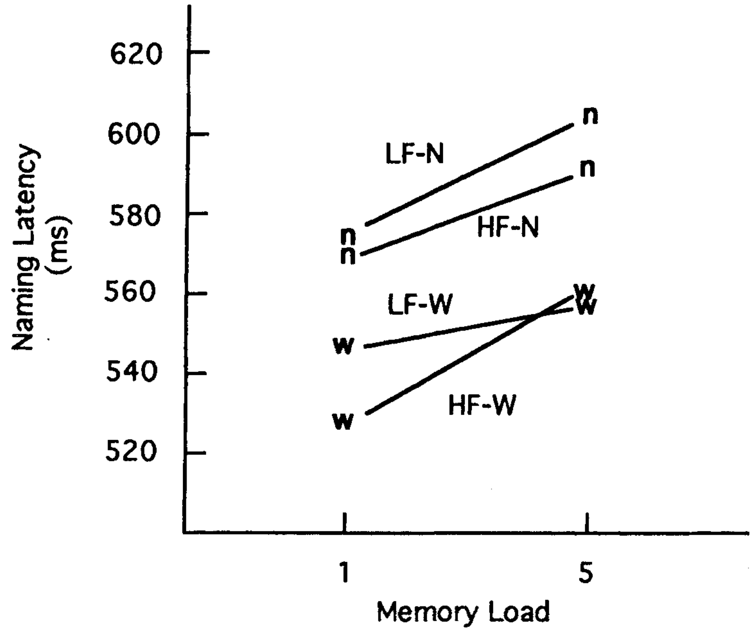

Figure 8. **Naming latencies as a function of memory load for four types of stimuli: high-frequency words (HF-W), low-frequency words (LF-W), nonwords formed from high-frequency base words (HF-N), and nonwords formed from low-frequency based words (LF-N).**

Contrary to our expectations the nonwords were not more vulnerable to secondary-task interference than words. The mean memory load effect of +27 ms was not significantly greater than the mean load effect of +22 ms for the words. Furthermore, high-frequency words, that which were expected to produce the smallest load effect because their pronunciation is primarily dependent upon the less resource-demanding routine for addressed phonology, actually generated the largest load effect. However, these expectations were all based on the original assumption that the memory task and naming task were tapping into a limited and common pool of processing resources, and the results of the Herdman study and that of Bernstein and Carr have already cast doubt on the central-resource assumption. If the two tasks interfere with one another through outcome conflict rather than competition for scarce resources, then perhaps it is reasonable to expect that rehearsing digits, which after all are common words, should interfere as much with the naming of words as nonwords.

Inspection of Figure 8 shows that a frequency effect is present for the words under low load, but that it disappears under high load. One possible explanation is that the outcome conflict is very specific, i.e., not only do words in memory interfere more with the naming of words than nonwords, but also high-frequency words interfere more with the naming of high-frequency words than low-frequency words. This seems unlikely since the vulnerability of high-frequency words to high memory load is specific to this experiment. It was not observed in the first experiment of Paap and Noel or either of the replications that used digits as the memory material.

An alternative explanation meshes nicely with the Baluch and Besner results described earlier. Suppose that under low load there are sufficient resources to fuel both the routines for addressed and assembled phonology. When the lexically-based route to addressed phonology is operating, a small frequency effect, in this case +20 ms, can occur despite the fact that it is likely to be somewhat attenuated by occasional winners on the route to assembled phonology. (Recall that the frequency manipulation was quite strong, with median frequencies of 214 and 5.85 for the high and low-frequency words, respectively.) Operating both routines is not logically necessary under these conditions: the nonwords necessitate the use of the assembled phonology route, but since the words are all regular they too could be pronounced without reference to the lexicon. However, the presence of a frequency effect is good evidence that the routine for addressed phonology is also in operation.

In contrast, when the memory load is high, the naming of high-frequency words is interfered with more than the naming of low-frequency words and the result is the elimination of the frequency effect. This pattern can be explained if we assume that subjects are being put between the proverbial rock and a hard place and abandon the use of addressed phonology. The presence of nonwords makes it necessary to continue to use the resource-demanding routine for assembled phonology. The high memory load makes resources scarce.The only option for conserving resources is to disable the routine for addressed phonology. Thus, although the lexical route is quite automatic and is rarely subject to strategic influence it may, in this unusual case, be turned off. This accounts for both the absence of a frequency effect and the fact that high-frequency words are hurt more by the high memory load than the low-frequency words.

Our results, and those of Baluch and Besner, suggest that readers may be able to disable the lexical route to addressed phonology. Pending a converging operation that further supports the disablement of the lexical route, we consider this as an intriguing working hypothesis. If there is strategic control over the operation of the addressed routine, one might anticipate that it would be rarely exercised since skilled readers spend most of their reading time reading for meaning. Since the direct route to the lexicon offers a fast and reliable pipeline to the meaning of high-frequency words in general, and high-frequency exception words in particular, it may have the characteristics of a highly automated process. Even when the task is simply reading English words aloud and all of the words are consistent, it may take more (e.g., a high memory load) than a logical possibility to induce subjects to abandon the lexical route. In contrast, with a shallow orthography like Persian, abandoning the lexical route may occur more readily.

In closing this section it should be noted that attentional and strategic effects discussed above are supported by the well-known dissociations reported for acquired dyslexics. In a recent summary Coltheart, et al. compared the success of dual-route theory in accounting for this data to the PDP model developed by Seidenberg and McClelland (1989). [See Paap & Noel for a discussion of how the PDP model fares in its ability to explain the attentional effects reviewed in this section.]

Following brain damage, surface dyslexics display normal nonword reading, but many exception words are read incorrectly. Furthermore, the incorrect responses would be predicted from the application of spelling-sound rules (e.g., pronouncing PINT as if it rhymed with MINT). Two of the clearest cases are MP (Bub, Cancelliere, and Kertesz, 1985) and KT (McCarthy & Warrington, 1986). Dual-route theory can account for this pattern by simply assuming that the lexical route to addressed phonology has been damaged and that the rule-based route to assembled phonology is intact. Attempts have been made to simulate surface dyslexia by deleting hidden units from the trained PDP model. These attempts have not succeeded and it seems highly unlikely that they ever will succeed since even the intact model reads nonwords very poorly. Besner, Twilley, McCann, and Seergobin (1990) determined that the intact PDP model gets only about 60% of the nonwords correct compared to 90% for normal readers. Yet on a somewhat different set of nonwords the dyslexic patients are 95% correct at reading nonwords! The predicament for the PDP model is to find a form of damage that will actually make the intact model much better at reading nonwords.

Phonological dyslexics show an opposite pattern of preservation and impairment: word reading is still good, but nonword reading is very bad. A good case was reported by Funnell (1983). Her patient could not read any nonwords at all and, in fact, could not even give the sounds that correspond to individual letters. Nonetheless, this patient achieved scores around 90% correct in tests of word reading. Dual-route theory provides an acceptable account of phonological dyslexia if the routine for assembled phonology was damaged while the routine for addressed phonology was spared.

The implemented portion of the PDP model can not account for phonological dyslexia. However, Seidenberg and McClelland suggest that their general model could explain the phenomena. The general model includes connections between orthography and meaning through another set of hidden units. The meaning units are, in turn, connected to the phonological units and hence an alternative pathway exists for reading aloud that goes from orthography to meaning to phonology. Suppose the direct connections between orthography and phonology were damaged, but the route mediated through meaning was still intact. Words, which have meaning, could still be named; but nonwords that lack meaning would have no connections to phonology. But, according to Coltheart, this explanation fails because in the case of phonological dyslexia reported by Funnell, the patient also had a semantic impairment and would certainly have shown semantic confusions if reading aloud was mediated by semantics. Since the patient did not make such confusions Coltheart concludes that the Seidenberg and McClelland reconciliation of phonological dyslexia with their model cannot be correct.

In summary, the independence-processes hypothesis is supported by several demonstrations of attentional and strategic effects. These include the results that low-frequency exception words are named faster with a memory load of five digits versus one digit (Paap & Noel; Herdman; and Bernstein & Carr) and when subjects are presented with only exception words compared to both exception words and consistent words (Paap & Noel). These results can be explained by assuming that the routine for assembled phonology is a more controlled process that requires more resources and can be strategically disabled. We have also seen that a frequency effect can be eliminated when transparent Persian words are read aloud by merely including nonwords in the experiment (Baluch & Besner). The presence of nonwords is insufficient to eliminate the frequency effect in English, but regular words read against a background of nonwords and under a high concurrent memory load also fail to show a frequency effect. The absence of frequency effects under these conditions is consistent with the view that the lexical route to addressed phonology can also be strategically disabled. These phenomena provide a strong challenge to single-route theories that assume that exception words, consistent words, and nonwords are all read by the same process.

Acknowledgment
Preparation of the manuscript was supported in part by a grant from the National Science Foundation (BNS-9023472) awarded to the first author. Although they might find it hard to believe, visits by Greg Stone and Guy Van Orden to New Mexico State University were very helpful in shaping many of the arguments. The final draft also benefited from a timely workshop on dual-route theory at the University of Arizona sponsored by Ken Forster; Max Coltheart, as always, renewed the enthusiasm of dual-route proponents.

References

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition, 10,* 565-575.

Ball, J. T. (1991). *A computational psycholinguistic model of language comprehension based on a relational analysis of written English.* Unpublished doctoral dissertation. New Mexico State University, Las Cruces, New Mexico.

Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 644-652.

Bernstein, S. E., & Carr, T. H. (1991). Dual-route theories of reading: When does a concurrent task facilitate pronunciation? 32nd Annual Meeting of the Psychonomic Society, San Francisco, CA.

Besner, D., Twilley, L., McCann, R. S., & Seergobin, K. (1990). On the association between connectionism and data: are a few words necessary? *Psychological Review, 97,* 432-446.

Bowey, J. A., & Shatte, A. J. (1991). Orthographic rime frequency effects are independent of word frequency. 32nd Annual Meeting of the Psychonomic Society, San Francisco, CA.

Bub, D., Cancelliere, A., & Kertesz, A. (1985). Whole-word and analytic translation of spelling to sound in a non-semantic reader. In Patterson, K., Marshall, J. C., and Coltheart, M. (Eds.) *Surface dyslexia: Cognitive and neuropsychological studies of phonological reading.* London: Lawrence Erlbaum Associates Ltd.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.) *Strategies of information processing* (pp.151-216). London: Academic Press.

Coltheart, M., Curtis, B., Atkins, P., & Schreter, Z. (1991). Computational modeling of reading: Dual or Single Route? 32nd Annual Meeting of the Psychonomic Society, San Francisco, CA.

Funnell, E. (1983). Phonological processes in reading: New evidence from acquired dyslexia. *British Journal of Psychology, 74,* 159-180.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 104-115.

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

Langacker, R. (1987). *Foundations of cognitive grammar.* (Vol. 1). Stanford, CA: Stanford University Press.

McCarthy, R. & Warrington, E. K. (1986). Phonological reading: phenomena and paradoxes. *Cortex, 22,* 359-380.

Navon, D. (1984). Resources - A theoretical soup stone? *Psychological Review, 91,* 216-234.

Navon, D. (1985). Attention division or attention sharing? In M. Posner & O. Marin (Eds.), *Attention & performance XI* (pp. 133-146). Hillsdale, NJ: Erlbaum.

Norman, D. A., & Bobrow, D. J. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44-64.

Paap, K. R., Chen, L., & Noel, R. W. (1987). Word recognition: Is the sky falling on top-down processing? Paper presented at the annual meeting of the Psychonomic Society, Seattle, Washington.

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 & performance XII* (pp. 221-244). Hillsdale, NJ: Erlbaum.

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: The word-superiority effect. *Psychological Review, 89,* 573-594.

Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research, 53,* 13-24.

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.

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26,* 608-631.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words.* New York: Columbia University Press.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97,* 488-522.

Venezky, R. L. (1970). *The structure of English orthography.* The Hague, Netherlands: Mouton.

Zinna, D. R., Liberman, I. Y., & Shankweiler, D. (1986). Children's sensitivity to factors influencing vowel reading. *Reading Research Quarterly, 21,* 465-480.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 16

Strategies and Stress Assignment: Evidence from a Shallow Orthography

Lucia Colombo
University of Padua

Patrizia Tabossi
University of Bologna

Reading is a well practiced activity for a normal adult, at least in some societies and social classes; the processes involved in reading words have accordingly become quite automatic. In one classical view of the reading process, the identification of a word is always phonologically mediated (Rubenstein, Lewis, & Rubenstein, 1971; Van Orden, Pennington, & Stone, 1990). As opposed to this view, access to meaning can be accomplished through the use of a word-specific process exploiting the direct association between a whole-word orthographic pattern and its meaning. The outcome of this process has been termed direct access (Baron & Strawson, 1976; Coltheart, 1978). Following experimental evidence favoring both positions, the dual route model has been proposed (Carr & Pollatsek, 1985; Coltheart, 1978, 1980; Forster & Chambers, 1973; Morton & Patterson, 1980; Paap, McDonald, Schvaneveldt, & Noel, 1987). In this type of model, reading words is based on multiple routines linking print, sound and meaning. More precisely, the lexical routine is the process which exploits the direct association between a whole word orthographic pattern and its meaning, on one hand, and the word's pronunciation, on the other hand. There is also an indirect way of accessing meaning, through the use of phonology. In this case the system exploits a sublexical routine, which relies on so called spelling-sound correspondence rules, or grapheme-to-phoneme conversion rules (GPCs), that is, associations between smaller-sized orthographic units, such as letters or clusters of letters, and the corresponding sounds. Phonology obtained in this way has been termed "assembled," to distinguish it from "addressed" phonology where the word pronunciation is retrieved from the lexicon accessed directly from the orthographic pattern of the word. The output of the sublexical routine can be used to start articulation or to access meaning. The present paper deals with the computation of phonology, and so presentation of the different characteristics of the models will be mainly concerned with this aspect of reading.

Theorists supporting the dual route model claimed that the existence of two independent mechanisms, one based on the whole word, and one based on sub-word units, was required to explain how irregular words and nonwords are pronounced. The application of spelling-sound rules to irregular words will result in an incorrect pronunciation, and so for these words the visual route must be used. Nonwords do not have a lexical entry, and so they are pronounced according to the GPC rules. In most conceptualizations of this model, the direct visual pathway to meaning is assumed to run in parallel with the phonological pathway, leading to a race between them. For regular words the outputs of the two routes produces a consistent pronunciation, while for irregular words the two pathways lead to different outputs, yielding interference. This fact explains why regular words are named faster than irregular words (Baron & Strawson, 1976).

It has been shown, however, that the effect of regularity interacts with the word's familiarity. The regularity effect is mainly, or only, present, with low frequency words, for which the lexical routine, although it must be waited for because it guarantees the correctness of the derived pronunciation, is slower with respect to the sublexical routine (Seidenberg, 1985a; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Taraban & McClelland, 1987). Following this result, showing the selective influence of one variable, frequency, on one of the two routines, in the models proposed afterwards the existence of multiple pathways was generally accepted, but the independence of the routines was challenged. Thus, in the time course model proposed by Seidenberg (1985 a; Seidenberg et al., 1984) orthographic and phonological information is automatically activated, and whether the effects of phonology become apparent depends on the relative time course of the codes' activation. More recently, a connectionist model has been presented, where a single process is responsible for the compilation of the pronunciation for both regular and irregular words, and for nonwords (Seidenberg & McClelland, 1989; but see Besner, Twilley, McCann, & Seergobin, 1990)

In the dual-route model, nonlexical phonology is assumed to be under the reader's strategic control. This would explain why phonological effects are smaller in lexical decision than in naming (Balota & Chumbley, 1984) as well as why they are influenced by the composition of the experimental list (Carr, Davidson, & Hawkins, 1978; Coltheart, Besner, Jonasson, & Davelaar, 1979; Henderson, 1982; Seidenberg et al. 1984; Shulman, Hornak, & Sanders, 1978). This view has however been challenged by Seidenberg (1985 a, b) who claimed that the initial processes used in decoding words are completely automatic, and so cannot be under the reader's control. The clarification of this issue is likely to have theoretical consequences for models of reading performance (Henderson, 1985; Seidenberg, 1985b).

It is generally accepted that the phonological code produced by the lexical routine, based on learned associations between a visual form and its phonological correspondent, and on extensive practice with the specific forms of words, is generated automatically, if some widely accepted criteria for defining automaticity is considered (Cohen, Dunbar, & McClelland, 1990; LaBerge & Samuels, 1974; MacLeod, 1991; Posner & Snyder, 1975). There is less agreement however on whether the nonlexical process is automatic too. While it is true that phonological codes are also formed automatically at the sublexical level, because they are based on learned spelling-sound associations, to what extent the process of assembling the phoneme sequences can be considered automatic is less clear.

Paap and Noel (1991) argued that the process of assembling the word's pronunciation requires attentional resources, and should therefore be sensitive to processing load. Indeed, when their subjects were asked to read regular and irregular words of high and low frequency while rehearsing a number of digits to be recalled later on, an unusual pattern of interactions between regularity and frequency was obtained. The most striking results were that low frequency irregular words were named faster under high (five digits) than under low (1 digit) load, and that a substantial frequency effect appeared with regular words. These results are consistent with the interpretation proposed by Paap and

Noel that the output of the sublexical routine is slowed down as a consequence of the task requirements of remembering the digits, and therefore the interference caused by the almost simultaneously outputs of the sublexical and the lexical routine is avoided. Thus they were interpreted by the authors as supporting the idea that the assembly of phonology is effortful and requires resources.

The question of the extent to which the two routines are effortful rather than automatic, and therefore open to strategic manipulations can be posed differently. If the readers know in advance some aspect of the phonological representation of the words they have to name aloud, can they strategically exploit this knowledge, to optimize performance? For instance, is it possible for a reader to exploit the information that all stimuli to be presented in a list are irregular words, or, alternatively, that the list consists only of nonwords (see for instance Carr et al., 1978)? If the lexical and sublexical routines yield independent outputs, subjects might be able to attend to one or another, depending on the experimental conditions.

The idea of flexible strategies has been recently considered also by Monsell, Patterson, Hughes, and Milroy (1991). According to the authors, the ability to use specific strategies exploiting advance information about the stimuli to be presented can indicate whether the lexical and the sublexical routines are functionally distinct. If a specific strategy can be used only on one of the two routines, then a pattern of dissociation should emerge, with only one or the other of the routines to be influenced by a specific variable. In Monsell et al.'s study (see also Frederiksen & Kroll, 1976; Paap & Noel, 1991) subjects had to read blocks of either irregular words, or irregular words mixed with nonwords. Irregular words can only be pronounced through the use of the lexical routine, on the basis of the phonological representation stored in memory, whereas for nonwords there is no stored representation, and their pronunciation is given by the most frequent (regular) spelling to sound correspondence. If it is known that all stimuli in a list will be irregular, then the output of the sublexical routine might be momentarily ignored in order to give the lexical process a headstart, thus avoiding interference. On the other hand, if only nonwords are used in a list, then there would be no point in trying to find a lexical representation to match them. If subjects are able to dissociate the two routines, then it should be faster to read nonwords and irregular words when they are presented by themselves, than when they are mixed together. Moreover, subjects should tend to make more regularization errors when the words are mixed with nonwords, than when they are presented separately. Indeed, this is what Monsell et al. found.

In this chapter the influence of strategies on the different processes involved in the computation of phonology is studied. This paper will focus on whether and how strategies can have selective influence with respect to the operation of stress assignment. First, a brief outline of a possible model of the representation of stress, and of the processes involved in stress assignment in reading will be exposed, based on data from Italian (Colombo, in press) and English (Monsell, Doyle, & Haggard, 1989). Following this, two experiments will be presented, in which the extent to which strategies of stress assignment can be selectively used in different experimental contexts is investigated.

The representation of stress

Colombo has recently suggested that the information relevant for stress assignment may be of different types, and claimed that a description of the operations involved in stress assignment must be strictly tied to the specific process (lexical or sublexical)[1] involved in reading (Colombo, 1991; in press). According to Colombo lexical stress is coded either suprasegmentally or segmentally (for a discussion of these two types of representation see also Black & Byng, 1985). The suprasegmental representation captures the rhythmic structure of a word as a whole, is not intrinsically represented in the single phonemes of a word, but can be superimposed on the sequence of phonemes correspondent to a letter string.

This information is part of the implicit knowledge about stress encoded in one's general knowledge about the words of a given language. Italian speakers know, for instance, that in Italian the dominant stress in multisyllabic words is on the penultimate syllable. Words with stress on the antepenultimate syllable (for words with three syllables or more) are a minority. This implicit knowledge should bias the assignment of stress in word pronunciation, in the sense that the most common stress pattern becomes "regular." It could therefore be expected that regularly stressed words are pronounced faster than irregularly stressed words, that is, than words with stress on the antepenultimate syllable. This phenomenon has been called the stress effect (Colombo, in press; Monsell et al., 1989). Moreover, if a word's pronunciation is not known it will be pronounced with the most frequent stress pattern (Colombo, in press, Experiment 5). The type of representation of stress just described will be labelled the *general bias* of a specific language.

Unlike the suprasegmental representation, the segmental representation captures those properties of stress intrinsic to the phonemes. For instance, in Italian the phonemic representation of a vowel may specify the fact that it is a carrier of stress by its duration (Bertinetto, 1980). Segmental information is of two sorts. At one level, the stored knowledge about a word will include its stressed vowel. When phonology is retrieved directly from the lexicon, information about stress is completely specified in the segmental properties of the word. In fact *lexical* knowledge is the only reliable source for stress assignment, for languages like Italian and English where stress is not predictable.

In addition to being represented in the lexicon, segmental information about stress is also part of the phonological correspondences of subword spelling patterns that are common to a group of words in a language ("neighborhood," Glushko, 1979; Brown, 1987). Particularly relevant for stress assignment is the similarity among words that share the final letter cluster, formed by the nucleus of the penultimate syllable and the last syllable (e.g., "-E-RO" in "int'ERO"). These words identify a neighborhood and may influence the assignment of stress. For instance, in the neighborhood of words sharing the "-ero" cluster the vowel nucleus of this segment (-e-) will be pronounced with longer duration for words which have a dominant stress (like "int'ero"), with respect to those which have an irregular stress (like "l'ibero"). The importance of the final cluster for stress assignment has been shown also for English on both linguistic (Chomsky & Halle, 1968) and psychological grounds (Baker & Smith, 1981).

The type of knowledge provided by similar words, the *neighborhood* information, is represented in the strength of the learned phonological correspondences of the endings of words, which depends on the proportion of words in which they are present and may

[1] In the remaining part of the paper the terms lexical or word-specific process and sublexical or assembly process will be used to denote, respectively, the mechanisms responsible for retrieving the stored word's pronunciation in the lexicon, and compiling a phonological representation of a word based on the spelling-sound correspondences of different size.

influence the assignment of stress. Colombo (in press, experiments 4 and 5) has found that the characteristics of the neighborhood can both determine whether the stress effect, that is, the difference in naming time between regularly and irregularly stressed words, is found, and the probability that a nonword is assigned an irregular stress pattern. When words are irregularly stressed, the cost of irregularity (the stress effect) only appears if the words have inconsistent neighbors.

According to Colombo, the type of information on which people rely to assign stress depends on the reading process that is involved in pronouncing the words. When the lexical process is used, and the pronunciation is simply retrieved from the lexicon, stress is only represented segmentally. In contrast, when the assembly process is used, stress assignment derives from the interactions of two sources of information. One, the *general bias* of the language, is derived from one's knowledge of the stress type distribution in the language, and is abstractly represented by a rhythmic structure superimposed on a word.

The second source of information about stress that can be exploited by the sublexical routine comes from the neighborhood of the to be pronounced word, which provides segmental information. This type of representation is also probabilistically determined, reflecting the distribution of orthographic patterns and of their phonological correspondences, on the basis of which a neighborhood is defined. Both sources of knowledge are pooled during the assembly process, and can determine the probability that a word is stressed according to the most common regular pattern.

Stress and strategies

The above characterization of stress assignment and of its relation to the mechanisms involved in pronunciation may come useful, if it is correct, to clarify whether the use of a specific reading process is under the reader's strategic control. If the information used to assign stress depends on what mechanism, lexical or sublexical, drives pronunciation, it should be possible to selectively influence one or the other, depending on the stress characteristics of the experimental list. Naturally this hypothesis is based on the assumption that the different types of information come from processes that are to some extent independent. If this is not the case, then to explain the effect of a selective influence of one variable on a specific process would be more difficult.

In Italian, which is a language with a shallow orthography, the only source of irregularity is dependent on lexical stress assignment. In normal reading conditions, the word-specific process must be involved in phonology retrieval for correct stress assignment, and so it is difficult to distinguish the extent to which each type of representation of stress is involved. It should however be possible to create conditions in which only one or the other are used. We have therefore manipulated the list structure so that the stress pattern of the words in it included was completely consistent, and therefore predictable.

The effect of blocking were tested with respect to two phenomena: the semantic priming effect (Meyer & Schvaneveldt, 1971; Neely, 1977, 1991) and the stress effect (Colombo, in press; Monsell et al., 1989). Both effects require the involvement of the lexical process to become apparent. Their nullification in the blocked condition should therefore imply that the pronunciation of the words in the naming task has been accomplished by assembling the pronunciation through the sublexical process, rather than by directly retrieving it from lexical memory.

The semantic priming experiment

Typically, semantic priming is taken to reflect the fact that activation spreads in the lexicon (Collins & Loftus, 1975), so that the prior activation of the prime sends activation to other words related to it, facilitating their recognition when they are presented shortly afterwards.

Clearly, in a semantic priming paradigm word pronunciation can be affected by the prior occurrence of a related word only insofar as the task is performed on the basis of lexical knowledge rather than simply assembling the word's phonology with no lexical involvement. This fact presumably explains why the sensitivity of naming to semantic priming is not as large as in a task which is more likely to require identification of the whole word, like lexical decision. Nonetheless, in English as well as other languages with a deep orthography the priming effect in a naming task, although small, is consistently found (Becker & Killion, 1977; Forster, 1981; Frost, Katz, & Bentin, 1987; Lupker, 1984; Meyer, Schvaneveldt, & Ruddy, 1975; Neely, 1991; Seidenberg, Waters, Sanders, & Langer, 1984).

In languages with a shallower orthography, however, the effect of semantic priming in word pronunciation is less consistent, and while observed in some studies (for Serbo-Croatian, Carello, Lukatela, & Turvey, 1988; Lukatela, Feldman, Turvey, Carello, & Katz, 1989; for Dutch, de Groot, 1985) it has failed to be observed in others (Frost et al., 1987; Katz & Feldman, 1983).

One way of reconciling these seemingly contradictory results is to assume that whether or not semantic priming effects are observed in word pronunciation depends, at least in languages with a shallow orthography, upon the strategies adopted in performing the task. Support for this hypothesis comes from Tabossi and Laghi (in press). They looked at semantic priming effects in English and Italian, and found that while in English semantic priming effects were always obtained, in Italian their occurrence depended upon the characteristics of the experimental lists. When the list included only words or included nonwords together with words with different stress patterns, for which lexical knowledge was required, semantic priming effects were observed. In contrast, when only regularly stressed words and nonwords occurred in the list, no priming was obtained.

Thus, although it can be assumed that, regardless of the spelling-to-sound characteristics of a particular language, most common words are read on the basis of a visual code, nevertheless under special circumstances a non-lexical strategy of reading may be adopted. In these cases, word pronunciation is performed on the basis of assembled phonology, assuming that stress is assigned by determining the sequence of phonemes forming a partial specification of the phonological representation of the word, dividing the sequence of phonemes into syllables, and then superimposing the prosodic structure on it. Whereas the most frequent stress pattern in a language, that is, the *general bias*, provides the strongest bias, the requirements of the experimental manipulation should be able to shift the bias toward the less frequent stress pattern. This mechanism presupposes no lexical involvement, and hence no priming effects are expected.

In order to test this hypothesis trisyllabic Italian words stressed on the antepenultimate syllable were selected for use as targets. (For the sake of clarity, these types of words will be labeled "irregular," to contrast with the "regular" ones, stressed on the penultimate syllable). Each of these words was presented to the subjects shortly after a prime, which was also a trisyllabic word with the same stress pattern. The prime was either semantically related (e.g., PULPITO, "pulpit") or unrelated (e.g., TRAGICO, "tragic") to the target (e.g., PREDICA, "sermon"). The subjects' task was to read the prime silently and then to name the target aloud as quickly as possible.

Two groups of subjects were tested. One group (Blocked Group) received the test materials in a list where all words—either primes or targets—were trisyllables stressed on the antepenultimate syllable. Moreover, trisyllabic nonwords were added to the list, constructed in such a way as to bias the pronunciation toward the irregular stress pattern. This was accomplished by constructing nonwords so that the last two syllables were letter clusters that are mostly present in irregularly stressed words, as suggested by Colombo (in press).[2] Thus, the stress pattern in the experimental list was entirely consistent. In addition, the experimental list was preceded by a group of practice letter strings, which had the same stress pattern as the experimental items and had the purpose of creating specific expectations about the stress type in the list.

In the second group (Mixed Group) test pairs were the same as in the Blocked Group. Word/pseudowords and word/word filler pairs were also the same, except that 15% of primes, 15% of the nonword targets and 15% of the filler targets were replaced by trisyllabic items whose stress was on the penultimate rather than on the antepenultimate syllable. The same structure of the experimental trials was also used in the practice trials.

The aim of the experiment was to see whether subjects could be induced to use the sublexical, rather than the lexical routine, assigning stress in agreement with the list bias. In fact, the composition of the list was such as to obviate the need to use lexical information for correctly assigning stress, which was always on the first syllable of each item. Moreover, no reliable information about stress, in terms of cue validity, could be derived from the neighborhood, because only about half of the irregular words had a large consistent neighborhood, while the remaining words either had many neighbors with regular stress pattern, or shared the final cluster with a mixed stress neighborhood. If in the experimental conditions here created pronunciation is driven sublexically, then no semantic priming effect ought to be observed in the blocked group. In contrast, semantic priming should be found in the mixed group, where the occurrence of few regular words renders the assembling strategy unreliable, giving place to an interaction between the two groups.

Method

Subjects. A total of 36 Italian native speakers, students from the University of Bologna, volunteered for the experiment, which lasted about 15 min. None of the subjects had previously participated in an experiment of this sort.

Materials. Materials for the blocked group were constructed as follows. A total of 24 test pairs was selected. They had the following characteristics. All primes and targets were trisyllabic Italian words with stress on the antepenultimate syllable, of medium to high frequency of occurrence (mean frequency of the target words=99, from a corpus of

[2]Nonwords were made up in the following way (Colombo, in press). The final letter clusters of regular and irregular words, with a high number of consistent or of inconsistent neighbours (many friends or many enemies) were excised from the words. The nonwords were created by preceding each cluster with a nonsense radical, so as to form a pronounceable letter string that did not specifically resemble any real word. For instance, from the word **"grafici,"** irregular and with many friends, was derived the nonword **"boltici"** sharing the cluster **"-ici"** which is present in many words irregularly stressed. It was found that this type of nonword was more likely to be pronounced irregulalry, with respect to a nonword like **"rilota,"** formed by a cluster **"-ota"** which is present in a large neighbourhood of regularly stressed words.

1,500,000 words, Instituto di Linguistica Computazionale, 1988). Primes and targets in each pair were associated or semantically related (see Appendix). These materials were then re-paired to form 24 unrelated pairs.

In addition, 36 filler pairs were created. In each of these pairs, prime and target were trisyllabic Italian words with the same stress pattern as in the test pairs, and matched in frequency to the latters. In none of the filler pairs, however, was there a semantic relation between primes and targets.

Finally, 60 word-pseudoword pairs were included. Again, primes were trisyllabic words with irregular stress pattern. Pseudowords, which were also trisyllabic, were pronounceable strings. Because it was critical that the bias created by the consistent stress pattern throughout the list would be contrived by pseudowords as well, these were constructed following precise criteria. Namely, the nonwords which were used in Colombo's (in press) Experiment 5, and yielded the higher probability of being pronounced with the less frequent stress pattern were selected.

The assignment of stress for these pseudowords was controlled by a pre-test. Ten judges were asked to read aloud a list of 200 items. The list included bisyllabic as well as trisyllabic words, with both types of stress patterns. In addition, it contained bisyllabic and trisyllabic pseudowords. All the pseudowords in the experiment were selected among those that were stressed by at least nine of the ten judges on the antepenultimate syllable.

Two counterbalanced lists were constructed. Each list contained half of the related and half of the unrelated pairs, in addition to all the filler and pseudoword pairs. Within each list, that included a total of 120 trials, each item occurred only once. The two lists were divided in 3 blocks of 40 trials each. Within each block, there was an equal number of related pairs (4), unrelated pairs (4), filler pairs (12), and nonword pairs (20).

The materials for the mixed group were exactly the same as for the blocked group, with two exceptions. In the filler trials, nine primes and nine targets were replaced by trisyllabic words of comparable frequency, but with regular stress. Likewise, in the word-pseudoword pairs, nine primes and nine targets were replaced by trisyllabic regularly stressed items. The pseudowords were chosen among those that in the pre-test described above received stress on the penultimate syllable by at least nine of the ten judges. Thus, in the mixed group, 15% of all the primes, and 15% of all the target words and pseudowords were regularly stressed.

Practice trials consisted of a list of 20 trials, half of which were formed by semantically unrelated word-word pairs, whereas the remaining half were formed by word-pseudoword pairs. For the blocked group, all the items were irregularly stressed. For the mixed group three primes, and one pseudoword and two word targets of the whole list were regularly stressed.

Procedure and design

The stimuli were displayed on the monitor of an Apple IIe microcomputer. Each trial was preceded by a tone which was followed, after 0.5 s, by the prime presented for 200 ms. The target appeared on the screen 100 ms after the offset of the prime and stayed on for 1.5 s. Simultaneous with the onset of the target was the starting of a timer, which was either stopped by the subject's response or (if the subject did not respond) reset automatically after 1.5 s. Subjects were instructed to silently read the prime and name aloud only the target letter string.

After the appropriate practice list, each subject received one of the experimental lists. The importance of being fast and accurate was stressed with each subject, who received the three blocks of his/her experimental list in a counterbalanced order. Both practice and experimental trials were presented to the subjects in a randomized order.

An equal number of subjects was randomly assigned to the blocked and mixed Group, and within each group, an equal number of subjects was randomly assigned to one of the experimental lists. There were two factors in the experiment: Group (Blocked vs. Mixed), which was between subjects, and semantic relation (Related vs. Unrelated), which was within subjects.

Results and discussion

Mean reaction times for each condition were calculated. In order to reduce variability, data that were more than 2 SDs from the mean RTs of each subject (4.6% of responses) were excluded from analyses. The mean percentage of errors was 1.27.

Mean reaction times for correct responses and mean percentages of errors in the related and unrelated conditions for the blocked and mixed group are presented in Table 1.

When subjects were treated as the random variable, the overall analysis showed no reliable effect of group: blocked vs. mixed, $F(1,34)=2.24$, n.s., Mse=61279. In contrast, the effect of semantic relation and the interaction between group and semantic relation were both significant: Related vs. Unrelated, $F(1, 34)=6.49, p< .05$; Group × Semantic Relation, $F(1, 34)=5.16$, Mse=291, $p < .05$.

When materials were treated as the random variable, the overall analysis showed a reliable effect of group: Blocked vs. Mixed, $F(1, 23)= 4.37$, Mse=392, $p < .001$. In contrast, the effect of semantic relation was not significant: Related vs. Unrelated, $F(1,23)= 2.33$, n.s., Mse=1433. The interaction between the two main factors was reliable: Group x Semantic Relation, $F(1,23) =7.16$, Mse=638, $p < .005$,.

Since the interaction was the only significant source of variance in the analysis by both subjects and items, separate analyses for the blocked and the mixed group were performed. In the former group, the difference between related and unrelated condition was significant neither by subjects nor by items: $t(17)=.28$, ns.; $t(23)=.23$, n.s., respectively. In the latter group, however, the semantic relation was significant in both analyses: $t(17)=3.48$, p<.001, one-tailed; $t(23)=2.83, p < .001$, one-tailed.

The results showed that subjects were reliably faster at reading related than unrelated words in the mixed group, but not in the blocked group, where no semantic priming effects were observed. Test trials were exactly the same in the two groups, the only difference being that in the former, but not in the latter, the stress pattern in the list was entirely predictable. Thus, the findings support the hypothesis that the structure of an experimental list can determine whether Italian readers pronounce words lexically or nonlexically, as reflected in semantic priming, and stress is a critical factor in producing this effect.

Table 1

Mean reaction times and error percent in Experiment 1 for semantically related and unrelated words in the mixed and blocked groups.

	MIXED	BLOCKED
RELATED	544 (0)	526 (1.85)
UNRELATED	564 (1.85)	527 (1.38)

In fact, when the occurrence of pseudowords in the list requires the use of assembled phonology, whether people use this strategy consistently across word and pseudoword trials, or else limit it to pseudowords, employing a whole-word strategy for lexical items, depends on how confident they are that stress can be assigned correctly to words without relying on lexical knowledge.

These results are consistent with those obtained by Tabossi and Laghi (in press). There, in all the Italian lists words were regularly stressed on the penultimate syllable. This pattern was altered only in one experiment (Experiment 4), where, in order to discourage people from using a nonlexical strategy, the list of materials, in addition to pseudowords, contained a small proportion of trisyllabic words. These words were stressed on the antepenultimate syllable, and required lexical knowledge in order to be pronounced correctly. Indeed, the occurrence of semantic priming effects in that experiment suggested that those few irregular words were sufficient to determine the subjects' reading strategies.

That strategies may affect naming is also indicated by Baluch and Besner (1991), who studied word pronunciation in Persian. In this language the spelling-to-sound correspondence is consistent. Some written words, however, include the specification of their vowels, whereas others do not. Thus, while the former words are phonologically transparent, the latter are opaque. Baluch and Besner showed that, when the experimental list only included word trials, semantic priming effects were detected for both opaque and transparent words. When, however, pseudowords were included in the experimental list, semantic priming effects were found for the opaque but not for the transparent words.

The present results, in addition to being consistent with those obtained by Baluch and Besner (1991) and Tabossi and Laghi (in press), extend those findings in two ways. First, it has often being claimed that strategic effects take time. Hence, a very short SOA between prime and target is likely to minimize strategic effects (Neely, 1977; Posner & Snyder, 1975). The present results suggest that subjects can adopt a reading strategy and use it consistently across trials even at an SOA (300 ms) that is usually considered to be too short to allow strategic effects.

Second, the findings indicate that people are very sensitive to list structure. In fact, when the experimental conditions are such to render the use of the non-lexical strategy entirely safe (blocked group), the ordinary lexical strategy is not attended to, even though the bias created by the blocking manipulation goes against the bias existing in the language. As soon as a small margin of uncertainty is introduced, the lexical process becomes the only reliable source.

Stress and neighborhood effects

The lack of semantic priming in the blocked group in the first study was interpreted to suggest that pronunciation was derived sublexically, rather than lexically. In the following experiment we asked whether the blocking manipulation, which should induce the subjects to use the sublexical process, can reduce the difference in naming time between regularly and irregularly stressed words, consequently reducing the size of the stress effect. According to results reported in some recent papers (Colombo, in press; Monsell et al., 1989) words with regular stress are named faster than words with the less frequent stress, because of the general, implicitly stored, knowledge that most words in the language have a typical stress pattern and because of the bias that this knowledge creates. However, if subjects know in advance that the words they have to pronounce will all be irregularly stressed, they may be able to avoid the cost associated with irregularity. If the output of the assembly process is devoid of stress specification, and the prosodic structure can be superimposed on it, reflecting the specific bias of the experimental condition rather than the general bias of the language (favoring the most frequent stress pattern), then the cost associated with the irregular condition should be reduced or disappear.

A blocking manipulation was already included in Colombo' s (in press) Experiment 1, but it was not successful, and the stress effect appeared as strongly in the blocked condition as in the mixed. In the present experiment some factors were controlled that had not been properly controlled in that experiment. For instance, each block was preceded by practice trials. Moreover, in order to avoid confoundings that might obscure the pattern of results, words in each group were matched for initial phonemes, and were controlled with respect to stress neighborhood.

As suggested above, one type of stress information is specified in the phonological correspondences of subword spelling patterns that are common to a neighborhood. In particular, the final two syllables seem to carry information relevant to stress assignment. In order to see whether the blocking manipulation was successful, there had to be a significant stress effect in the mixed group, and therefore it was necessary to control for the characteristics of the neighborhood of the Italian words used in the experiment. This is because when irregularly stressed words have a large consistent neighborhood the stress effect disappears (Colombo, in press, Experiment 4). Thus, the irregular words used in the second experiment shared a large number of neighbors with regular stress, which biased their pronunciation toward a regular stress pattern. This manipulation also allowed to verify whether the neighborhood information can be disregarded, or must be attended to. Under the conditions of the present experiment this type of information is interfering because it biases the assignment of stress towards a pattern which is not the same as that of the words in the list. Optimal performance would then require that it is not attended to. However, if the phonological codes emerging from the spelling-sound correspondences are formed automatically then the influence of neighborhood would be unavoidable. This follows from a widely accepted criterion used to distinguish controlled from automatic processes, namely that a process is automatic when it is faster with respect to another process and interferes with it (LaBerge & Samuels, 1974; Posner & Snyder, 1975).

In order to maximize the probability that pronunciation would be driven by the sublexical rather than the lexical process low frequency words were used (Seidenberg, 1985 a). If high frequency words had been selected, and the word's pronunciation had been directly retrieved from the lexicon, stress information would have been completely specified, thus making the blocking manipulation useless and unsuccessful.

In the present experiment three groups of subjects were used. All groups were presented with a set of practice trials that were meant to prepare the subjects to the stress pattern of the words they were going to name aloud. The first group was presented with regularly and irregularly stressed words mixed together. For the second group the words were presented in blocks, in which words were consistent through the list for stress and frequency. In the third group of subjects the conditions were the same as in the blocked group, but this time they received instructions that strongly stressed speed. If subjects had to compile the pronunciation in a very short time, they might be more easily induced to take advantage of the blocking manipulation that made stress consistent throughout a block, without bothering to check the correctness of each aspect of the phonological representation. In the blocked groups the experimental conditions were then such as to allow subjects to attend only to the information derived from the list, where the stress pattern was consistent throughout, assuming that they could possibly disregard information coming from the endings of the words, relative to the consistency of the stress neighborhood.

Method

Subjects. Thirty students of the University of Padua served as volunteer subjects. None of them had previously participated in experiments involving stress assignment.

Material and procedure. A group of trisyllabic Italian words was selected with the following characteristics. They were of low frequency (mean frequency=7 for irregular words; for regular words the mean frequency was 14, but when an exceptional high frequency item was removed ("sereno") the mean frequency decreased to 9.6) and were stressed either on the penultimate (regular) or on the antepenultimate (irregular) syllable. Irregular words were chosen among those with a large typically regular neighborhood, as shown in the following example. For instance, if the word "s'igari" is divided into the three syllables "si-ga-ri," the nucleus of the penultimate syllable (-a-) plus the final syllable (-ri) form the unit "-ari." If the number of words ending in "-ari" and stressed regularly is much larger than the number of words stressed irregularly, as "sigari," then the word is regular consistent. Because of these constraints, the number of words used in the experiment was necessarily reduced. 24 words with regular stress were selected, all with consistent neighbors, and 24 words with irregular stress and inconsistent neighbors.

The average length of the words was 6.12 for the regular list, and 6.16 for the irregular list. The words were well matched for initial phonemes (see Appendix).

The two sets of words formed two lists that were presented to the blocked groups. The same words, randomly mixed, were divided into two lists, which were both presented to the mixed group. In addition to the experimental stimuli, a set of words were selected to be used as practice. These were formed by 14 words with regular stress, and 14 words with irregular stress.

Subjects were randomly assigned to three groups. The mixed group was presented with a set of 14 practice trials, with regular and irregular words, then with the first list. Afterwards they were presented with a second set of 14 more practice trials, followed by the second list. In both practice and experimental trials both types of stress pattern were present.

The blocked group was presented with the first set of practice trials, which could be formed of either regular or irregular words, and then the first list of words with stress consistent with practice. Following the first list, the second set of practice and experimental trials was presented. The third speeded blocked group had the same material and presentation type as the second group. The order of the lists was counterbalanced throughout all the groups, so that regular and irregular stressed words were both presented in the first and in the second block an equal number of times.

Subjects were seated in front of a display connected to an Apple IIe microcomputer. Words were presented on the center of the screen, white on black. A voice-key connected to the computer was triggered by the subject's voice and stopped the timer. An auditory cue was presented before each word, followed by an interval of 500 ms. The words remained on the screen until the voice-key was triggered, or an interval of 1100 ms had elapsed. When the word disappeared, the experimenter pressed one button for correct trials, and another for incorrect trials. Immediately afterward the reaction time for the trial was displayed followed by the auditory cue for the next trial. Trials which had been marked as wrong by the experimenter, or were above 1100 ms, were automatically eliminated from RT's analyses and considered as errors.

Table 2

Mean reaction times and error percent (in parentheses) in Experiment 2, for words stressed regularly and irregularly, in the different groups: mixed, blocked, and speeded blocked.

	STRESS	
	REGULAR	IRREGULAR
MIXED	573 (2.5)	599 (12.5)
BLOCKED	526 (3.7)	609 (9.6)
BL/SPEEDED	481 (2)	538 (9)

Subjects were instructed to name words aloud as soon as possible. For the speeded blocked group instructions particularly stressed speed. Thus subjects in the latter group were told that they had to try and be very fast. If they were slower than 600 ms, the following feedback was displayed: "Please, be faster" and they were recommended once again to increase speed. After the experiment, subjects were questioned about how they had tried to accomplish the task, and whether they had followed particular strategies.

Results and discussion

The data consisted of the mean correct reaction times for each group and each condition. 4.0% of the data were eliminated for failures of the microphone to trigger the voice key. Trials were considered errors when they were outliers, or when words were mispronounced ("tortura" rather than "tortora"), or there were hesitations, false starts or wrong stress assignment. Errors in stress were few, so they were not analyzed separately.

An analysis of variance was carried out on the mean correct reaction times, with two factors, Group (mixed, blocked, speeded blocked), between subjects, and Type of stress, within subjects.

In the subjects analysis, the factor Type of Stress was significant, $F(1,26)= 42.4$, Mse= 1046, $p<.001$, and the factor Group was also significant, $F(2,26)= 3.6$, Mse=9097, $p<.05$. Subjects in the mixed group were 18 ms slower than in the blocked group, and those in the latter were 57 ms slower than in the speeded group. The interaction between group and stress type was significant too, $F(2,26)= 3.7$, Mse=1046, $p<.05$.

An analysis of the simple main effects showed that the factor Group was significant for regular words, $F(2,32)= 4.3$, MSe= 5071, $p<.05$. The speed gain from the mixed to the blocked condition was 47 ms for regular words while there was none for irregular (-10 ms), while the gain from blocked to speeded was 45 ms for regular and 71 ms for irregular. Moreover, the same analysis showed that the factor Type of stress was significant for the blocked group, $F(1,26)= 32.6$, Mse=1046, $p<.001$, and also for the speeded group, $F(1,26)=17.6$, Mse=1046, $p<.001$, while it was marginal for the mixed group, $F(1,26)= 3.1$, Mse=1046, $p<.1$, (57 ms, 83 ms, and 26 ms, respectively).

An analysis of the blocked and speeded blocked groups considered separately from the mixed group showed a significant effect of Group, $F(1,18)= 9.3$, Mse= 3712, $p<.01$, and an effect of Type of Stress, $F(1,18)= 32.5$, Mse=1557, $p<.001$, while the interaction was not significant, $F=1.0$. The analysis conducted on the mixed and blocked group showed that the stress effect was once again reliable, $F(1,18)= 22.4$, MSe=1312, $p<.001$, and the interaction was also significant, $F(1,18)= 6.1$, $p<.05$. The stress effect was significant in the blocking condition, $F(1,18)= 26{,}06$, $p<001$, but not in the mixed condition, $F(1,18)=2.53$.

Further analyses were carried out to see whether the order of presentation in the blocked conditions had an effect. For the blocked group the only significant factor was Type of Stress, $F(1,8)= 14.46$, Mse=2364, $p<.01$. For the speeded blocked group, Type of Stress was again significant, $F(1,8)= 43.7$, Mse=375, $p<.001$, and so was the interaction, $F(1,8)= 6.7$, $p<.05$. The effect of presentation order for irregular words, which were 53 ms faster when presented as the first block than as the second block, was reliable, whereas there was only a small effect (9 ms) for regular words presented first.

In the item analysis, the factors Type of Stress and Group were analyzed in a between subjects design. Type of Stress was significant, $F(1,138)= 44.08$, Mse= 2645, $p<.001$, and also Group, $F(2,138)= 34.6$, $p< .001$. The interaction was also significant, $F(2,138)= 3.2$, $p< .05$.

In the analysis of errors the only significant factor was Type of stress, $F(1,27)= 33.6$, Mse=1.4948, $p< .001$.

The results of the second experiment show an effect of both stress type and block manipulation. Words with regular stress were pronounced faster than words with irregular stress. Moreover, subjects in the mixed stress group were slower than subjects in the blocked stress group, and subjects in the speeded blocked condition were faster than those in both the other conditions. Thus far, the results support the predictions. Making the type of stress consistent throughout a list was expected to produce an advantage, because subjects would know in advance what prosodic structure they had to assign to the words within the list. In fact, the blocking manipulation did produce a gain in speed, which was detectable not only in the speeded group but also in the blocking group that had not received special speed instructions. Thus, there was an advantage in knowing the stress type in advance.

The interaction between stress and group was however significant, showing a pattern of data somehow in opposition to what had been predicted. The interaction shows that blocking stress had an effect only on regular words. Knowing in advance the stress pattern of the words to be read did not produce a reduction or nullification of the stress effect, as had been expected. Rather, the stress effect was much larger in the blocked condition than in the mixed condition.

As pointed out above, the irregular words had been selected among those inconsistent with respect to neighbors. Moreover, in the irregular-word condition of the blocked group the bias given by the list was in the opposite direction from that produced by the final part of the words. This was not the case for regular words, for which both the list bias and the neighborhood bias were in the same direction. Thus, it appears that inconsistent information coming from the neighborhood produced interference for irregular words. The output of the assembly process was presumably formed more quickly for the regular words in the blocked group, producing the gain speed. In the irregular-word condition, although presumably the assembly process was involved, as can be inferred from the interference coming from the neighborhood, the lexical output also had to be waited for, because contradictory information emerged from the sublexical process. This explanation predicts exactly that the advantage of the blocked vs. the mixed condition would only be apparent for regular words (47 ms) whereas the irregular words should show no advantage

(-10 ms), which is what was found. In the mixed group stress type was varied, and so the output of the lexical process, though probably slow because the words were low in frequency, was necessarily consulted for both regular and irregular words to provide a correct pronunciation. This explains why the stress effect here was reduced (26 ms). The effect of neighborhood or of the list bias would only influence the assembly process.

A comparison of the blocked conditions shows that both regular and irregular words were read faster in the speeded condition. Moreover, the effect of stress tends to be smaller when subjects are speeded, but is still reliable. The latter finding is consistent with predictions of the time course model (Seidenberg, 1985; Waters & Seidenberg, 1985). According to this model, whether the phonological output is involved in identifying a word, and whether it is driven by the sublexical or by the lexical process, depends on the relative time course of the processes involved. Thus, regularity effects depend on the familiarity of the orthographic pattern that drives the lexical output, and are smaller in faster subjects. Apparently, inducing subjects into a faster mode of processing has similar consequences.

The question of where the speed gain comes from remains, however. If subjects had not waited for the lexical output, a higher proportion of errors would have been expected in the speeded condition than in the other two, but in fact this is not what occurred. Only one subject in the speeded group made a high number of errors (25%), while the average error rate of the remaining subjects was around 5%. Thus, it appears that subjects were able to work very fast without qualitatively altering their performance.

General discussion

The results of the experiments of the present paper, although in part contrary to expectations, provide additional insight concerning how stress is represented and processed in phonological retrieval.

The existence of different types of representation of stress, which was suggested elsewhere (Colombo, in press), is confirmed as far as Italian is concerned. Let us consider each, in turn. First, it was suggested that stress is part of the phonological representation of the word stored in memory. As such, when reading is mainly driven by the lexicon, as is the case in the mixed groups of both Experiments 1 and 2, the cost associated with irregularity is reduced, because stress is already specified in the phonological entries. However, whether the naming process is driven by the lexical or the assembly process seems to be open to strategic adjustments. Thus, in the blocked condition of the semantic priming experiment the output of the lexical process was altogether disregarded.

The existence of a suprasegmental level at which prosody is represented, and which can be imposed on the syllabic structure of a word, seems to be confirmed by the data of the first experiment. When the output of the lexical process can be circumvented, and subjects can rely exclusively on the output of the assembly process, stress can be assigned suprasegmentally, without attending to information from the lexicon, as shown by the lack of a semantic priming effect. This is made possible also by the characteristics of the stimulus set in the semantic priming experiment, where information derived from neighborhood did not consistently bias one particular stress pattern. The data from this experiment also support the idea that this level of representation is somehow abstract, because it is not intrinsically coded in the phonemes of a specific word and can be independently assigned to the phonological representation of a word.

The influence of the suprasegmental representation of stress, however, can only be seen when the different pieces of information coming from the sublexical process are concordant, as was the case in the regular-words condition of Experiment 2, where both list bias and neighborhood bias were congruent. But in the irregular-words condition the presence of a large typically regular neighborhood produced interference, increasing the size of the stress effect. This result confirms the idea that stress is also represented sublexically at the segmental level, and that the operation of stress assignment is also sensitive to this type of information.

This finding has some interesting implications. First of all, it seems that subjects could not avoid attending to the information, coming from the final part of the words, regarding the dominant stress in the neighborhood. As mentioned above, because of the stress type consistency through the list, it might have been more useful for optimal performance not to attend to the information coming from the final cluster. Subjects might, for instance, have started articulation pronouncing the first phonemes, without reading the whole string. The fact that they did not, and could not, do so (in fact a few subjects remarked that this was very difficult to do) seems to suggest that neighborhood information is automatically encoded. This is consistent with the idea that phonological correspondences of sublexical segments of various size are phonological patterns that have been learned and associated with the relative orthographic pattern. Thus, this component of the sublexical process is automatically driven, and that explains why it emerges naturally and must be attended to.

What is perhaps less automatic, or could even be effortful, and thus presumably more sensitive to strategic manipulations, is the component of the sublexical process which is devoted to assembling the emerging phonological outputs and selecting the most appropriate for correct pronunciation. The assumption here is that when phonology is driven by the sublexical process, different phonological patterns emerge, corresponding to parsed orthographic units of different size, and whose connection with the corresponding orthographic unit have different weights (Seidenberg & McClelland, 1989; Shallice & McCarthy, 1985). What the assembly process must do is select the units which are more appropriate in order to form a phonological pattern that resembles most the representation of the word to be pronounced and assemble them into a whole. It is at this point that strategic processes can operate, not eliminating the emerging automatically derived outputs but adjusting the time of articulation depending on the contextual conditions. Blocking the stress type, and including nonwords in a list, makes it more functional not to wait for the lexical output, and articulation can be started on the basis of the sublexical output. When information at the latter level is incongruent, the lexical output must be waited for and consulted. This interpretation appears to be consistent with what has been proposed by Monsell et al. (1991). It is also consistent with the idea that phonological codes are derived automatically, but can be subject to a strategic mechanism when they do not map consistently to one of the activated codes present in memory.

According to traditional conceptualizations of automaticity (La Berge & Samuels, 1974; Posner & Snyder, 1975) a process which has become automatic through practice can produce interference on a less automatic process. Thus, a slower process which is susceptible to interference could be either less well-learned, or requiring resources. Recent conceptualizations of automaticity have particularly underlined the fact that the distinctions between automatic and controlled processes is not a sharp one. Processes are more likely to belong to a continuum of automaticity and can display characteristics of automaticity in different degrees depending on the context. A process **A** which appears automatic if compared to process **B** may appear less automatic when compared to process

C (Cohen et al., 1990; MacLeod, 1991). If, as it appears, the information about stress derived from the neighborhood is automatically encoded, and interferes with the application of stress suprasegmentally, then the latter process should be either less automatic than the former, or controlled.

On this ground, the fact that the operation of suprasegmentally applying a stress pattern on a sequence of phonemes can be interfered with may not necessarily imply that it is a controlled process. One possible interpretation of the susceptibility to interference of this process is related to the processing time course of the different codes used to compile the phonological representation. Cohen et al. (1990) claimed that the relative speed of two processes, which is one of the important parameters for the presence of interference, is determined by their strength of processing, which is then responsible of the relative automaticity of the two processes. This account fits nicely with the present data. When the neighborhood of a word is large, and predominantly representing one type of stress, the strength of processing is very high, because the particular segment common to the neighbors has been encountered very frequently, and consistently mapped to a specific spelling-sound correspondence. Information relative to this segment is therefore quickly encoded. Information about stress type in a list, which would be used to apply stress suprasegmentally, is instead derived during the experimental situation, and presumably not practiced enough, in particular with respect to the irregular stress condition, and so strength of processing is lower. The time to apply stress suprasegmentally is consequently delayed.

Whether these data are univocal evidence in favor of a dual-route model in which the processes involved in pronunciation are independent is probably a still open issue. Single process models, where computations at the lexical and sublexical levels are not differentiated (Seidenberg, 1985a,b; Seidenberg & McClelland, 1989; Shallice & McCarthy, 1985), cannot in the present form explain the differential effects of strategies, but could probably be modified to do so (Monsell, et al., 1991; Paap & Noel, 1991).

Acknowledgment

The research was supported by grants MURST 60%, 1991, to Lucia Colombo and MPI 40%, 1991, to Patrizia Tabossi. We are grateful to Petros Giovanis and Assunta Zito for running the subjects in Experiment 1. We also wish to thank Ram Frost, Steve Lupker and Carlo Umiltà for their helpful comments on a former version of the paper, and Sandro Bettella for his technical assistance and for writing the software for the experiments.

References

Baker, R. G., & Smith, P. T. (1976). A psycholinguistic study of English stress assignment rules. *Language and Speech, 19,* 9-27.

Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17,* 4, 644-652.

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 & Performance, 10,* 340-357.

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception & Performance, 2,* 386-393.

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

Besner, D., Twilley, L., McCann, R. S., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review, 97,* 432-446.

Bertinetto, P. M. (1980). The perception of stress by Italian speakers. *Journal of Phonetics, 8,* 385-395.

Black, M., & Byng, S. (1986). Prosodic constraints in lexical access in reading. *Cognitive Neuropsychology, 3,* 369-409.

Brown, G. D. A., (1987). Resolving inconsistency: A computational model of word naming. *Journal of Language and Memory, 26,* 1-23.

Carello, C., Lukatela, G., & Turvey, M. T. (1988). Rapid naming is affected by association but not by syntax. *Memory & Cognition, 16,* 187-195.

Carr, T. H., Davidson, B. J., & Hawkins, H. L. (1978). Perceptual flexibility in word recognition: Strategies affect orthographic computation but not lexical access. *Journal of Experimental Psychology: Human Perception & Performance, 4(4),* 674-690.

Carr, T. H., & Pollatsek, A. (1985). Recognizing printed words: A look at current models. In D. Besner, T. G. Waller, & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 1-82). New York: Academic Press.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* No. 3, 332-361.

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

Colombo, L. (1991) The role of lexical stress in word recognition and pronunciation. *Psychological Research, 53,* 1, 71-79.

Colombo, L. (in press). Lexical stress effect and its interaction with frequency in word pronunciation. *Journal of Experimental Psychology: Human Perception & Performance.*

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing.* London: Academic Press.

Coltheart, M. (1980). Reading, phonological recoding and deep dyslexia. In M. Coltheart, K. E. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia.* London: Routledge & Kegan Paul.

Coltheart, M., Besner, D., Jonasson, J. T., & Davelaar, E. (1979). Phonological encoding in the lexical decision task. *Quarterly Journal of Experimental Psychology, 31,* 489-507.

de Groot, A. M. B. (1985). Word-context effects in word naming and lexical decision. *Quarterly Journal of Experimental Psychology, 37A,* 281-297.

Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for automatic processing. *Quarterly Journal of Experimental Psychology, 33A,* 465-495.

Forster, K. I., & Chambers, S. M.. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12,* 627-635.

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception & Performance, 2,* 361-379.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographic depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception & Performance, 13,* 1, 104-115.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception & Performance, 5,* 6474-691.

Henderson, L. (1982). *Orthography and word recognition in reading.* London: Academic Press.

Henderson, L. (1985). Issues in the modelling of pronunciation assembly in normal reading. In K. E. Patterson, J. C. Marshall, & M. Coltheart, (Eds.), *Surface dyslexia* (pp. 459-508). London: Erlbaum.

Istituto di Linguistica Computazionale (1988). *Corpus di italiano contemporaneo.* Unpublished manuscript.

Katz, L., & Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory & Cognition, 9,* 157-166.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293-323.

Lukatela, G., Feldman, L. B., Turvey, M., Carello, C., & Katz, L. (1989). Context effects in bi-alphabetical word perception. *Journal of Memory & Language, 28,* 214-236.

Lupker, S. J. (1984). Semantic priming without association: A second look. *Journal of Verbal Learning & Verbal Behavior, 23,* 709-733.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative view. *Psychological Bulletin, 109(2),* 163-203.

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

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Frequency effects in word recognition: Where are they? *Journal of Experimental Psychology: General, 118,* 43-71.

Monsell, S., Patterson, K. E., Graham, A., Hughes, C. H., & Milroy, R. (1991). Lexical and sub-lexical translation of spelling to sound. I. Strategic anticipation of lexical status. Paper presented at the Meeting of the Psychonomic Society, New Orleans, Nov. 1991.

Morton, J., & Patterson, K. (1980). A new attempt at an interpretation or an attempt at a new interpretation. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia.* London: Routledge & Kegan Paul.

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. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264-336). Hillsdale, NJ: Erlbaum.

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*. Hillsdale, NJ: Erlbaum.

Paap, K. R., & Noel, R. W. (1991). Dual route models of print to sound: Still a good horse race. *Psychological Research, 53,* 13-24.

Patterson, K., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. Patterson, J. C. Marshall & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading*. London: Lawrence Erlbaum Associates Ltd.

Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance V.* New York: Academic Press.

Rubenstein, H. H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 10,* 645-657.

Seidenberg, M. S. (1985a). The time-course of phonological code activation in two writing systems. *Cognition, 19,* 1-30.

Seidenberg, M. S. (1985b). The time course of information activation and utilization in visual word recognition. In D. Besner, T. G. Waller, & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice, Vol. 5* (pp. 199-252). Orlando, FL: Academic Press.

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., & McCarthy, R. (1985). Phonological reading: From patterns of impairment to possible procedures. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading.* London: Erlbaum.

Shulman, H. G., Hornak, R., & Sanders, E. (1978). The effects of graphemic, phonetic, and semantic relationships on access to lexical structure. *Memory & Cognition, 6,* 115-123.

Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory & Cognition, 20,* 315-328.

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26,* 608-631.

Van Orden, G. C., Pennington, B. F., & Stone, G. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97(4),* 488-522.

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time course and decision criteria. *Memory & Cognition, 13,* 557-572.

Appendix

List of semantically related pairs of words, with stress on the antepenultimate syllable, used in Experiment 1 (in their original form and in English translation).

Italian	English
ANGELO-DIAVOLO	ANGEL-DEVIL
ATOMO-PICCOLO	ATOM-SMALL
CELIBE-NUBILE	BACHELOR-SINGLE
CLASSICA-MUSICA	CLASSIC-MUSIC
CREDITO-PRESTITO	CREDIT-LOAN
FORMULA-CHIMICA	FORMULA-CHEMICAL
GRACILE-DEBOLE	DELICATE-WEAK
MACCHINE-TRAFFICO	CARS-TRAFFIC
MIGNOLO-POLLICE	LITTLE FINGER-THUMB
MINIMO-MASSIMO	MINIMUM-MAXIMUM
PASCOLO-PECORA	PASTURE-SHEEP
PESSIMO-OTTIMO	WORST-BEST
PULPITO-PREDICA	PULPIT-SERMON
SEGGIOLA-TAVOLO	CHAIR-TABLE
SEMPLICE-FACILE	SIMPLE-EASY
SOLIDO-LIQUIDO	SOLID-LIQUID
SPAZZOLA-PETTINE	BRUSH-COMB
SPIRITO-ANIMA	SPIRIT-SOUL
STERILE-FERTILE	STERILE-FERTILE
SUOCERO-GENERO	FATHER IN LAW-SON IN LAW
TIPICO-SOLITO	TYPICAL-USUAL
TRAGICO-COMICO	TRAGIC-COMIC
UNDICI-NUMERO	ELEVEN-NUMBER
VINCITA-PERDITA	WIN-LOSS

List of words with stress on the penultimate syllable (labelled regular) and on the antepenultimate syllable (labelled irregular) used in Experiment 2.

REGULAR STRESS	IRREGULAR STRESS
ABUSO	ACINI
ALBUME	AFONA
BIDONE	BIFORE
CANDORE	GAMBERI
CANTINE	CAUTI
DORATA	DINAMO
EVASO	INCAVO
ILLUSO	IMPETO
LETALE	LIGURE
MALVAGI	MACINA
NOCIVE	NOMINE
OLIVO	ORBITA
OTTAVA	ORFANE
PULITO	PETALI
REMOTO	RESINE
RICAMO	RICINO
ROSAIO	RECLUTA
SONORA	SATURA
SALAME	SENAPE
SERENO	SIGARI
SPAZIOSA	STIPITE
TELAIO	TIMPANI
TENACI	TORTORA
TORTURA	DALMATA

PART 3

Orthography and Lexical Structure

This Page Intentionally Left Blank

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 17

Morphological Analysis in Word Recognition

Laurie B. Feldman
State University of New York at Albany

Darinka Andjelković
University of Belgrade

One of the most refined techniques for investigating morphological processing in word recognition is the variant of the lexical decision task known as repetition priming (Stanners, Neiser, Hernon & Hall 1979). It provides a primary source of evidence, according to Henderson (1989), of facilitation between words formed from the same morpheme (i.e., morphological relatives). Generally, target (second presentation) decision latencies and error rates are reduced in the context of morphologically-related primes (first presentation). Words related to the target (e.g., HEALS) can be forms that are unaffixed (e.g., HEAL), inflected (e.g., HEALED) or derived (e.g., HEALER) in either the same or different modalities (e.g., print or speech) and they can be separated by as many as fifty intervening items. Effects of morphological relatedness have been observed in the lexical decision task across a variety of languages including Serbo-Croatian (Feldman & Fowler, 1987) and Hebrew (Bentin & Feldman, 1990), English (Fowler, Napps, & Feldman, 1985; Feldman, 1991) and American Sign Language (Hanson & Feldman, 1991; see also Emmorey, 1989).

In this chapter, we review evidence of morphological processing in word recognition. In the first two sections, studies of morphology that employ the repetition priming techniques are described. In section one, morphological and orthographic similarity effects are contrasted because alternative accounts of morphological effects in word recognition often minimize the role of the morpheme unit and focus on orthographic and phonological similarity of morphologically-related words. In section two, morphological effects are differentiated from effects due to semantic association, although both are based on word meaning. The repetition priming procedure is a viable tool for studying how morphological relationships among words are represented in the lexicon; however, a confounding episodic contribution to the pattern of facilitation can also occur (e.g., Feustel, Shriffrin, & Salasoo, 1983). It is important, therefore, to differentiate morphological effects from episodic and other types of facilitation and to provide converging evidence of morphological analysis from other word recognition tasks. Morphological effects have also been observed in sentence verification and oral reading tasks. For example, facilitation due to shared morphemes (and/or shared syntactic structure based on the ordering of subject, object and verb constituents) in prime and target sentences have been obtained. In section three, morphological and syntactic facilitation effects are examined in a sentence comprehension task. Finally, it is useful to ask whether analysis of a word's constituent structure at the level of the morpheme is linked to analysis of that word at other linguistic levels. For example, the association between morphological and phonological processing in beginning readers has been examined by comparing performances on such tasks. In section four, a common underlying skill for morphological and phonological analysis is revealed.

1. Distinguishing between orthographic and morphological effects

The standard morphological formation processes in English typically entail prefixation and suffixation to a base morpheme. As a consequence, forms with a common base morpheme generally share orthographic and phonological as well as morphological structure. For regular forms, there is structural transparency (Henderson, 1989) in that related forms are structured around the same base morpheme (e.g., COMPUTE-COMPUTER). Moreover, to the extent that compositionality is present, related words will commonly also have similar meanings. Covariation of morphological and orthographic structure in related forms invites an orthographic account of the morphological effects observed in tasks such repetition priming.

In the repetition priming task, first (prime) and second (target) presentations are separated by an average of ten and sometimes as many as 50 intervening items. Target latency as a function of type of prime is examined. Changes in spelling or pronunciation tend not to significantly diminish the effect of morphological relatedness in this task so that prime-target pairs such as SLEPT-SLEEP or HEALTH-HEAL produce facilitation equivalent to SLEEP-SLEEP or HEALED-HEAL respectively (Fowler et al., 1985; Feldman & Moskovljević, 1987). Similarly, formal similarity of morphologically-unrelated prime and target (viz., phonologically and orthographically but not morphologically similar pairs such as DIET and DIE) does not result in priming at these long lags (Feldman & Moskovljević, 1987; Hanson & Wilkenfeld, 1985; Napps, 1989; Napps & Fowler, 1987). Despite an absence of effects due exclusively to spelling/pronunciation and orthographic form in the repetition priming literature, there persists a tendency to try to interpret morphological effects as effects of orthographic structure. For example, Seidenberg (1987) following Adams (1981) suggested that patterns of high and low bigram frequency could account for morphological patterning because transitional probabilities of letter sequences that straddle a (syllabic or) morphological boundary tend to be low (bigram troughs) relative to probabilities of sequences internal to a unit (cf. Rapp, 1992).

The similar response patterns for regular and irregular forms described above provide evidence against an orthographic account of morphological facilitation in repetition priming. Other morphological effects inconsistent with an orthographic account are based on (1) manipulations of alphabet and on the (2) the absence of an effect when a target is preceded by an unrelated word with a similar orthographic and morphological structure. These findings will be reviewed in the remainder of section one.

Many readers of Serbo-Croatian are fluent in both the Roman and Cyrillic alphabets. For such readers, this situation has been exploited in the study of morphological processing (Feldman & Moskovljević, 1987; Experiment 1). The rationale was that if the facilitation observed in repetition priming arises in a relatively early stage of processing and represents repeated analysis of the same orthographic pattern, then the facilitation due to morphological relatedness should differ when successive presentations of the target word alternated alphabet versus when they preserved alphabet. In that study (Feldman &

Moskovljević, 1987), lags ranged from 7 to 13 items and alphabet was manipulated between subjects so that a one alphabet (preserved) and a two alphabet (alternated) condition existed. Because it is possible that the time course of activation of visual form varies with lag (Monsell, 1985; Ratcliff, Hockley, & McKoon 1985), a second study was conducted in which alphabet was manipulated within subjects and a more expanded range of lags (3 to 20 intervening items) was included (Feldman, 1992). In both, words and pseudowords were presented twice, with a lag of intervening items. Subjects who were students at the University of Belgrade were instructed to perform a lexical decision to each letter string as it appeared. In the alternated alphabet condition, prime and target were transcribed in different alphabets (e.g., НОГА-NOGOM). In the preserved alphabet condition, prime and target were in the same alphabet (e.g., NOGA-NOGOM). Decision latencies to targets that were preceded by primes where target and prime either alternated or preserved alphabet were compared in an attempt to find evidence for facilitation based on repetitions of specific visual patterns.

Results of the two alphabet alternation experiments are summarized in Table 1. In neither experiment was the effect of alphabet (preserved/alternated) significant. Moreover, in the latter experiments, it was the case that for words, neither the effect of lag nor the interaction of alphabet by lag was significant. Stated generally, significant target facilitation occurred when primes appeared in either the same alphabet or in a different alphabet from the target and target facilitation was no greater in the alphabet preserved condition than in the alternated condition. Obviously words presented and represented in the same alphabet are more visually similar than are the Roman and Cyrillic alternatives of a word. Yet, in the repetition priming task where several items intervened between first and second presentations, no significant increment to facilitation was observed on the alphabet preserved trials relative to the alphabet alternating trials. This finding was observed at lags as short as 3 and as long as 20 items.

A second strategy for differentiating morphological and orthographic effects entailed comparing morphologically-related primes to unrelated primes that share orthographic structure. In Serbo-Croatian, it is possible to identify pairs of unrelated words that are formed around homographic base morphemes. For example, the word "BOR" in nominative singular, meaning "pine," is masculine in gender while the word "BORA," meaning wrinkle, in nominative singular is feminine. They have homographic base morphemes spelled BOR but, because of gender differences, require different sets of inflectional affixes. In a repetition priming task (Feldman & Andjelković, 1991; Experiment 3), targets (e.g., BOROVI from BOR) were preceded an average of ten items earlier in the list by an identical repetition (e.g., BOROVI), by a morphologically-related form (e.g., BOR) or by a morphologically-unrelated but homographic form (e.g., BORAMA from BORA). An analysis of variance revealed a significant effect of prime type in both reaction time and errors. Results are summarized in Table 2. Target decision latencies were 589 ms in the identity condition, 617 ms in the morphological condition and 656 ms in the orthographic condition. Decision latencies in the no prime condition were 661 ms. Target error rates were 3.3% in the identity, 4.4% in the morphological and 12.6% in the orthographic condition. Error rates in the no prime condition were 16.7%. The effect of prime type was significant with both subjects (*F1*) and item (*F2*) as random effect variables, with both reaction time and errors as dependent measures. Post-hoc tests indicated that the morphological and orthographic prime conditions were significantly different.

Table 1

Mean decision latencies (ms) and errors for words in the alphabet preserved and alphabet alternated conditions of the repetition priming task.

	First Presentation	Lag	Repetition: Alphabet Alternated	Repetition: Alphabet Preserved	Difference
(Feldman and Moskovljević, 1987; Experiment 1.)					
	642	10		552	90
	678	10	588		90
(Feldman, 1992; Experiment 1a)					
words	651	10		601	50
	12.7		6.6		6.1
				592	59
				7.3	5.4
		20	607		44
			4.5		8.2
				595	56
				7.3	5.4
(Feldman, 1992; Experiment 1b)					
words	628	3	562		66
	10.8		8.3		2.5
				562	66
				5.9	4.9
		10	567		61
			7.9		2.9
				573	55
				7.9	2.9

The orthographic and no prime conditions did not differ with either the reaction time or the error measure. That is, no facilitation with orthographically similar but morphologically unrelated words was observed when an average of ten items intervened between first and second presentations in a repetition priming task. These results are consistent with an earlier experiment conducted with Serbo-Croatian materials (Feldman & Moskovljević, 1987; Experiment 2) in which decision latencies for target words such as STAN meaning "APARTMENT" were reduced by the prior presentation of a derivationally-related prime such as the word STANČIĆ meaning "LITTLE APARTMENT." By contrast, the word STANICA did not reduce target latencies. Note that this word is morphologically unrelated but orthographically similar to the target, is composed of one morpheme and means "BUS STATION." In contrast to Feldman and Moskovljević (1987), in the present study, orthographic primes, as morphological primes, were morphologically complex forms consisting of a base morpheme and an inflectional affix. Nevertheless, orthographic primes were equivalent to the no prime condition. Collectively, these studies refute the hypothesis that orthographic similarity underlies morphonological facilitation in repetition priming.

Table 2

Mean reaction times and errors for targets (e.g., BOROVI) following identity, morphological and orthographic primes and for first presentations of the target in repetition priming.

Prime Type	Example	Latency	Errors
Identity	borovi	589	3.3
Morphological	bor	617	4.4
Orthographic	borama	656	16.7
First Presentation		661	16.7

In summary, relative to the no prime condition, both morphological relatives and identical repetitions facilitated target recognition. The orthographic prime condition was not significantly different from the no prime condition. Finally, and most important to the present discussion of morphological effects in repetition priming, target latencies and errors following morphological primes and orthographic primes at long lags were significantly different both in the analysis by subjects and in the analysis by items.

Orthographic and morphological primes also differentially influenced target latency when presented immediately in succession. In a traditional immediate priming task, morphological primes facilitate and orthographic primes may inhibit. However, the orthographic effect is sensitive to the density of the orthographic neighborhood of the prime (Forster, Davis, Schoknecht, & Carter, 1987) as well as the relative frequency of prime and target and the presence or absence of a mask (Segui & Grainger, 1990). For example, without a mask, lower frequency orthographic primes tend to inhibit whereas with a mask, it is the higher frequency prime that shows inhibition.

A recent report with French materials is consistent with this characterization of morphological as contrasted with orthographic primes (Grainger, Colé & Segui, 1991). In that study, masked primes consisted of morphological, orthographic or unrelated derivations of the target. Decision latencies for targets were fastest in the morphological condition (619 ms), (numerically but not statistically) slowest in the orthographic condition (653 ms) and intermediate (639 ms) in the unrelated condition. In those materials, however, orthographic primes (e.g., AFFIRMÉ-REFORMÉ) tended to be less similar to the target than were morphological primes (e.g., DEFORMÉ-REFORMÉ) and this might account for the marginally significant inhibition in the orthographic condition. Nevertheless, the critical point is that morphological primes showed facilitation whereas orthographic primes showed weak inhibition, at best, and conservatively, no difference from the unrelated condition.

A study recently completed in Serbo-Croatian replicates the difference between orthographic and morphological primes presented in immediate succession with an unmasked prime (Feldman & Andjelković, 1991). In a series of two experiments, targets consisting of morphologically-complex forms were preceded by either a morphological relative, an unrelated word formed from a homographic base morpheme or an orthographically and morphologically unrelated word. For example, the target BOROVI was preceded by (1) BOR which is inflectionally related, (2) BORI which is not related morphologically although it is orthographically similar because its base morpheme is homographic and by (3) KRV which is unrelated along both morphological and orthographic dimensions. In one experiment (Experiment 2), primes were of higher frequency than targets. In a second (Experiment 3), primes were of lower frequency than targets. In both experiments, primes without masks were presented to university students in Belgrade for 250 ms and followed by a blank for 50 ms after which the target appeared for 1000 ms. Latencies greater than 2SD from the mean were treated as errors. Results are summarized in Tables 3 and 4.

Morphological and unrelated primes differed significantly at short lags in both experiments and this outcome replicates the morphological facilitation observed at longer lags with similar materials. When primes were less frequent than targets, significant orthographic inhibition (*F1* and *F2*) was evident in the error measure but not in the reaction time measure. This finding is consistent with the results of Segui and Grainger (1990) using morphologically simple materials and unmasked primes. When primes were more frequent than targets, orthographic inhibition was not statistically significant although the reaction time pattern was quite similar to the pattern obtained when primes were less frequent that their targets.

Table 3

Mean reaction times (and percent errors) for targets (e.g., borovi) following morphological, orthographic and unrelated primes. Primes were higher in relative frequency than their targets.

Prime Type	Example	Latency	Errors
Morphological	bor	684	24
Orthographic	borovi	754	45
Unrelated	krv	738	38

Table 4

Mean reaction times (and percent errors) for targets (e.g., bori) following morphological, orthographic and unrelated primes. Primes were lower in relative frequency than their targets.

Prime Type	Example	Latency	Errors
Morphological	borovi	687	15
Orthographic	bor	743	46
Unrelated	krv	746	30

In Italian, inhibitory effects between homographic base morphemes (e.g., FINA-FINIRE which mean "thin" in feminine singular and "to finish," respectively) have been reported in both a double lexical decision (viz., are both letter strings words?) and in a lexical decision task where prime and target are presented successively (Laudanna, Badecker, & Caramazza, 1989). Results were interpreted as evidence of inhibitory connections between homographic base morphemes in the lexicon. By this account, a differential effect of the relative frequencies of prime and target is not anticipated although it could be accommodated. More problematic is the failure to observe inhibition among homographic forms concurrent with facilitation among morphologically-related forms at longer intervals between presentations viz., at the lags incorporated into the repetition priming task. If inhibition reflects a principle of lexical organization then a justification for its sensitivity to lag is needed. Regardless of whether orthographic primes slow or impair accuracy to targets relative to unrelated primes and whether homographic base morphemes relative to other types of orthographic controls pose a special problem for the representation of morphology, it is useful to focus on the reliable facilitation obtained when items are morphologically-related as compared to either unrelated or orthographic conditions. This is true both in repetition priming with average lags of ten items and in successive priming with or without a mask.

In summary, morphologically-related words that undergo changes in spelling and/or pronunciation so that the base morpheme is partially obscured produce the same pattern of facilitation as do related forms that are structurally transparent. This finding has been observed in Serbo-Croatian (Feldman & Fowler, 1987) as well as English (Fowler et al., 1985; see also Kelliher & Henderson, 1990; Nagy, Anderson, Schommer, Scott, & Stillman, 1989) and presents a challenge for an orthographic account of facilitation in the repetition priming task. In addition, for morphologically-related prime-target pairs, the effect of presenting repetitions in the same alphabet was no different than the effect of alternating alphabet (Feldman & Moskovljević, 1987; Feldman, 1992). This outcome suggests that the basis of facilitation must be sufficiently abstract to tolerate changes in visual form introduced by manipulations of alphabet. Finally, when homographic base morphemes appeared in prime and target, facilitation was observed among forms that shared a base morpheme but not among unrelated forms. In fact, the contrast between homographic and no prime conditions sometimes revealed marginal inhibition. Evidently, morphological effects cannot be described in terms of shared orthographic structure.

2. Distinguishing between semantic and morphological effects

Related forms, by definition, share a base morpheme. Because morphemes are generally defined as units of meaning, it is plausible that morphological facilitation reflects the semantic similarity of prime and target. Linguists distinguish between two types of morphological relatives, inflections and derivations, and these forms differ with respect to the productivity of rules and the predictability of their meaning from a semantic analysis of the base and its constituents (Aronoff, 1976). Whereas inflections rarely produce new shades of meaning, derivations are much less constrained semantically and historically often change meaning once formed (e.g., TERRIFIC and TERROR). For example (from Henderson, 1985), the prefix UN typically modifies the base adjective or verb in a predictable manner (e.g., UNCLEAR, UNDRESS) but some forms are derived from obsolete or rare bases (e.g., UNKEMPT) (Lakoff, 1971; Jackendoff, 1975). Moreover, forms such as UN+base+ABLE are semantically ambiguous insofar as the prefix can modify either the base verb (V) or the (V+ABLE) adjective (e.g., UNDOABLE may mean UNDO+ABLE or UN+DOABLE). Inconsistencies of semantic composition are obvious in semantic comparisons of base-derivation pairs such as DISCUSS-DISCUSSION, CONGREGATE-CONGREGATION and PROFESS-PROFESSION and point to the unpredictability inherent in a semantic analysis of some complex forms from their constituents.

Nevertheless, morphologically-related words tend to have similar meanings. Moreover, because semantic facilitation has been so thoroughly studied (see Neely, 1991, for a review), it is important to contrast facilitation due to morphological relatedness and to other types of semantic relatedness. Semantic contributions to the pattern of facilitation in repetition priming were explored with derivational relatives in English and in Hebrew. In a repetition priming study with English materials (Feldman, 1991), semantic overlap was first assessed by a separate group of subjects who scaled the items for semantic distance. (Due to the constraints of English, word class changes and other aspects of semantic predictability were not well controlled.) For each target, a morphological relative close and remote in meaning was selected. Small but statistically equivalent effects of the prior presentation of the same morpheme in a related word were observed for derivational relatives that were semantically close (36 ms) and semantically remote (30 ms) whereas identical presentations produced robust facilitation (93 ms). Evidently, extent of semantic overlap did not influence the pattern of facilitation in repetition priming.

In a second study (by C.A. Fowler, reported in Feldman, 1991), a target (e.g., HOT) was preceded at least ten items earlier in the list by the same item (e.g., HOT) or by a strong antonym (e.g., COLD). No facilitation was observed in the antonym condition relative to the identity and no prime (i.e., initial) conditions. In the sense that these items were highly predictable and semantically constrained, it is difficult to argue that semantic similarity underlies facilitation in repetition priming.

A third study (Bentin & Feldman, 1990) compared facilitation by associative and morphological primes at long and at short lags. Materials were Hebrew words; prime conditions consisted of morphological, semantic or both morphological and semantic relatives of the target. For example, the word meaning "LIBRARY" was preceded by the word for "NUMBER" which is formed from the same root or base morpheme, by the word for "LIBRARIAN" which is also formed from the same root and by the word for "READING" which is semantically but not morphologically related. Magnitude of morphological facilitation did not change over lags whereas that for semantic facilitation did. Moreover, when the prime immediately preceded the target, semantic and semantic-plus-morphological primes showed greater facilitation than did morphological primes but when an average of ten items intervened, morphological and semantic-plus-morphological showed equivalent facilitation. Evidently the patterns of associative and morphological facilitation are differentially affected by lag (see Table 5).

In conclusion, similarity of meaning between morphologically-related prime and target does not affect the pattern of facilitation in repetition priming. Morphological relatives closely related and remotely related semantically both produced the same pattern of facilitation at long lags. Moreover, closely related antonym pairs produced no effect under conditions where morphological relatedness effects were observed.

Table 5

Mean reaction times and errors for targets (e.g., ספריה meaning "library") following morphological, semantic plus morphological, semantic and no prime in repetition priming (Bentin & Feldman, 1990).

			Lag			
			lag 0		lag 15	
Prime Type	Example	Meaning	RT	Errors	RT	Errors
Morphological	מספר	number	589	1.8	587	1.0
Semantic/Morphological	ספרן	librarian	559	1.0	583	2.6
Semantic	קריאה	reading	563	2.1	611	1.1
Filler			606	1.2	606	1.2

Patterns of morphological facilitation in this task are, therefore, not easily described in terms of semantic similarity. In addition, morphological effects occurred at separations between prime and target that considerably exceed those at which semantic/associative priming has been demonstrated (Bentin & Feldman, 1990; see also Dannenbring & Briand, 1982; Emmorey, 1989; Henderson, Wallis & Knight, 1984; Napps, 1989). Evidently, morphological and semantic facilitation reflect different underlying mechanisms.

3. Morphological effects in sentence contexts

The study of word recognition is sometimes represented as the interface between the domains of perception and language processing. It is the case, however, that the status of linguistic codes in word recognition is problematic (Henderson, 1989). In the first two sections of this chapter, we showed that morphological effects could be experimentally differentiated from effects of (1) shared orthographic and (2) shared semantic structure. In the next two sections, the relation between morphology and other types of linguistic patterning are examined. In section three, patterns of facilitation due to morphological and syntactic similarity between prime and target sentences are explored. In section four, the association between morphological and phonological analysis is examined in beginning readers.

Effects of morphological relatedness have been observed in sentence contexts as well as in isolated words. In one study (Feldman & Andjelković, 1990), students from the University of Belgrade were presented with pairs of sentences that either shared the same syntactic structure (subject verb (SV) or verb object (VO)) and/or shared the same base morphemes. In one experiment, subjects were required to read the sentence aloud and onset to vocalization was measured. A similar task has been reported to show effects of syntactic structure (Bock, 1986; 1990). In two related experiments, subjects were required to judge whether the sentence made sense and latency and errors were measured. For example, subjects saw target sentences such as VODIČI PLIVAJU which means "The guides swim" and has a subject (S) verb (V) structure. Across subjects, that sentence was preceded by four types of prime sentences. These included (1) Morphologically unrelated and structurally dissimilar primes consisting of sentences such as ČITA KNJIGU which means "He reads a book." This sentence has a verb (V) object (O) structure (which is grammatical in Serbo-Croatian because pronouns need not be specified outside the verb). Also included were (2) morphologically unrelated but structurally similar primes consisting of sentences such as ŽENA ČITA which has a SV structure and means "The woman reads." and (3) morphologically related and structurally similar primes consisting of sentences such as VODIČ PLIVA which has a SV structure and means "The guide swims." Finally, there were (4) morphologically related but structurally dissimilar primes consisting of sentences such as VODI PLIVAČA which has a VO structure and means "He guides the swimmer." Primes and targets were constructed so that the same order of the two base morphemes (i.e., VOD- and PLIV-) was maintained over all prime and target sentences in which they appeared.

Foil sentences were morphologically related or unrelated and had either the same or a different constituent structure. Morphologically related foils contained the same base morphemes in illegal combinations such as verbal affixes on nouns and nominal affixes on verbs. Morphologically unrelated foils were composed of legal morphological combinations but were semantically anomalous. The primes for foil sentences were always semantically and syntactically acceptable.

In the oral reading task, the prime and target members of a pair were presented in different alphabets. Primes were always presented in Roman and targets in Cyrillic. In this way, the visual similarity of successive presentations of a morpheme was reduced. Significant effects of morphological relatedness were observed in the oral reading task. Effects of structural similarity were absent (see Table 6). While this outcome can be interpreted as evidence of facilitation due to repetition of base morphemes in prime and target sentences, an alternative account of morphological effects in this task focuses on the repetition of the initial syllable (e.g., VOD-) in all related sentences but not in unrelated sentences. Therefore, it is important to replicate the morphological effect in a task where an advantage based on repetition of the first syllable effect seems unlikely to occur.

In the verification task, subjects had to decide whether each target sentence made sense. Therefore, latencies presumably reflect more than just processing of the first syllable. All prime sentences were meaningful as were half of the target sentences. Primes and targets were printed in the same alphabet. Otherwise, materials were identical to those of the previous experiment. Results indicated that latencies were significantly longer following morphologically unrelated primes than following morphologically related primes (see Table 6). In addition, latencies were longer following structurally dissimilar sentences than following structurally similar sentences in both the morphologically related and the unrelated conditions. The interaction was significant by *F1* but not by *F2*. As in the previous experiment, the effect of morphology could reflect priming of morphemic units over different sentence structures. Alternatively, it could simply reflect episodic repetitions of the same letter sequence (e.g., VOD-) in sentence initial position. The episodic account seems unlikely in a verification task where the entire sentence must be processed before responding. Moreover, it does not explain the significant difference between morphologically related prime sentences with same and different structures (i.e., 901 vs. 975).

Table 6

Oral reading and verification times (errors in parentheses) for target sentences primed by morphologically and structurally similar sentences (From Feldman & Andjelković, 1990).

		Sentence structure	
		Same structure	Different structure
target: VODIČI PLIVAJU (SV)			
Morphology			
	Related	VODIĆ PLIVA (SV)	VODI PLIVAČA (VO)
	Unrelated	ŽENA ČITA (SV)	ČITA KNJIGU (VO)
Task			
Oral reading:			
	Related	676	690
		(5)	(8)
	Unrelated	718	730
		(9)	(9)
Verification:			
	Related	901	975
		(3)	(16)
	Unrelated	973	1002
		(12)	(16)

In addition, when morphemes were not repeated (i.e., for morphologically unrelated sentences), structurally dissimilar primes and structurally similar primes had significantly different effects on target latencies. This outcome indicates that the verification task is sensitive to sentence structure defined over different surface forms. In summary, when the experimental task requires that the entire sentence be processed, facilitation can arise between sentences with similar structures.

An examination of the materials used in the previous two experiments revealed that morphologically related and structurally related primes consisted of words related by inflection to the target sentence whereas morphologically related but structurally dissimilar sentences generally consisted of words related to the target by derivation. In a final experiment in this study, effects of sentence structure and morphology were again investigated in a sentence verification task. In contrast to the previous experiments, here all critical items consisted of morphologically related prime-target sentences. Moreover, in addition to sentence structure, type of morphological relatedness (viz., inflection/derivation) was manipulated. The essence of stimulus construction entailed identifying base morphemes that could function as part of either a noun or a verb. For example, subjects saw target sentences (constructed around the morphemes PLIV- which means "swim" and VOD- which means "guide") such as PLIVAJU VODIČI which means "The guides swim" and has a VS structure. As in the previous experiments, that sentence was preceded by four types of prime sentences across different groups of subjects. All primes were morphologically related by either inflection or derivation to the target. In inflected sentences, the word class of the base morphemes was preserved over prime and target sentences. In derived sentences, the word class of the base morphemes changed in prime and target sentences. In addition, primes and targets varied with respect to similarity of sentence structure. Four combinations of sentence structure and morphology were possible: (1) structurally dissimilar inflectional primes consisting of VO sentences such as PLIVA KA VODIČU which means "He swims toward the guide." (2) structurally dissimilar derivational primes consisting of VO sentences such as VODIŠ PLIVAČA which means "You guide the swimmer." (3) structurally similar inflectional primes consisting of SV sentences such as VODIĆ PLIVA which means "The guide swims" and finally, (4) structurally similar derivational primes consisting of SV sentences such as PLIVAČ VODI which means "The swimmer guides" (see Table 7). Primes and targets were constructed so that the same order of sentence constituents (viz., S,V,O) was preserved throughout a set. The advantage of constructing materials in this way is that repeated base morphemes (i.e., VOD- and PLIV-) do not always appear in the same position in prime and target sentences. For example, both VODIĆ PLIVA and PLIVAČ VODI have subject before verb but the ordering of base morphemes in these sentences differ. The disadvantage is that by preserving the ordering of elements in a pair with similar structure, the effect of changing syntactic role for a particular base morpheme may be lost. Prime and target sentences were printed in different alphabets.

Results indicated that the effect of morphology was significant (by both $F1$ and $F2$) for both the latency and the error measures such that derivations produced less facilitation than did inflections. This finding suggests that the effect of sentence structure observed in the previous verification experiment can be attributed to different effects on targets of prime sentences related by inflection and by derivation. Because inflectionally related primes differed only in number from the target whereas derivationally related sentences transformed the base morpheme of the noun into a verb and the base morpheme of the verb into a noun, it was always the case that inflectional sentences were semantically more similar to the target than were derivational sentences.

Table 7

Verification times (and errors in parentheses) for target sentences primed by morphologically and structurally similar and dissimilar sentences.

	Sentence structure	
target: PLIVAJU VODIČI (VS)	Same structure	Different structure
Morphology:		
inflection	VODIĆ PLIVA (SV)	PLIVA KA VODIČU (VO)
derivation	PLIVAČ VODI (SV)	VODIŠ PLIVAČA (VO)
inflection	958 (6)	949 (10)
derivation	1084 (20)	1100 (21)

The effect of sentence structure was not replicated. The failure to obtain an effect of sentence structure for either inflectionally or derivationally related primes indicates that facilitation based on repetition of the initial syllable is not in itself a plausible account of facilitation. The absence of a sentence structure effect most likely reflects the way in which structure was manipulated in the present experiment. Specifically, the order of (subject, object and verb) constituents ways not always preserved across structurally related primes and their targets.

Morphological effects occur in sentence contexts as well as in isolated words. In the first two experiments in this series, all related targets started with the same initial syllable (base morpheme). Therefore, effects of morphological relatedness could simply be anticipation effects based on repetition of the first syllable. This account is not plausible in the third experiment, however. In fact, in that experiment, same structure and different structure sentences started with different initial syllables (morphemes) and yet there was no effect of sentence structure of the prime (958 ms vs. 949 ms). In sentences related by inflection, the absence of an effect of structure was not anticipated and needs to be investigated further. Specifically, in contrast to English, in Serbo-Croatian it is possible to independently manipulate repetition of sentence constituents (subject, object and verb) and the ordering of those constituents.

Even in contexts and tasks that focus on sentence processing it appears that the morphemic constituents of words are analyzed. That is, morphological analysis is not restricted to isolated words in the word recognition task. In these experiments, it is evident that activation among the morphological constituents of prime and target sentences is not necessarily tied to their syntactic role in a sentence nor to the ordering of morphemes. In the next section, associations between morphological processes and other analytic processes are investigated.

4. Associations between phonological and morphological analysis

The beginning reader provides a window through which to evaluate the relation between morphological and phonological analysis in word recognition. In one study (Feldman, Andjelković, & Fowler, in preparation), children between seven and eight years of age who were native speakers of Serbo-Croatian were administered both a morphological and a phonological task. In the morphological task, children were auditorily presented with a source word and a sentence frame and their task was to complete the sentence by adjusting the morphological affixes on the source word to make it fit semantically and syntactically with the sentence frame. Sentence frames were constructed so that depending on the source word, either an inflectional or a derivational substitution was required and frames were paired so that for one source word both an inflectional and a derivational adjustment were necessary. For example, some children were required to fit the source word KUVAR meaning "a cook" (agent in nominative singular) into the sentence frame MAMA MI POMAŽE DA ____ which means "Mother helps me to ___." This sentence requires the first person singular verb form KUVAM which is related by derivation to the source word KUVAR. For other children, the infinitive KUVATI was presented as the source word for the same sentence frame. Here the source word and the response are related by inflection. Still other groups of children viewed the same source words (i.e., KUVAR, KUVATI) on different sentence frames. For example, OVAJ RESTORAN IMA DOBROG ___ which means "This restaurant has a good___" requires the accusative singular form of the agent KUVARA. This response is inflectionally related to KUVAR and derivationally related to

KUVATI. All subjects were tested on all four combinations of sentence frame and source word and across subjects the same base morpheme appeared in all conditions. This design minimized effects due to sentence frame and to the morphological complexity and the familiarity of the source word as well as the correct response.

Thirty six sentences were constructed. Each contained between four and six words. The target word was always in final position and varied with respect to word class. Sentences were read aloud by the experimenter. The source word was presented both before and after the sentence frame and all source word-sentence frame combinations required at least one morphological substitution. Forty children randomly selected from an urban elementary school in Belgrade, Yugoslavia, were tested individually by a native speaker of Serbo-Croatian and the subjects' responses were transcribed by that experimenter. Results indicated that inflectional responses tended to be correct more frequently than were derivational responses. Mean errors were 1.0 and 3.5 respectively out of a maximum of 18.

Subsequent to the sentence completion task, all subjects participated in a phoneme deletion task (Rosner & Simon, 1971). Subjects heard words and pronounced them aloud without the designated phoneme. All words became orthographically legal but meaningless sequences after phoneme deletion. The position of the deleted letter was balanced across words and, in the source word, it always constituted part of a cluster. Responses were transcribed by the same adult native speaker of Serbo-Croatian. Performance on the phoneme deletion task was correlated with performance on the inflectional and derivational conditions (summed over sentence frame) of the sentence completion task.

Results indicated a significant correlation between phonemic deletion and each morpheme condition $r = .37$ for inflections and $r = .52$ for derivations, respectively. Finally, to each child was administered verbal and nonverbal intelligence tests and these served as covariant controls. Verbal intelligence accounted for 33% of the variance and nonverbal intelligence accounted for 25% of the variance on derivational morphology score. The contribution of intelligence to the inflectional score was not significant but this outcome may reflect the near perfect performance and consequent lack of variability on the inflectional task. Most important, results revealed a significant relationship between phonological and derivational performance even when effects of intelligence were partialled out. That is, with controls for verbal and nonverbal intelligence, performance on a phonological task was still a significant indicator of performance on a (derivational) morphology task. It accounted for a significant 14% of the variance.

Evidently, the ability to explicitly manipulate phonemic segments is associated with the ability to complete sentences with a syntactically correct form. This relationship is independent of intelligence and can be interpreted as evidence of a general linguistic style of analysis that is not tied to particular units. This outcome has also been observed for learners of English which conveys syntactic information through fixed word order in contrast to Serbo-Croatian where word order is relatively free to vary (Fowler, 1988; 1990).

It is often claimed that metalinguistic skill is the single most important factor in learning to read (e.g., Tunmer, 1988). While there is ample compelling evidence that reading an alphabetic orthography requires phoneme awareness, evidence that awareness of linguistic units above the level of the phoneme makes an independent contribution to reading skill is sparse (but see Fowler, 1988; 1990). In the present study, the awareness of morphemes in young readers is associated with phoneme awareness and the relationship cannot be explained by general intelligence or by vocabulary knowledge.

5. Morphological effects reflect linguistic analysis

It is sometimes claimed that three related skills underlie the language user's command of morphology (Tyler & Nagy, 1989). Primary is an appreciation of the *internal structure* of a word such that the presence of a shared component among morphological relatives is recognized, either explicitly or implicitly. Experimental evidence for morphological analysis of a word's structure comes primarily from patterns of facilitation in a priming task in which the same base morpheme is repeated. Recognition by skilled readers is facilitated when the same morphological components recur and the basis of this facilitation can be neither semantic nor orthographic in origin.

Once words can be analyzed with respect to morphological components then it is reasonable to ask whether skilled readers are sensitive to the constraints on *combinations* of morphemes or to the *syntactic implications* of appending particular affixes to a base morpheme. The design of the foils in the sentence verification and oral reading tasks forced adult subjects to attend to these dimensions because some sentences were composed of illegal morphological combinations. Although the foils were not analyzed, morphological analysis was evident in the facilitation to target sentences composed from the same morphological constituents as their prime sentences. Here, effects of visual similarity were eliminated by presenting members of a pair in contrasting alphabets. Interestingly, equivalent facilitation occurred when, across prime and target sentences, a base morpheme changed word class (derivational relatives) and when it did not (inflectional relatives) even though they differed with respect to semantic similarity. It has been reported that skills of morphological analysis emerge before combinatory or syntactic skills (Tyler & Nagy, 1989). In the sentence completion task, however, seven and eight year olds were able to produce the appropriate inflectional and derivation affixes for a variety of syntactic contexts. In order to respond accurately in this task, subjects had to segment the base morpheme from its source word as well as generate a syntactically appropriate affix. Sometimes this entailed forming a verb from a noun or a noun from a verb and it always required the addition of an inflectional affix. Evidently, the metalinguistic demands of this task allowed the children to utilize their knowledge of morphemes and how they combine in particular syntactic contexts.

In conclusion, evidence for morphological analysis in word recognition is not tied exclusively to the repetition priming task although that task has allowed a differentiation between morphological analysis and effects of orthographic or semantic structure. Morphological effects are also evident in a task where constituents of a sentence are experimentally manipulated. Although the mechanism of syntactic effects in the verification task is not clear, it is certain that facilitation occurs when the constituents of words are repeated. This finding is surprising because the task fosters analysis at a level more abstract than the morpheme or the combination of morphemes that comprise the words of a sentence. Evidently, readers cannot refrain from engaging in analysis at the level of the morpheme.

It has recently been demonstrated that time to recognize a target word is influenced by its frequency relative to the other words in its orthographic neighborhood (Grainger, 1990). Similarly for a morphologically simple word, the frequency of words that are inflectionally and derivationally related to it influences the pattern of reaction times in lexical decision (Katz, Rexer, & Lukatela, 1991; Nagy, Anderson, Schommer, Scott, & Stallman, 1989; Taft & Forster, 1975) and it is not necessary that the shared base morpheme of those words be explicitly represented in the surface form (Kelliher & Henderson, 1990). These findings suggest that recognition of any particular word is influenced by properties of other words that are related along some dimension. This similarity is often captured in terms of organizational properties of the lexicon that are distributed rather than tied to one lexical entry. Perhaps the illusion of a shared orthographic or semantic component among morphological relatives has misguided the investigation of morphological processing and undermined our understanding of the status of the morpheme as an abstract linguistic unit.

It has been observed that children who are good at phonological analysis also tend to be good at morphological analysis and that their performance cannot be attributed to either verbal intelligence (and good vocabulary) or to general intelligence. This finding helps to elucidate the value of morphological analysis. It is well established that good and poor beginning readers differ in their ability to grasp the phonological structure of a word in a variety of experimental tasks. Phonological analysis helps the beginning reader map unfamiliar written words into a spoken form which may be familiar even when the written form is not. The essence of this process is an explicit appreciation of the linguistic components of a written word and how they map onto phonemes. It is important because it underlies the ability to read unfamiliar words and combinatorial productivity of the writing system in general. Morphological analysis may serve a similar function. Insofar as words can be constructed from and decomposed into meaningful components and those components can be recombined into new words, morphological analysis enhances the productivity of the reader.

Acknowledgment

The research reported here was conducted at the Laboratory of Experimental Psychology at the University of Belgrade and at Hebrew University in Jerusalem and was supported by funds from National Institute Of Child Health and Development Grant HD-01994 to Haskins Laboratories and by funds from the University of Belgrade.

References

Aronoff, M. (1976). *Word formation in generative grammar.* Cambridge, MA: MIT Press.

Adams, M. J. (1981). What good is orthographic redundancy? In O. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology.* Hillsdale, NJ: Erlbaum.

Bentin, S., & Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology, 42A,* 693-711.

Bock, J. K. (1986). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12(4),* 575-586.

Bock, J. K., & Loebell, H. (1990). Framing sentences. *Cognition, 35,* 1-39.

Dannenbring, G. L., & Briand, K. (1982). Semantic priming and the word repetition effect in a lexical decision task. *Canadian Journal of Psychology, 36,* 435-444.

Emmorey, K. D. (1985). Auditory morphological priming in the lexicon. *Language and Cognitive Processes, 4,* 73-92.

Feldman, L. B. (1992). Bi-alphabetism and the design of a reading mechanism. In D. M. Willows, R. S. Kruk, & E. Corcos (Eds.), *Visual processes in reading and reading disabilities.* Hillsdale, NJ: Erlbaum.

Feldman, L. B. (1991). Morphological relationships revealed through the repetition priming task. In M. Noonan, P. Downing, & S. Lima (Eds.), *Linguistics and literacy.* (239-254). Amsterdam/Philadelphia: John Benjamins.

Feldman, L. B., & Andjelković, D. (1990, November). *Morphemic and syntactic facilitation in sentence contexts.* Paper presented at the Psychonomic Society, New Orleans, LA.

Feldman, L. B., & Andjelković, D. (1991, November). *Morphological priming at long and short lags in visual word recognition.* Paper presented at the Psychonomic Society, San Francisco, CA.

Feldman, L. B., Andjelković, D. & Fowler, A. (ms in preparation). Morphological and phonological analysis in beginning readers.

Feldman, L. B., & Fowler, C. A. (1987). The inflected noun system in Serbo-Croatian: Lexical representation of morphological structure. *Memory & Cognition, 15,* 1-12.

Feldman, L. B., & Moskovljević, J. (1987). Repetition priming is not purely episodic in origin. *Journal of Experimental Psychology: Learning, Memory & Cognition, 13,* 573-581.

Feustel, T. C., Shriffrin, R. M., & Salasoo, A. (1983). Episodic and lexical contributions to the repetition priming effect in word identification. *Journal of Experimental Psychology: General, 112,* 309-346.

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

Fowler, C. A., Napps, S. E., & Feldman, L. B. (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory & Cognition, 13,* 241-255.

Fowler, A. (1988). Grammaticality judgments and reading skill in grade 2. *Annals of Dyslexia, 38,* 73-84.

Fowler, A. (1990). Factors contributing to performance on phonological awareness tasks. *Haskins Laboratories Status Report on Speech Research, SR-103/104,* 137-152.

Grainger, J. (1990). Word frequency and lexical neighborhood effects in lexical decision and naming. *Memory & Language, 29,* 28-244.

Grainger, J., Colé, P., & Segui, J. (1991). Masked morphological priming in visual word recognition. *Memory & Language, 30,* 370-384.

Hanson, V. L., & Feldman, L. B. (1989). Language specificity in lexical organization: Evidence from deaf signers' lexical organization of ASL and English. *Memory & Cognition, 17,* 292-301.

Hanson, V. L., & Wilkenfeld, D. (1985). Morphophonology and lexical organization in deaf readers. *Language and Speech, 28,* 269-280.

Henderson, L. (1985). Toward a psychology of morphemes. In A. W. Ellis (Ed.), *Progress in the psychology of language* (pp. 15-72). London: Erlbaum.

Henderson, L. (1989). On the mental representation of morphology and its diagnosis by measures of visual access speed. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp 357-391). Cambridge, MA: MIT Press.

Henderson, L., Wallis, J., & Knight, D. (1984). Morphemic structure and lexical access. In H. Bouma & D. Bouwhuis (Eds.), *Attention and performance X* (pp 211-224). Hillsdale, NJ: Erlbaum.

Jackendoff, R. (1975). Morphological and semantic regularities in the lexicon. *Language, 51*, 639-671.

Katz, L., Rexer, K., & Lukatela, G. (1991). The processing of inflected words. *Psychological Research, 53*, 25-31.

Kelliher, S., & Henderson, L. (1990). Morphologically based frequency effects in the recognition of irregularly inflected verbs. *British Journal of Psychology, 81*, 527-539.

Lakoff, G. (1971). *Syntactic irregularity*. New York: Rinehart & Winston.

Laudanna, A., Badecker, W., & Caramazza, A. (1989). Priming homographic stems. *Journal of Memory & Language, 28*, 531-546.

Monsell, S. (1985). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language* (pp. 147-195). London: Erlbaum.

Nagy, W., Anderson, R. C., Schommer, M., Scott, J. A., & Stallman, A.C. (1989). Morphological families in the internal lexicon. *Reading Research Quarterly, 24*, 263-282.

Napps, S. E. (1989). Morphemic relationships in the lexicon: Are they distinct from semantic and formal relationships? *Memory & Cognition, 17*, 729-739.

Napps, S. E., & Fowler, C. A. (1987). Formal relationships among words and the organization of the mental lexicon. *Journal of Psycholinguistic Research, 16*, 257-272.

Rapp, B. C. (1992). The nature of sublexical orthographic organization: The bigram trough hypothesis examined. *Journal of Memory and Language, 31*, 33-53.

Ratcliff, R., Hockley, W., & McKoon, G. (1985). Components of activation: Repetition and priming effects in lexical decision and recognition. *Journal of Experimental Psychology: General, 114*, 435-450.

Rosner, J., & Simon, D.P. (1971). The auditory analysis test: An initial report. *Journal of Learning Disabilities, 4*, 384-392.

Seidenberg, M. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy? In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. (pp 245-263). Hillsdale, NJ: Erlbaum.

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.

Stanners, R. F., Neiser, J. J., Hernon, W. P., & Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior, 18*, 399-412.

Tunmer, W. (1988). Metalinguistic abilities and beginning reading. *Reading Research Quarterly, 23*, 134-158.

Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory & Language, 28*, 649-667.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 18

Units of Representation for Derived Words in the Lexicon

Cristina Burani
Alessandro Laudanna
Istituto di Psicologia del C.N.R., Roma

Increasing experimental evidence on the processing of morphologically complex words in a number of languages (English, Hebrew, Serbo-Croatian, Italian, French, Spanish) has converged on the suggestion that morphological relationships are represented in the lexicon. This seems to hold not only for inflected words but also for derivatives, that is, for words derived from a base through processes of affixation or infixation. The representation of a derivative seems to be connected to that of its base and possibly to the representations of the base's inflected forms and of the other words derived from the same base. The processing and the representation of derived words show both analogies and differences with those of inflected words.

Most of the authors who have studied the processing of derived words would agree on these general issues. When they try to develop more detailed accounts of empirical results, however, their proposals are far more varied. Some models assume that morphologically complex words are represented in decomposed form as root plus affixes, with morphologically related words sharing the same root representation (see, e.g., Caramazza, Laudanna, & Romani, 1988; Taft & Forster, 1976; Tyler, Behrens, Cobb, & Marslen-Wilson, 1990), while others suggest that morphologically related words are represented as whole undecomposed forms interconnected through morphological links (see e.g., Lukatela, Gligorijević, & Kostić, 1980; Fowler, Napps, & Feldman, 1985). Some accounts locate morphological decomposition at a pre-access stage (Taft & Forster, 1975), while others consider it a property of permanently stored lexical entries (Caramazza et al., 1988; Tyler et al., 1990).

In this chapter, we do not discuss the different theoretical proposals in detail. Rather, we focus on the specific issue of how derivatives are represented. We present our experimental studies on visual lexical access to morphologically complex stimuli in Italian. Our account is to a great extent specific to the processing and representation of Italian but includes, in part, more general processing mechanisms. This choice arises from more general assumptions, which we discuss before focussing on the specific question at issue.

First, the issue of morphological processing and representation should be addressed within an architecture of lexical processing in which different components are posited. Specifically, we assume that morphological organization is a characteristic of the lexicon in which lexical entries are permanently represented and stored. This representational component ought to be independent of, although related to, other components of processing or representation. At least two other components of lexical access are assumed. The first component can be conceived as a recognition device which works on access representations, thus interfacing the actual stimulus with its lexical entry (or entries). This component provides addresses to the second component, namely, the internal lexicon, where lexical entries are permanently stored (see section 2 for details). A third component is the semantic system in which representations of meaning and conceptual aspects are stored (for a similar proposal, see Tyler et al., 1990). This architecture implies that the representation of morphological relationships is largely independent of the representation of both semantic relationships on the one hand and orthographic or phonological relationships on the other, as showed by cumulative evidence in the literature (e.g., Bentin & Feldman, 1990; Emmorey, 1989; Napps, 1989; Tyler, Marslen-Wilson, & Waksler, in press). We return to some of these issues, in interpreting the data within our own model.

Our second suggestion is that the characterization of the component in which lexical entries are represented ought to be largely specific to the language (or language-type) considered. In other words, the representational format and the organization of lexical entries cannot but reflect the characteristics of the language under investigation. This implies that the lexical processing system adapts to or models itself on the characteristics of the language it is exposed to and has to process. This assumption methodologically implies what can be considered our third suggestion, which will be formulated now but will be elaborated in the final section of the chapter.

The third suggestion is that when processing and representation of morphologically complex words are studied, both their linguistic properties (e.g., productivity of the included affixes) and their distributional characteristics (e.g., word and morpheme frequency) should be assigned a central role. This is implied by the assumption that representational choices are not only specific to a given language, but may also differ according to the types of morphemes considered. This applies to all types of morphologically complex words, but is particularly evident in the case of derived words and derivational affixes, which are far less homogeneous than inflections. The fact that derived words are rather heterogeneous does not imply that each derivative has its own independent lexical representation. On the contrary, some systematicity or generalizations can be found for derived words as well. However, it should not be assumed that one representation modality is valid for all types of derived words, or for both inflected and derived words. Rather, the characteristics of different types of derivatives must be seriously taken into account in determining principles of morphological relatedness. To this end, linguistics should not be assigned a privileged role. What we have called "distributional" factors, which will be further specified in the course of the paper, might also play a role, and in certain cases might even subsume "linguistic" aspects like productivity.[1] Ascertaining which factors contribute (and to what extent) to the establishment of a morphological relationship in the speaker's (or reader's) lexicon can be seen as the specific goal of psycholinguistic investigations.

These assumptions underlie our studies on the processing of morphologically complex words, which will be described in this chapter with the main emphasis on derived words. The chapter is organized in the following way: In the first section, after a brief sketch of the morphological characteristics of Italian, the issue of the format and organization of

[1] In this chapter we do not discuss the property of productivity of affixes. For an interesting perspective in which productivity is accounted for in terms of distributional factors, see Baayen (1991). A discussion of the role of affix productivity in lexical processing and representation can also be found in Burani (in press a), as well as in Frauenfelder and Schreuder (1991).

representations of derived words are discussed with reference to this language. Specifically, the question of whether the unit around which Italian derived words are organized is the *root* or the *stem* (definitions are presented below) is introduced. Then a first set of data from our work on Italian derived words are described. These data come from lexical decision experiments in which the effects of root frequency and morphemic repetition priming on derived words were compared to effects on inflected words. The first set of data point to analogies in the processing and representation of inflected and derived words, which are phonologically transparent with respect to their base roots, are of medium-low frequency and include a commonly used affix. A second set of data, however, shows differences in the processing of inflected and derived words (although of the same described type) which can be compatible with the hypothesis of different representational modalities.

In the second section, the representations of derived vs inflected words is discussed with reference to the Augmented Addressed Morphology model of lexical access (see Caramazza et al., 1988; Caramazza, Miceli, Silveri, & Laudanna, 1985; Laudanna & Burani, 1985). The role of both distributional and lexical properties (such as frequency of whole words and morphemes, number of word-types containing a given morpheme, boundedness of the root morpheme) in affecting lexical representation is discussed. After the description of a third and a fourth set of more recent data which bear directly on the latter issues, we conclude that, unlike inflected words, whose representation is homogeneous, the representations of derived words may differ according to the types of derived words and derivational affixes considered, and to the frequency of both the words and the morphemes included.

Although the paper focusses on the processing and representation of derived words, evidence and arguments concerning derived words will be compared to those concerning inflected words to suggest a model which can account for both.

Roots and stems as units of representation for derived words

Before addressing the issue of the lexical representation of derived words, we present a brief outline of the main morphological characteristics of Italian. Italian words, with very few exceptions (mainly function words), are inflected by means of inflectional suffixes; that is, Italian words are usually morphologically complex. This has two main implications. On the one hand, the Italian inflectional system is very rich, with many paradigms varying according to grammatical category and, within a grammatical category, according to conjugations (for verbs) or to inflectional classes (for nouns and adjectives). On the other hand, the bare or uninflected *root* never corresponds to an actual word: roots are sequences of phonemes (or graphemes) which become actual words only in combination with inflectional affixes.

This constitutes an important difference from English, in which there are words that formally coincide with the base form (e.g., the singular of nouns, or persons other than the third singular in the present tense of verbs). By contrast, within the class of morphologically complex Italian words, no actual word can unequivocally correspond to the base form, that is, no actual word can be considered the base for the derivation of other words. In Italian, the base for a word's derivation is constituted by the (abstract) root, or by the "theme," namely the root followed by a thematic vowel, as suggested by some recent accounts (see Dressler & Thornton, 1991; Scalise, 1984).[2] From now on, we will adopt the

[2]The *thematic vowel* is a meaningless affix which marks an inflectional class. In Italian, for instance, verbs are divided into three inflectional classes which are characterized by different thematic vowels (-a-, -e- and

view that the derivation of a new word is accomplished by adding prefixes or derivational suffixes to the *root.*

Derived words are inflected like all other words. This implies that a derivative consists of at least three morphemes: the root, the derivational affix (prefix or suffix), and the inflectional suffix. Derivational suffixes attach directly to the root, and inflectional suffixes are added to their right. Not all derivational suffixes are easily distinguishable from the root. In fact, while some derivational suffixes do not modify the root, but are simply added to it, others induce some phonological modifications. Thus in some cases the addition of the derivational suffix determines a phonological sequence which consists of the root followed by the derivational suffix plus the inflectional ending (e.g., the derived word **osservazione** [observation] is formed by combination of the verbal root **osserv-** [to observe] and the derivational suffix -(**a**)**zion**-, plus the inflectional suffix -**e**). By contrast, in other cases the process of derivation induces partial modifications, giving rise to a new phonological sequence combining the root and the suffix (e.g., the derivative obtained from the verbal root **concep-** (to conceive) and the suffix -(**i**)**zion**-, plus the inflection -**e**, is **concezione** (conception), which is the result of morphophonological adaptations).

We will call *stem* the combination of the root with the derivational suffix (independently of its being a new phonological sequence or merely the addition of the suffix to the original root), following a suggestion made by some linguists (see, e.g., Bauer, 1983). According to this proposal, the stem is what is left in a word after the removal of only the inflectional affixes, while the word's root is what is left after the removal of *all* the affixes, both inflectional and derivational. Thus in a word like **osservazione**, the root is **osserv-**, as in the verb **osserv-are** (to observe), while the stem is **osservazion-**, composed by the combination of the root with the derivational suffix -(**a**)**zion**-. When inflected, the stem **osservazion-** gives rise to two actual words, **osservazione** (observation) and **osservazioni** (observations). In a word including only inflectional affixes, root and stem are identical, while in a derivative the root can normally be distinguished from the stem, at least in derivatives that are phonologically transparent with respect to their roots. The distinction between root and stem is more complex in derivatives whose base roots have undergone phonological modifications. This type of derivative is briefly considered in the next section. In the present section we deal exclusively with phonologically transparent derivatives, which constitute the object of most of our investigations.

In terms of lexical representation, while words with only inflectional affixes can be represented either as whole undecomposed forms or, if a decompositional view is adopted, as roots plus inflections, words with derivational affixes have another possibility. A derivative can be represented as a whole form or in one of two decomposed forms, namely as root plus derivational affix plus inflectional suffix (e.g., **osserv-azion-e**), or as stem plus inflectional suffix (e.g., **osservazion-e**). These possibilities hold in principle for both prefixed and suffixed derivatives, although there may be differences in the representation according to other differences between prefixed and suffixed words.

The characteristics of the Italian inflectional system, in addition to the consideration that inflections constitute a paradigm (with sub-paradigms), has led us to adopt a view in which the root is the shared unit of representation of morphologically related inflected words. This view is supported by a great deal of empirical evidence (see, e.g., Burani, Salmaso, &

-**i**- respectively). Thus, the infinitive of a 1st conjugation verb is e.g. **amare**, its theme for the imperfect tense is **amav-**, etc.; for a 2nd conjugation verb we have e.g. **vedere**, **vedev-**; for a 3rd conjugation verb we have **sentire**, **sentiv-**.

Caramazza, 1984; Caramazza et al., 1988). For derivatives, both data and theoretical accounts are far more problematic. Schematically, the questions that have been addressed but still deserve more investigation are the following: Are derivatives represented as whole forms or in decomposed form? If they are represented in decomposed form, do their lexical entries consist of the root plus derivational affixes plus inflectional suffixes, or do they consist of the stem plus its inflectional suffixes? In the latter case, how is the representation of the stem related to the representation of the root from which the word is derived?

The heterogeneity of derived words in many respects (various phonological and semantic relations with their bases, types of derivational affixes included, and other factors which will be discussed throughout the paper) suggests that it would be best not to adopt a unitary solution for the lexical representation of derivatives. For instance, the frequency of a derivative, as well as its semantic or phonological transparency with respect to the base, have been indicated as factors possibly affecting lexical representation, with high-frequency and (phonologically and semantically) opaque derivatives more likely to be represented independently from their bases (e.g., Bybee, 1985, 1988; Tyler et al., in press). In this paper, we do not discuss the semantic properties of derived words, which have only recently begun to be studied experimentally (see, e.g., Bentin & Feldman, 1990; Tyler et al., in press). Phonological transparency, which has been experimentally investigated in more detail (see, e.g., Fowler et al., 1985; Napps, 1989; Stanners, Neiser, & Painton, 1979) will be only briefly considered. By contrast, we will discuss, with reference to our studies on Italian, other distributional factors which may affect the processing and representation of derivatives.

We will now briefly describe some of our data, which show analogous effects in the processing of suffixed derived words and inflected words, and thus seem to support analogies in the representation of the two types of morphologically complex words. These data show evidence for the root as a shared unit of representation for both inflected and derived words. A first set of data (Burani & Caramazza, 1987) exploited the *root frequency effect* on lexical decisions for suffixed derived words, which was first shown to play a role in lexical decisions for inflected words (Burani et al., 1984; Taft, 1979;) and, with some exceptions, for suffixed derived words as well (Bradley, 1979).

Both Taft (1979) and Burani et al. (1984) found, for English and Italian respectively, that latencies and accuracy in lexical decisions for inflected words are affected by the frequency of the inflected word's root as well as by the frequency of the inflected word as a whole, namely its "surface" frequency. "Root frequency" means the cumulative frequency of all the words sharing the root, while "surface" or "whole-word" frequency refers to the frequency of a specific word containing that root. Thus a low-frequency word can belong to a morphological set, or family, whose cumulative frequency is high because it includes one or more other words which have a higher frequency. In this latter case, is lexical decision for the low-frequency word a function of its own (low) frequency, or is it a function of its (high) root frequency?

The two studies mentioned showed that, for inflected words, latencies and accuracy are affected by both surface and root frequency. This result seems to imply that frequency plays a role at two different stages or components of processing, one in which the word is represented as a whole and one in which it is represented as root plus affixes. Independently of the theoretical solutions adopted by the different models of lexical access, the root frequency effect shows that the root is activated in some component of lexical access.

Does root frequency play a role in lexical decision for derived words as well? If this were the case, there would be evidence for activation of the root as a processing/representation unit for derivatives too. A first study of English suffixed derived words (Bradley, 1979) showed root frequency effects for some types of derivatives, namely those suffixed by **-er**, **-ment** and **-ness**, which preserve the phonological characteristics of the base word, but not for those suffixed by **-ion**, a suffix which affects the root morpheme's phonology (e.g., **prevention** from **prevent**), sometimes affects stress placement (e.g., **discrimination** from **discriminate**), and very often also affects the root morpheme's spelling (e.g., **division** from **divide**).

Burani and Caramazza (1987) extended these findings to a set of Italian derivatives selected on the basis of criteria which tended to maximize the possible role of the derivative's root in lexical access. The root of the suffixed derivatives did not differ phonologically or orthographically from the base root; the derivational suffixes were mostly productive; the derived words were of medium-low frequency.[3]

Burani and Caramazza (1987) found that root frequency affected lexical decisions for these derived words. Subjects were quicker and more accurate when responding to the medium-low frequency derivatives which included a high-frequency root than when responding to derivatives of the same (medium-low) surface frequency, and with the same suffixes, but with a root of lower frequency. However, in a second experiment it was found, differently from Bradley (1979), that surface frequency too affected access to derivatives. Derivatives matched for root frequency were responded to more quickly and more accurately when they had higher surface frequency. These findings on suffixed derivatives were replicated for French by Col, Beauvillain and Segui (1989), who, on the other hand, failed to show the root frequency effect for prefixed derivatives. Thus access to suffixed derivatives, like access to inflected words, seems to imply activation of both the root's and the whole word's representations, in different processing components .

Further evidence pointing to analogous conclusions comes from lexical decision experiments which exploited the *morphemic repetition priming effect.* This effect has been found in many studies of visual lexical decision (see, e.g., Fowler et al., 1985; Stanners et al., 1979), and consists in the fact that lexical decision for a word is facilitated when the word has been preceded in the list (at a varying lag) by another word morphologically related to it. The facilitation has been found at different prime-target lags (usually of 8-12 intervening stimuli but extending up to 48 stimuli), and it holds for both inflectionally and derivationally related words, although its amplitude is sometimes larger for inflectionally related than for derivationally related (both prefixed and suffixed) words.

The usefulness of this paradigm in studying lexical access has been challenged for involving episodic memory and post-access or strategic mechanisms (see Monsell, 1985, for a discussion of the different components involved in this task). However, it has been argued that, if some experimental constraints are adopted, morphemic repetition priming can show genuinely lexical effects (see Fowler et al., 1985; Monsell, 1985; Napps, 1989). This paradigm has produced effects which are quite stable and large, and occur under many experimental conditions. They thus differ from other priming effects, like orthographic or semantic effects, which vary widely with variations in the materials or the experimental

[3]The derivational suffixes included **-evole** and **-(a)/(i)bile** (approximately corresponding to English **-able**); **-zione**, **-mento**, **-tore** (English: **-tion**, **-ment**, **-er**); **-(a)/(e)nza** (approximate English equivalents: **-(a)/(e)nce**, **-hood**, **-ness**, **-ity**, **-tion**, **-ency**). These same suffixes were used (in different proportions and included in different words) in all the experiments we describe in the first section.

conditions (see, e.g., Forster, Davis, & Schoknecht, 1987; Humphreys, Evett, Quinlan, & Besner, 1987; Napps, 1989). In summary, morphemic repetition priming effects are distinct from both formal (orthographic or phonological) and semantic effects, and converge with other data in suggesting that the lexical representation of morphological relatedness is not reducible to either formal or semantic relationships alone (see Bentin & Feldman, 1990; Emmorey, 1989; Napps, 1989; Tyler et al., in press).

For Italian, Burani and Laudanna (1988) showed that, at an 8-12 lag, suffixed derived words of the type studied in Burani and Caramazza (1987) primed target inflected words having the same root. This effect was not significantly different from the effect obtained by inflected primes sharing the root with the inflected targets (e.g., the derivative **osservazione** primed the inflected word **osserviamo** [we observe] in the same way as the inflected verbal form **osservavate** [you observed] primed **osserviamo**). These results, which were found to be distinct from orthographic effects (see Laudanna and Burani, 1986) were replicated by Laudanna, Badecker, and Caramazza (in press) with contiguous primes and targets (SOA: 200 msec). Morphemic repetition priming effects thus suggest that access to a morphologically complex prime (either inflectionally or derivationally affixed) leads to activation of morphological relationships in the lexicon which benefit the morphologically related target word when it occurs. Specifically, access to a (suffixed) derivative would lead to activation of the derivative's root, as demonstrated by the facilitation produced for the inflected target which shares the same root.

In summary, this first set of data on root frequency, repetition priming and contiguous priming effects point to analogies in processing and representation of inflected and suffixed derived words, and suggest that morphological relationships exist in the lexicon between inflected words and the derivatives that share the same unmodified root with the inflected words. Moreover, the data suggest that suffixed derivatives of this type are represented in the lexicon as root plus derivational and inflectional affixes. However, the presented data do not seem to exclude completely the hypothesis that the unit of representation for suffixed derived words is the stem rather than the root. In principle, the morphological effects that were discussed might also arise from morphological links existing within the lexicon between two distinct, although related, lexical entries, one corresponding to the stem and one to its base root. To distinguish between the two hypotheses we need more constraining results. A second set of data which will now be discussed seems to add further evidence.

These data (see Laudanna et al., in press) point to some differences between the processing of inflected and derived words, which might favor the hypothesis of stem representation for suffixed derived words. Experiments 2 and 3 in Laudanna et al. were based on the processing of words containing homographic roots (see also Laudanna, Badecker, & Caramazza, 1989). Homographic roots are roots that are orthographically identical but grammatically and semantically different. To give an example, the Italian word **portavano** (they carried) is formed by the combination of the verbal root **port-** (first conjugation) and the verbal inflectional suffix **-(a)vano** (past tense, third plural person). The root **port-** is orthographically ambiguous because it is identical to the nominal root **port-** included in the nouns **porta** (door) and **porte** (doors).

When presented simultaneously or with an interstimulus delay of 200 milliseconds, two inflected words containing homographic roots show a robust and consistent inhibitory effect when confronted in a lexical decision task with control word pairs which contain orthographically similar or unrelated roots (see Experiments 1-3 in Laudanna et al., 1989; Experiments 2 and 3 in Laudanna et al., in press). This inhibitory effect was interpreted as the result of a link existing between the homographic roots in the lexicon: specifically, it was argued that when the two words are presented simultaneously, the activations of the two root entries interfere with each other, while, when the two words are presented sequentially, the activation of the first root interferes with the following attempt to activate the orthographically identical root entry. In either case, it was assumed that morphologically decomposed representations are activated in the course of access to inflected words and that, among these representations, grammatically distinct roots having the same orthographic structure are connected by inhibitory links.

In a subsequent study (Laudanna et al., in press, Experiments 2 and 3), the study of root homographs was extended in order to investigate the effect on a target inflected form (e.g., **mute** [mute, fem.pl.], whose root is **mut-**) induced by a prime *derived* form with a homographic root (e.g., **mutevole** [changeable]) in comparison with the effect induced on the same target by a prime *inflected* form with a homographic root (e.g., **mutarono** [they changed]). The sequence **mut-** is the root of both **mutarono** and **mutevole**, whereas the stems of these two words are **mut-** and **mutevol-**, respectively. The assumption on which these experiments were based was that if the representation of derived words in the lexicon consists of their roots plus derivational as well as inflectional affixes, then when a derivative is presented as a prime for an inflected target root homograph, the root component of the derived word (e.g., **mut-** in **mutevole**) should inhibit the homographic root of the inflected form (e.g., **mute**). This effect should not differ from that observed when an inflected root homograph (e.g.. **mutarono**) is used as prime for the same target. Alternatively, if derived words are represented in terms of their stems and inflectional affixes, a word like **mutevole** would access its stem entry (**mutevol-**), which is not identical to the target's homographic root (**mut-**). According to this hypothesis, the inhibitory effect observed with inflectional root homographs should disappear when derived words are employed as primes.

The results of the two experiments showed that, while inflected root homographs inhibited their targets, derived primes did not; under the assumption described above, these results would support the stem representation hypothesis for derived words. Only in the case of inflected root homographs, would the addressed entries—the roots—have an identical orthographic structure, which in turn would activate an inhibitory link between them. By contrast, when derived and inflected root homographs are processed, the activated decomposed lexical entries correspond to stems and roots, respectively, which are not orthographically identical and therefore would not inhibit each other.

However, an alternative hypothesis might, in principle, be advocated without challenging the root representation hypothesis for both inflected and derived words. On this account, two homographic roots would still be represented as separate entries in the lexicon, both linked with their own sets of admissible inflectional and derivational suffixes. On this view, the inhibitory links would be posited either between each root and all the inflectional suffixes which are not combinable with it, or between different sets of inflectional suffixes (for example between those combinable with verbal and adjectival **mut-**, respectively). However, in order to account for the experimental results, no inhibitory link would be posited between a given root and the sets of derivational suffixes not combinable with it, or between inflectional and derivational suffixes. In other words, the sets of inflectional and derivational affixes would be independent. In this case it would be possible to explain the results of Laudanna et al. (in press) by maintaining the root representation hypothesis for derivatives as well (at least for derived word-types that match the characteristics described in this paragraph). This would imply that, whereas the activation of the representation for the inflected word **mut-arono** interferes with the subsequent activation of the inflectional suffix **-e** in the word **mut-e**, the same does not hold when the representation of the derived word **mut-evole** is activated before the inflectional suffix **-e** included in **mut-e**.

Variables affecting processing and representation of derived words

In the preceding section we argued that current data on Italian suffixed derivatives may be compatible with both the hypothesis of root representation and the hypothesis of stem representation. As already stated, derivatives constitute a rather heterogeneous set and their lexical representations might be constrained by many lexical and distributional factors. In order to assess the probability that one or the other forms of representation (either in terms of the root or of the stem) is privileged for a given type of derivative, these factors must be specified. In the present section, the role of some lexical and distributional factors will be taken into account in their interaction with the functioning of the processing components assumed by a model of lexical access.

We will first specify some characteristics of the Augmented Addressed Morphology model (hereafter: AAM model) which are relevant to the issue under discussion. We will then present some data from our more recent studies which assessed the role of distributional factors in affecting lexical representation.

In the AAM model, two main components are involved in lexical access.[4] The first is a processing component which works on access representations and constitutes an interface between the actual stimulus and its lexical entry (or entries). Lexical entries are stored in the second component, namely the lexicon. Lexical entries, most of which are morphologically decomposed, are addressed through activation of the first component, that is, the recognition mechanism or address procedure which takes the printed stimulus as input and gives as output an address to one (or more, in the case of morphologically complex stimuli) entries in the lexicon.

We will now discuss the first component (the Address System). It consists of two subsystems, which operate in parallel on two types of access representations at different time rates. One system operates on the whole string corresponding to a word, whereas the other one operates on sub-lexical units corresponding to morphemes. However, both procedures address morphologically decomposed lexical entries in the second (lexical) component. When a morphologically complex stimulus is presented as input to the lexical processing system, both the address representations (those corresponding to whole words and those corresponding to roots and affixes) are activated (for a similar view, see also Frauenfelder & Schreuder, 1991).

Experimental data supporting this architecture come mainly from studies on morphologically complex non-words (Caramazza et al., 1985; Caramazza et al., 1988). The first of these studies (Caramazza et al., 1985) showed that a dyslexic patient who could not read nonsense words was by contrast able to read them when they were formed by two real morphemes (a root and an inflection), although not legally combined (e.g., **cantevi**, formed by the real verbal root **cant-**, and by the real suffix **-evi**, incompatible with that root). The second study (Caramazza et al., 1988) showed that morphologically complex

[4] We do not treat the third component of lexical access, namely the semantic system in which semantic information is stored. As already mentioned, our discussion is limited to formal and structural characteristics of derived words, not directly treating the role of semantic factors, which constitute a relevant dimension but would require additional considerations.

stimuli of the **cantevi** type took longer to be rejected as non-words by subjects in a lexical decision task. Both these results converge in showing that, in the absence of whole-word units, recognition units corresponding to morphemes can be exploited to address the lexicon.

We will now consider the processing of real words, for which whole-word access units usually exist. A number of factors affect the time course of activation of the recognition units (or access representations) available for a real word, thus affecting the probability that one of the two systems will be more efficient and faster in leading the stimulus to contact the lexicon. Frequency is certainly one of the most important factors, in that high-frequency words are presumably treated more efficiently by the first mechanism as whole forms, whereas low-frequency words may be processed more quickly by the second system, in which their constituent morphemes are activated, at least when the words include morphemes whose frequency is much higher than the whole-word frequency. The assumption underlying this hypothesis is that the frequency of occurrence in the language of a given lexical or sub-lexical unit strongly affects its availability as an access or recognition unit.

If we now consider the processing of suffixed derivatives of the type studied in Laudanna et al.'s (in press) experiments, as well as in the other studies of ours that have been described, we can make the following observations. All these derived words were of medium-low frequency and included a root which was shared by other inflected words, as well as a commonly used suffix. Therefore the cumulative frequency of a derivative's root was higher than the frequency of the derivative as a whole form. At the same time, each derivative included a suffix of higher frequency than the derived word as a whole. In summary, the constituent morphemes of these derived words had systematically a higher frequency than the words in which they appeared. It might therefore be thought that in these cases, analogously to the inflected cases above discussed, the derived words are more efficiently processed by the recognition system which works on morphemes; thereby addressing the lexical entries for roots and derivational suffixes separately.

By contrast, higher-frequency derivatives might be processed more efficiently through the address procedure that operates on whole words. The complex balancing and correlation of different factors such as the frequency and the orthographical and phonological transparency of the derived word with respect to its base root is certainly relevant for establishing one form or another of representation at the access stage.

Let us consider, for example, suffixed derivatives which have undergone phonological modifications with respect to the roots which constituted their bases (e.g., **concezione** [conception], whose base root is **concep**-ire [to conceive]). Unlike the derivatives considered so far, which included unmodified base roots, derivatives like **concezione** cannot be processed by the address procedure, which works on recognition units corresponding to roots. In fact the root **concep-** is not present as a unit in the word **concezione**. These derivatives are plausibly processed and represented as stems (**concezion-**) and inflectional suffixes (-**e** and -**i**), independently of being of low or high frequency.

In summary, different representational possibilities should be admitted for different types of derivatives. However, frequency (of the whole form and of the morphemes) and phonological transparency are not the only relevant factors in assigning one form of representation or another to a given (type of) derivative. These factors must be considered in correlation with other lexical and distributional characteristics of words. Hereafter we present some of our more recent data which were intended to assess the role of some of the latter factors in determining if a potential morpheme is a unit of access and representation.

We began our assessment of the role of some distributional properties for both affixes (in particular prefixes) and roots occurring in derived words, with two sets of experiments that will be described here. It should be remembered that the rationale for investigating the role of distributional factors is based on the more general hypothesis that the access system should reflect (or adapt itself to) the distributional properties of the input, especially when a linguistically defined class of input units (e.g., prefixes) are distributed very heterogeneously in many respects.

In the first study (Laudanna, Burani, & Cermele, forthcoming), we addressed the issue of prefix processing and representation by manipulating two quantitative dimensions of prefixes in Italian that might be important in establishing a prefix as a potential unit for morphological decomposition. The first dimension taken into consideration was the absolute number of word types in the language which include a given prefix. Prefixes display a high degree of variability with respect to this quantitative parameter. An inspection of the sets of two- and three-letter prefixes in Italian showed that the number of word types containing each of them covers a wide range going from about 2300 occurrences (for the prefix **ri-**, approximately corresponding to the English **re-**) to only 20 occurrences (for the prefixes **fra-** and **su-**, approximately corresponding to the English **inter-** and **above-**, respectively). There are reasons for thinking that this simple measure (number of word types beginning with a given prefix) might be related to subjects' experience of that prefix in different word contexts and hence to the likelihood that the orthographic pattern corresponding to a given prefix is abstracted, becoming a permanently stored unit at some stage of the lexical processing system.[5]

Prefixes also differ from each other to a large extent according to a second distributional parameter experimentally manipulated in Laudanna et al.'s study. This parameter also appears to be potentially important in predicting the probability that a given prefix is available for morphological decomposition in the lexical processing system. This second parameter is defined as the quantitative relation between the number of truly prefixed word types beginning with a given prefix (e.g., **ri-dare** (to give again), where the orthographic sequence "ri" corresponds to the prefix **ri-**) and the number of word types in which the same orthographic sequence is present as a pseudoprefix (e.g., **rid-ere**, (to laugh), in which the orthographic sequence "ri" does not correspond to a real prefix, the morphemic constituents being the verbal root **rid-** and the inflectional suffix **-ere**). It goes without saying that when the proportion of real prefixes to pseudoprefixes with the identical orthographic sequence is relatively high, the process of morphological decomposition may be more often successful. Minimizing the number of false alarms (or incorrect morphemic decompositions) would correspond to a principle of economy (or functionality) of processing that should make it more likely for some prefixes rather than others to be represented in the lexical processing system (see Frauenfelder & Schreuder, 1991, for a discussion of the economy of processing constraints). From an inspection of the Italian lexicon it was found that two- and three-letter prefixes show a great variability from this point of view as well. The percentages of truly prefixed words with respect to the total number of prefixed and pseudoprefixed words beginning with a given orthographic sequence vary from 4% to 81%.

Both the variables described seem to be potentially relevant in determining the probability that a prefix will be activated in the course of lexical access: prefixes occurring in more

[5]This measure is highly correlated with prefix frequency in the language, even though it must be noted that some degree of deviation may occur between the two measures, since frequency is based on both word-token and word-type occurrences.

word types and/or more often than homographic pseudoprefixes should be more likely activated as processing units. The experiments in Laudanna et al. (forthcoming) were based on the consideration that it should be possible to assess the relative importance of the two variables by comparing in an experimental task prefixes which have divergent values for one or both of those variables. Therefore, two lexical decision experiments were carried out in which the experimental stimuli consisted of non-words resulting from the morphologically illegal combination of a prefix with a real word (e.g., **riviale**, formed by the combination of the prefix **ri-** with the noun **viale** [avenue]). The first experiment was based on a multiple regression design and the prefixes selected to form the non-words were chosen to allow the evaluation of single effects of each of the two independent variables described above as well as the combined effect of both. Each prefixed non-word formed by a prefix followed by a real word was matched with a non-word having the same embedded word preceded by an orthographic sequence of the same bigram or trigram frequency as the prefix in the critical non-word (e.g., **paviale**, which is formed by the noun **viale** preceded by the orthographic sequence "pa," as frequent as "ri" at the beginning of words in Italian). The results showed that the only variable which significantly predicted reaction time and error distributions in lexical decision was the ratio between the number of word types in the language containing a given prefix and the number of word types beginning with an identical orthographic sequence: the higher this ratio for a given prefix, the higher the reaction times and errors on non-words containing that prefix. Although positively correlated with reaction times and errors, the first variable, namely the absolute number of word types containing a given prefix, did not significantly predict lexical decision performance.

In the second experiment, the structure of both prefixed and control non-words was kept constant, but the effect of the two different distributional parameters was submitted to a more stringent test. Three categories of prefixed non-words were selected, each with its own controls: in the first category, prefixes at the beginning of the non-words had low values for both of the independent variables (a low number of word types and a low proportion of truly prefixed words). In the second category, prefixes occurring at the beginning of the non-words were equally included in a small number of word types in the language, but this number represented a high proportion of words with that initial sequence. Finally, the third category of non-words contained prefixes which had high values for both distributional parameters. Again, the results of the second experiment showed a significant effect on reaction times and errors obtained for just those non-words containing prefixes occurring in a high proportion of truly prefixed words.

Taken together, the results of the two experiments in Laudanna et al. (forthcoming) show that prefixes may be represented as units of processing and/or representation in the lexical access system and that they can be activated during the recognition process. However, the probability of their being represented and processed during access seems to be tied to their "salience," defined in terms of a distributional property, that is the numerical relationship between prefixed and pseudoprefixed words in the language. Moreover, independently of these first results, it is important to point out that this parameter, as well as other not yet investigated distributional properties of both whole words and their constituent morphemes, might be relevant for understanding the way in which derived words are processed and represented.

In the research just described (Laudanna et al., forthcoming), we investigated how some distributional properties of sub-lexical units affect the probability that a derivational affix, specifically a prefix, is taken as a processing unit in lexical access. We now present the second study carried out recently (Burani, Laudanna and Cermele, in press), in which we investigated how other distributional properties of words might also affect the lexical representation of root morphemes. We addressed the general issue by investigating a specific question, namely: Are prefixed words with *bound roots* represented in the lexicon analogously to prefixed words with *free roots*, or do they have a different type of representation?

In many languages, including Italian, prefixes may be added either to a free-standing root (as when **re-** is added to the English root **play** to form **replay**), or to a bound root, that is a root which does not constitute a word by itself, but only in combination with a derivational affix (e.g., **-gress in ingress, progress, regress**). (For a discussion of bound roots, see, e.g., Selkirk, 1982, p.98.) It might be argued that prefixed words with free roots are better candidates for the hypothesis of the root as the unit of processing and representation, in that their root is shared by other words, either inflected or derived (e.g., **plays, player**), while a bound root never occurs alone, but only in combination with a prefix (that is, as a stem). Moreover, the meaning of a free root tends to be constant across the various words in which it is included, while prefixed words having the same bound root tend to differ more in meaning than prefixed words sharing the same free-standing root. (For further discussion of these issues, see also Burani, in press a and b).

Because of the lack of a constant meaning, bound roots are not considered morphemes in a traditional view of morphology. However, reasons have been advanced, on linguistic grounds, for considering prefixed words with bound roots as morphologically complex. Aronoff (1976) argued that some words including the same phonological string, even though it has no constant meaning across different forms (e.g., **-mit** in the words **remit, commit, transmit, submit, permit, admit**) are morphologically related because they undergo the same morphophonological rule (giving rise to **remission, commission, transmission** and so on).

When lexical processing is considered, the principle for morphological organization provided by Aronoff (i.e., a morphophonological rule) does not necessarily play a central role. By contrast, other factors having to do with the distribution and frequency in the language of lexical and sub-lexical units may prove relevant in affecting the probability that a given phonetic (or orthographic) string will act as a morphemic unit (in this case, a root morpheme) even in the absence of meaning relations (see again Burani, in press a).

Specifically, and analogously to what was shown by Laudanna et al. (forthcoming) for prefixes, the number of word types in which a potential root is included might affect the probability that a bound root will act as a processing/representation unit. In other terms, this probability might be increased when a bound root is included in many different word types, that is, when it occurs in combination with various affixes.

The data we collected for assessing the latter hypothesis, which will now be briefly described (see also Burani et al., in press) were drawn from a free-recall task and consist in the analysis of the patterns of morphological errors produced by subjects. Thus these results are not of the same type as those discussed in the rest of this paper. However, we think that they may provide a useful view for investigating the more general question at issue.

The logic of analyzing the patterns of morphological errors (in this case, prefix substitution errors), produced by subjects in an experimental situation in which they had to recall lists of prefixed words, was the following. If prefix substitution errors do occur, and if experimental control can ensure that these errors do not have a mere semantic or phonological source but reflect aspects of morphological organization, then the distribution of these errors along the experimental variables which are supposed to affect processing may shed light on the principles governing organization and retrieval of morphologically complex words. Specifically, if prefix substitution errors are distributed differently according to different types of prefixed derivatives, this should indicate that prefixed derivatives are represented in partially different ways in the lexicon (see also Burani, in press b).

In Burani et al.'s (in press) study, two factors, namely root type (whether free or bound) and number of prefixed word-types sharing the same root, were investigated. Four sets of prefixed Italian words were selected: (i) words with free roots and a low number of prefixed words sharing the same root; (ii) words with bound roots and the same low number of related prefixed words; (iii) words with free roots and a high number of prefixed words sharing the root; (iv) words with bound roots and a high number of prefixed words sharing the root.

The expected outcomes were the following. If the lexical representation of prefixed words with bound roots is organized around the root in the same way as for free roots, then prefixed words with bound roots should produce about the same number of prefix substitution errors as words with free roots. If, on the contrary, prefixed words sharing the same bound root are not related or are more weakly related in the lexicon, they should give rise to fewer prefix substitution errors than words with free roots. An alternative prediction is that, among prefixed words with bound roots, only those whose morphological family contains many members give rise to as many morphological errors as those found for prefixed words with free roots.

The results showed that the only variable which affected the number of prefix substitution errors was the number of related prefixed words, with prefixed words possessing a higher number of prefixed relatives giving rise to more prefix substitution errors than words with a lower number of prefixed words sharing the same root. More interestingly, no interaction was found between the two variables, namely root type and number of related forms. In other terms, when the number of word types sharing the same root (whether free or bound) is kept constant, the activation of a related form is equally probable for words with free roots and words with bound roots. In more general terms, these data indicate that some linguistic distinctions, such as root-boundedness, may not play a differential role in word processing and representation, in cases where they show similar distributional properties.

Conclusions

In this chapter we have argued that there is ample evidence from experimental tasks that morphological relationships are represented in the lexicon not only for inflected words but also for derived words. However, we have also argued that, since derived words are less homogeneous than inflected words, their lexical representations may also be less homogeneous than those of inflected words. Specifically, in addressing the issue of whether Italian derived words are represented in terms of their roots or in terms of their stems, we have argued that, while there is clear evidence that the shared unit of representation for inflected words is the root, existing evidence for derived words seems to be compatible with both the hypotheses of root and of stem representation.

The evidence we have obtained from lexical decision tasks for low-frequency suffixed derivatives which are orthographically and phonologically transparent with respect to their base roots, and include frequently used derivational suffixes, although it is compatible with the root representation hypothesis, does not clearly favor this account over the stem representation hypothesis. Other types of derivatives (phonologically opaque and high-frequency derivatives) might be better processed and represented as stems plus inflectional endings. However, phonological transparency and frequency are not the only relevant factors in determining the probability that a given derivative is represented in terms of its root or of its stem. We have argued that other distributional properties of both words and morphemes may play an important role.

Some of our more recent results, drawn from lexical decision and free-recall tasks on prefixed stimuli, show that distributional factors such as the number of word types sharing a given sub-lexical unit, or the ratio of truly prefixed to pseudo-prefixed words in the language, affect the likelihood that a potential morpheme will establish itself as a unit of processing and representation. All these factors, which are related to the orthographical/phonological (or formal) characteristics of words (orthographical and phonological transparency, word and morpheme frequency or distribution), strictly constrain the functioning of the lexical access system. The role of these formal factors must obviously be considered in interaction with the semantic properties of both words and morphemes (and with the representational components which store this kind of information), which have not been taken into account in this chapter.

Acknowledgment
The authors would like to thank Anna Maria Thornton for her helpful comments on an earlier version of this chapter.

References

Aronoff, M. (1976). *Word formation in generative grammar.* Cambridge, MA: MIT Press.
Baayen, H. (1991). Quantitative aspects of morphological productivity. In G.E. Booij & van Marle, J. (Eds.), *Yearbook of morphology.* Dordrecht: Foris Publications.
Bauer, L. (1983). *English word-formation.* Cambridge: Cambridge University Press.
Bentin, S., & Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology, 42A,* 693-711.
Bradley, D. (1979). Lexical representation of derivational relation. In M. Aronoff & M.L. Kean (Eds.), *Juncture.* Cambridge, MA: MIT Press.
Burani, C. (in press a). What determines morphological relatedness in the lexicon? In G. T. M. Altmann & R. C. Shillcock (Eds.), *Cognitive models of speech processing: The Sperlonga meeting II.* Hove: Erlbaum.
Burani, C. (in press b). The lexical representation of prefixed words: Data from production tasks. In F. Kiefer (Ed.), *The interface between morphology and syntax.* Amsterdam: John Benjamins.
Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes, 2,* 217-227.
Burani, C., & Laudanna, A. (1988). Relazioni morfologiche tra forme derivate e forme verbali. *Giornale Italiano di Psicologia, 15,* 625-640.
Burani, C., Laudanna, A., & Cermele, A. (in press). Errors on prefixed verbal forms: effects of root type and number of related forms. *Rivista di Linguistica.*
Burani, C., Salmaso, D., & Caramazza, A. (1984). Morphological structure and lexical access. *Visible Language, 4,* 348-358.
Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form.* Amsterdam: John Benjamins.
Bybee, J. L. (1988). Morphology as lexical organization. In M. Hammond and M. Noonan (Eds.), *Theoretical morphology.* San Diego: Academic Press.
Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28,* 297-332.

Caramazza, A., Miceli, G., Silveri, M. C., & Laudanna, A. (1985). Reading mechanisms and the organization of the lexicon: Evidence from acquired dyslexia. *Cognitive Neuropsychology, 2,* 81-114.

Col, 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.

Dressler, W. U., & Thornton, A. M. (1991). Doppie basi e binarismo nella morfologia italiana. *Rivista di Linguistica, 3,* 3-22.

Emmorey, K. D. (1989). Auditory morphological priming in the lexicon. *Language and Cognitive Processes, 4,* 73-92.

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.

Fowler, C. A., Napps, S. E., & Feldman, L. B. (1985). Relations among regularly and irregularly morphologically related words in the lexicon as revealed by repetition priming. *Memory and Cognition, 13,* 241-255.

Frauenfelder, U. H., & Schroeder, R. (1991). Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In G. E. Booij & van Marle, J. (Eds.), *Yearbook of morphology.* Dordrecht: Foris.

Humphreys, G .W., Evett, L. J., Quinlan, P. T., & Besner, D. (1987). Orthographic priming: Qualitative differences between priming from identified and unidentified primes. In M. Coltheart (Ed.), *The psychology of reading.* London: Erlbaum.

Laudanna, A., Badecker, W., & Caramazza, A. (1989). Priming homographic stems. *Journal of Memory and Language, 28,* 531-546.

Laudanna, A., Badecker, W., & Caramazza, A. (in press). Processing inflectional and derivational morphology. *Journal of Memory and Language.*

Laudanna, A., & Burani, C. (1985). Address mechanisms to decomposed lexical entries. *Linguistics, 23,* 775-792.

Laudanna, A., & Burani, C. (1986). Effetti di ripetizione e priming su forme verbali morfologicamente collegate. *Giornale Italiano di Psicologia, 13,* 123-139.

Laudanna, A., Burani, C., & Cermele, A. (forthcoming). What makes a prefix a recognition unit?

Lukatela, G., Gligorijević, B., & Kostić, A. (1980). Representation of inflected nouns in the internal lexicon. *Memory and Cognition, 8,* 415-423.

Monsell, S. (1985). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 2). London: Erlbaum.

Napps, S.E. (1989). Morphemic relationships in the lexicon: Are they distinct from semantic and formal relationships? *Memory & Cognition, 17,* 729-739.

Scalise, S. (1984). *Generative morphology.* Dordrecht: Foris.

Selkirk, E. O. (1982). *The syntax of words.* Cambridge, MA: MIT Press.

Stanners, R. F., Neiser, J. J., Hernon, W. P., & Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior, 18,* 399-412.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition, 7,* 263-272.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14,* 637-647.

Tyler, L. K., Behrens, S., Cobb, H., & Marslen-Wilson, W. (1990). Processing distinctions between stems and affixes: Evidence from a non-fluent aphasic patient. *Cognition, 36,* 129-153.

Tyler, L. K., Marslen-Wilson, W., & Waksler, R. (in press). The representation and access of derived words. In G.T.M. Altmann and R.C. Shillcock (Eds.), *Cognitive models of speech processing: The Sperlonga Meeting II.* Hove: Erlbaum.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 19

Representation and Processing of Morphological Information

Cecile Beauvillain
Juan Segui
Laboratoire de Psychologie Experimentale C.N.R.S.
René Descartes University

The problem of the representation and processing of morphologically complex words has given rise over the past 15 years to a large number of theoretical and empirical studies in linguistics, psycholinguistics, and cognitive neuropsychology. Excellent critical reviews have been published by Henderson (1985) and Taft (1985).

In this chapter, we consider certain aspects of the question of knowing how morphologically complex words are represented in the internal lexicon and accessed during their recognition. Our comments are based on recent work carried out in our laboratory.

In accord with the most widespread definition, we shall consider two words as morphologically related if they share the same root. The set of words sharing a common root constitutes a morphological "family." Generally, the members of a same family share properties of form and meaning. This is the case, for example, of words in the couples "fear" / "fearless" or "animate" / "inanimate." Certain exceptions to this rule exist, however, and the degree of formal or semantic "transparency" between morphologically related words may be very variable. This is particularly true in the domain of derivational morphology.

The work we shall examine in this chapter bears solely on the processing of "transparent" derived words from the point of view of orthography and semantics and aims to answer the following three principal questions: a) in which form are derived words represented in the internal lexicon? b) are the morphological relationships between derived words of a same family represented in the internal lexicon in an explicit way that is not reducible to orthography and semantic relations? c) to what extent does the structure of derived words determine procedures for lexical access?

The first question raises the problem of the mode of representation of derived words and in particular that of knowing if these words are represented in the internal lexicon in their whole form or in a decomposed form. The second question concerns the existence in the internal lexicon of morphological relations between words that are not reducible to orthographic/phonological or semantic relations. Finally, the third question essentially concerns the nature of the procedures and representations used in order to access the lexical representation of derived words.

The format of the representations of morphologically complex words

The format in which morphological information is represented in memory is the subject of an important debate in psycholinguistics. This is a fundamental issue, for the manner in which the nature of lexical representations are envisioned conditions the hypotheses that may be advanced with respect to the representations and the access procedures employed.

An initial hypothesis, mainly advocated by Taft, consists in proposing that the complete representation of derived words is only present in the central lexicon. The important feature of Taft's hypothesis is that these representations of the central system may only be accessed on the basis of information corresponding to their stem. Indeed, according to Taft, it is the stems of derived words that constitute the entry of the input system.

Such a conception of the organization of the input system implies that derived words are "decomposed" into their constituent morphemes prior to access. In particular the affix must be "stripped off" in order to isolate the stem. Following this decomposition procedure, when the full information about the word is retrieved from the central system, this can be checked back to the presented word to verify that the correct lexical entry has been accessed.

An alternative view affirms that derived words have independent lexical entries (the full-listing hypothesis; for example, Butterworth, 1983) and that these entries may be accessed in a "direct" and "continuous" manner without any prior decomposition procedure (for example, Manelis & Tharp, 1977).

Finally, Caramazza et al. (Burani & Caramazza, 1987; Caramazza, Laudanna, & Romani, 1988) have contended that morphologically decomposed entries in the orthographic input lexicon may be accessed either by a whole word procedure (for known words) or by a morpheme address procedure (for novel words or non words).

With regard to this general question, Segui and Zubizarreta (1985) have advanced certain linguistic arguments which tend to demonstrate the necessity of postulating that the lexical entries for derived words must possess information bearing on their whole word form. This necessity is related to the fact that certain processes of word formation (affixation) as well as other morpho-phonological processes require, for their application, information concerning the whole word form of derived words. Thus, for example, the stress pattern of a derived item such as "instrumentality" may not be predicted with reference to that of the basic morpheme "instrument" but rather on the basis of the stress pattern of another derived item of its family which is "instrumental." This example suggests that precise information concerning the surface form of derived words must be indexed in their lexical entry. Furthermore, not only must derived words have their own lexical entry in which this information regarding their whole word form figures, but these entries must also be able to reflect the nature of their internal structure. Segui and Zubizaretta (1985) have proposed that this information must be represented in the mode organization of the morphological families, that is, in the way in which the different items sharing the same root or "head" are related. According to this interpretation, each morphologically derived form constitutes a lexical entry but it is not an isolated lexical entry. A similar interpretation was proposed by Lukatela, Gligorijević, Kostić, and Turvey, (1980) for inflected nouns in Serbo-Croatian. However, according to these authors, noun entries are organized around a particular central form, the nominative case.

Segui and Zubizaretta have suggested that the mode of organization of the morphologically family is determined by general morphological principles and, in particular, by the nature of the affix-frames which encode the phonological, morphological, and syntactic properties of the affixes. It is mainly the knowledge of the affix-frames which allows subjects to understand novel words and to judge the validity of certain combinations of morphemes. From this point of view, it is conceivable that the procedures employed for processing new words or non-words are primarily based on the use made by the subject of the properties of the affix-frames rather than those of the lexical entries themselves.

With respect to the relations maintained between whole word forms themselves, on the one hand, and between whole word forms and their morphemic components, on the other, an alternative proposition consists in expressing these relations at two different levels of representation. This may be done by proposing, in the framework of a hierarchical model of a connectionist type, a level of morphological representation different from the lexical level per se (for example, Fowler, Napps, & Feldman, 1985; Grainger, Cole, & Segui, 1991).

Despite the importance of research carried out in psycholinguistics and cognitive neuropsychology on the mode of representation of morphologically complex words, none of the previously examined interpretations has received sufficient empirical support and the debate remains largely open. If it is legitimate to think that information concerning, on the one hand, the whole word form of derived words and, on the other, their morphemic components, is represented at one or different levels in the processing system, the problem remains to establish in what measure this information is effectively used during the visual or auditory recognition of words; in other words, to what extent does this information constrains the on-line processing of derived words?

In the third part of this chapter we shall examine certain recent findings suggesting that the nature of the information used by the system during the recognition of derived words varies according to its internal structure and in particular according to the respective position of the root and the affix. Before examining this point, we shall present in the following section certain data which suggest that morphological relations are represented in an explicit manner in the internal lexicon.

Are morphological relations different from orthographic and semantic relations?

The favored procedure for tackling the study of relations between words in the internal lexicon is probably that of priming and a great deal of research has been devoted to evaluating the extent to which the effects of morphological priming differ from that of formal or semantic priming.

Most of these studies have shown that the effects of morphological priming are more robust and of greater duration than the effects of formal or semantic priming (for example, Henderson, Wallis, & Knight, 1984; Murrell, & Morton, 1974). However, as Napps (1989) has recently indicated, in these "classical" priming experiments the potential role of strategic processes, post-lexical effects as well as those related to episodic memory, makes their interpretation particularly delicate.

In an attempt to reduce the importance of these non-lexical factors, Fowler, Napps and Feldman (1985) and Napps (1989) investigated the effects of priming repetition by using a reduced number of morphologically related items presented within a long list of unrelated words. In these experimental conditions, the authors observed an important priming effect for morphologically related words. Under the same experimental conditions formal or semantic priming effects are not observed (see also Napps & Fowler, 1987). These studies on the whole suggest that the effects of morphological priming do not result from the simple convergence of orthographic, phonological and semantic relations. Morphological relations, thus, seem to be represented in an explicit way in the internal lexicon.

Despite their undeniable interest, the use made in these studies of relatively large SOA does not permit the exclusion of the intervention of strategic factors in determining the observed effects. In order to eliminate the possibility of the establishment of these strategies it is necessary to drastically reduce the potential role of episodic memory. A procedure capable of attaining this objective is that of masked priming. Using a forward pattern mask (500 ms) and a very brief presentation duration of the prime (64 ms), prime visibility is reduced to a level that excludes the application of any predictive strategy. The lexical nature of the priming effects obtained under these conditions is suggested by the results obtained with respect to the repetition effect (Forster & Davis, 1984; Segui & Grainger, 1990) and formal priming (Forster, 1987; Forster, Davis, Schoknecht, & Carter, 1987; Grainger, 1990; Segui & Grainger, 1990).

In regard to morphological relations, Forster et al. (1987) were the first to demonstrate, using this procedure, a clear effect for facilitation for short words (four or five letters) when the prime and the target are morphologically related. Thus, the masked presentation of "made" facilitates the processing of the target "MAKE" in a manner comparable to that obtained in a situation of identity priming. However, under the same experimental conditions comparable effects are not observed when the prime and the target only share orthographic properties. Indeed, important inhibitory effects and not facilitatory effects were obtained for orthographically similar words when the masked prime is more frequent than the target (for example, "blue" - "BLUR") (Grainger & Segui, 1990a).

Given this contrast between the effects of morphological and orthographic priming under conditions of masked priming we decided to approach these two types of effects in a comparative way (Grainger, Cole, & Segui, 1991). In a series of studies carried out using the masked priming procedure we compared lexical decision times for derived target words (example, "graveur") according to whether these were preceded by a prime composed of a morphologically related word (example, "gravure"), an orthographically related word (example, "gratuit"), or an unrelated word (example, "cristal").

The results obtained in these studies showed that with reference to the situation in which words are orthographically similar the words morphologically related introduce a clear facilitatory effect. However, if the effects of morphological priming are assessed with reference to a standard control situation comprising two unrelated words, no facilitatory effect is observed when the morphologically related words share initial letters. This absence of facilitation seems to be due to the simultaneous existence of inhibitory effects, that are orthographic in nature, and facilitatory effects, that are morphological in nature. An inhibitory effect that is orthographic in nature has also been observed by Henderson et al. (1985) in an unmasked presentation of words. As is the case in our experiments, Henderson et al. obtained response times that were greater for orthographically related words than for unrelated words.

These results, overall, suggest that in order to demonstrate strictly morphological effects in priming experiments may require, in certain cases, that they be delimited from eventual inhibitory effects that are orthographic in nature. Whatever the case may be, the principal finding obtained in the experiments mentioned above suggests that morphological relations between words are represented in an explicit way in the internal lexicon. Facilitatory effects of a morphological nature may not be the result of the conjunction of orthographic and semantic effects.

The role of the sequential organization of constituent morphemes of derived words during their processing. Besides the classical distinction between inflectional morphology (which has an essentially grammatical function) and derivational morphology (which may modify the form and meaning of the stem), it is important to take into consideration the fact that this latter includes two types of affixes, prefixes and suffixes. In prefixed words the affix precedes the root whereas in suffixed words it follows the root. This difference in the sequential organization of prefixed and suffixed words may have important implications with respect to access procedures and this not only in the auditory modality but also in the visual modality. It should be added that from a linguistic standpoint the nature of the information conveyed by prefixed and suffixed words is very different. In particular, it is generally suffixes rather than prefixes which determine the syntactic category of the derived word (see on this point, Cutler, Hawkins, & Gilligan, 1985 and Hawkins & Cutler, 1988).

Given the above, it seems important to approach the study of morphologically complex words taking into consideration these two types of derived words.

In an initial series of studies (Cole, Beauvillain, Pavard, & Segui, 1986), we observed that the recognition of a suffixed word is more rapid when this word is immediately preceded by another suffixed word than when it is by a pseudo-affixed word. Thus, the suffixed word "fautif" is more rapidly recognized following "tardif" (suffixed) than after "nocif" (pseudo-suffixed). On the other hand, no difference was observed for the processing of prefixed words according to the morphological nature of the context word. For example, the lexical decision time for the prefixed word "prénom" is the same when it is preceded by "préface" (prefixed word) or by "préfet" (pseudo-affixed word).

We interpreted this difference in sensitivity of suffixed and prefixed words to morphological context in formulating the hypothesis that whereas prefixed words are accessed on the basis of their whole word form, the presence of the root at the beginning of suffixed words induces an access procedure based on this root. The root constitutes the head of a morphological family and its presence at the beginning of suffixed words allows for the utilization of this root as an entry point into this family. Taking into consideration of the suffix component will then permit the selection of the family member corresponding to the presented stimulus.

According to this interpretation, the morphological structure of suffixed words is exploited on-line by the processing system in order to delimit a "morphological family" of lexical candidates. In contrast, lexical candidates for prefixed words are orthographic in nature and information concerning the morphological nature of these words will only be available and used at a post-access stage.

This general hypothesis which relates in a strict way the nature of the processing carried out to the sequential organization of the morphemic components of the derived word would be all the more plausible when the visually presented word is long. In effect, the left-to-right directionality of processing is all the more marked the greater the length of the word.

If our general hypothesis is correct, the prediction may be made that the access speed of prefixed words will be determined essentially by their surface frequency while that of suffixed words will also be sensitive to their base frequency or "cumulative" frequency.

We tested this hypothesis by comparing the effect of these two factors, surface frequency and cumulative frequency, on the lexical decision time for visually presented prefixed and suffixed words (Cole, Beauvillain, & Segui, 1989). Figure 1 presents the mean lexical decision times for prefixed and suffixed words as a function of their surface frequency and cumulative frequency.

It may be observed that the reaction times for suffixed words depend both on their surface frequency and cumulative frequency whereas only the former affects the reaction times for prefixed words. This finding is in accord, therefore, with the hypothesis advanced above and suggests that while accessing for suffixed words is sensitive to their internal morphological structure, for prefixed words it is only sensitive to their surface organization.

In an additional study we showed that reaction times for suffixed words belonging to the same morphological family also depend on their surface frequency. Thus, the response to the word "vendeur" is faster than for the word "vendable." This finding confirmed that the accessing of members of the family also depends on their respective relative frequency. The morphological family could be organized in such a way that its more frequent members are more rapidly "selected" than its less frequent members.

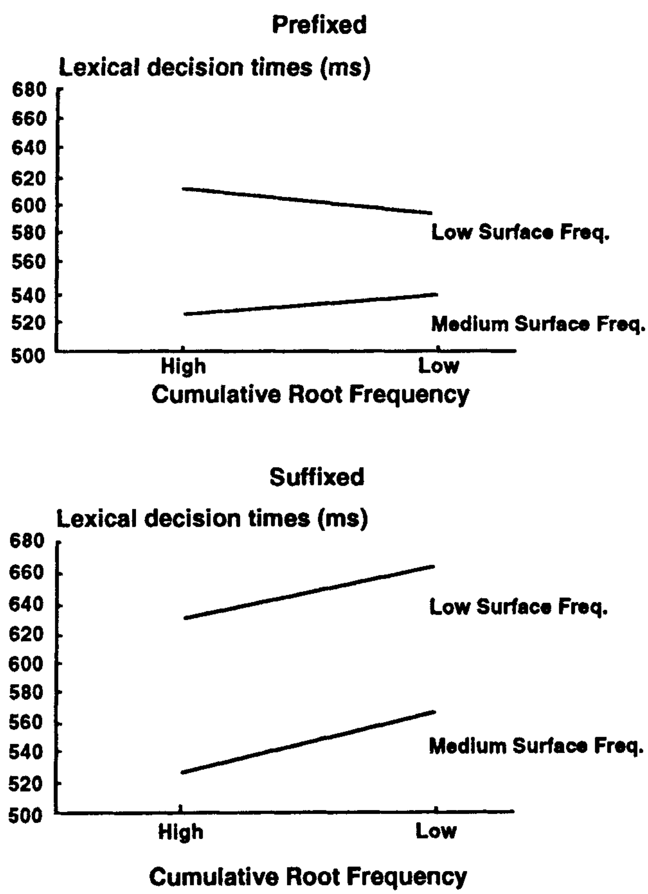

Figure 1. **Mean Lexical Decision Times for Prefixed and Suffixed Words as a function of Surface and Cumulative Root Frequency.**

In a series of studies analogous to those presented above, P. Cole (1987) examined in a detailed way the existence of a correlation between the lexical decision time and the logarithm of the surface and cumulative frequency of suffixed and prefixed derived words. The results obtained confirm the existence of a significant correlation between the reaction times of suffixed words and the logarithm of their cumulative and surface frequency. For prefixed words, alone surface frequency is significantly correlated to reaction times.

Taken altogether these researches demonstrate an important asymmetry in the processing of prefixed and suffixed words and suggest, in particular, that contrary to the hypothesis formulated by Taft, prefixed words are not accessed on the basis of their root but rather on the basis of their whole word form. We assumed that the different sequential organization of the morphemic constituents in affix + root (for prefixed words) and root + affix (for suffixed words) determines an access via the root only for the suffixed words within which it is the beginning part of words. Since for prefixed words the word candidates are selected on the basis of the initial part of the word, the information carried by the root is not exploited in the selection of the appropriate lexical candidate. Consequently, latencies to prefixed words are not affected by their cumulative root frequency. For suffixed words, access by the root at the beginning of the word entails accessing a morphological family of words sharing this root. This process is sensitive to the cumulative root frequency of suffixed words.

Certain recently obtained results in the area of spoken language recognition where the sequential nature of the stimulus presentation is strict lead also to a rejection of the hypothesis of prefix stripping.

Thus, Tyler, Marslen-Wilson, Rentul, and Hanney (1988) showed that the point of recognition of prefixed words is determined by the properties of the whole word form and not by those of the stem as the prefix stripping theory would predict. These authors concluded that if morphemic decomposition of these words exists, this decomposition occurs after access and not as part of the access process itself. More recently, Schriefers, Zwitserlood, and Roeloffs (1991) were also led to reject the hypothesis of a prelexical decomposition of prefixed words in the auditory domain. However, certain of the results obtained by these authors suggest that the identification of these words may depend on the properties of their morphemic components. They propose, therefore, that the lexical entries of derived words contain information concerning not only their whole word form but also their morphemic components and that this information is used during identification. These results suggest that the process of identifying complex spoken words is neither blind to internal structure nor is it mediated by a pre-lexical decomposition of speech input into stems and affixes. Complex words should be processed essentially in a left-to-right manner.

However, as noted above, in visual word recognition the way in which the visual input is mapped onto representation in the mental lexicon is probably not so directly constrained by a left-to-right sequential process. Directionality in time is a fundamental property of the speech signal. Indeed, spoken words are spread along the time axis moving from beginning to end in a way that is not true of the written words. In visual word recognition, there is no strong evidence for a left-to-right mapping of the sensory input onto lexical representations as it applies to the spoken word recognition process where the input is temporally distributed. Such a left-to-right process in visual word recognition should be a consequence of the oculo-motor constraints of the reading process, particularly for long words within which eye fixations are generally distributed from left (or slightly left of center) to right.

In this respect, in Cole et al.'s study (1989), 9- to 11-letter long prefixed and suffixed words were tested and a left-to-right reading process could have determined an access by the root only for suffixed words. Since for long prefixed words the processing of the root does not precede that of the whole word form, the information carried out by the root cannot be exploited in order to select the appropriate lexical candidate.

Another reason for a left-to-right access process could be that the lexical search process operates faster with the beginning part of words. According to this possibility, there should be directionality in lexical representations in that they should be accessed faster from left-to-right than from right-to-left; thus, to determine the membership of the word candidate, the lexical search process should operate faster on all words in the language beginning with an initial sequence rather than with an ending sequence.

To answer this question it is important to evaluate the extent to which directionality may be a property of the lexical representations themselves; such a property would entail a slowing down of the recognition procedures during access by a root located at the end of the word.

Recently, C. Beauvillain (submitted for publication) used a new experimental procedure to examine how the reader makes use of morphemic information early in the visual recognition process of prefixed and suffixed words. By manipulating the luminance of morphemic units within words displayed on a cathode-ray tube, it is possible to modify the speed of transmission of the information in the visual system and, consequently, in the higher levels of processing. The stimulus energy was manipulated by using two types of within-word display: high and low contrast. It was expected that a high contrast on a sublexical unit would facilitate word recognition if the lexical search space is organized in such a way that this unit is an access unit. To limit the effect of left-to-right processing determined by word length, only short (5 to 6 letters) prefixed and suffixed words were used. Preliminary data obtained using the same experimental procedure have shown, in effect, an absence of word beginning superiority for monomorphemic words of this length.

In a first experiment we tested the relevance of the root as an access unit, whether its position is at the beginning (suffixed) or the end (prefixed) of the word. This was carried out by comparing the effect of a high contrast on the root and on the root but one letter on prefixed and suffixed words. For example, the prefixed word "reflux" ("flux," the root, is a word) appeared in two stimulus conditions, "re**flux**" that emphasized the root (the 4-letter high contrast condition) and "ref**lux**" that emphasized the root but one letter (the 3-letter high contrast condition). If the root (that is, "flux" in "reflux") constitutes a processing unit of the word, the latency times should be shorter when the root is highly contrasted than when the root but one letter is highly contrasted. The same comparison was done for pseudo-prefixed words such as "reflet," displayed in the two following conditions: "re**flet**" and "ref**let**," and for monomorphemic words orthographically matched to affixed words and displayed in the two following conditions: "gi**fler**" and "gif**ler**."

In the encoding phase of the experimental task subjects were instructed to read a test word and to press a button once the word was read. Then, a second word appears on which the subject perform a "same" / "different" comparison task (comparison phase). The dependent variable was the time for encoding the test word.

As can be seen in Figure 2 encoding times for affixed words were significantly faster in the 4-letter high contrast condition that emphasized the root than in the 3-letter high contrast condition. Nevertheless, for pseudo-affixed and monomorphemic words this difference was negligible and non significant. Only within-word units that correspond to the root seem to be used by the recognition system. It is important to note that the information conveyed by the root is exploited by the recognition system whether the root is at the beginning (suffixed words) or the end (prefixed words) of the word. Such a result is compatible with a general root access hypothesis for prefixed as for suffixed words.

A second experiment was carried out to test directly this hypothesis in comparing the encoding time for affixed words in which the high-constrast segment corresponds to the root or to the affix. The important point is that for prefixed words, the root access hypothesis predicts that latencies may be shorter when the root is highly contrasted, as in "re**flux**," than when the beginning morpheme, the prefix, is highly contrasted as in "**re**flux."

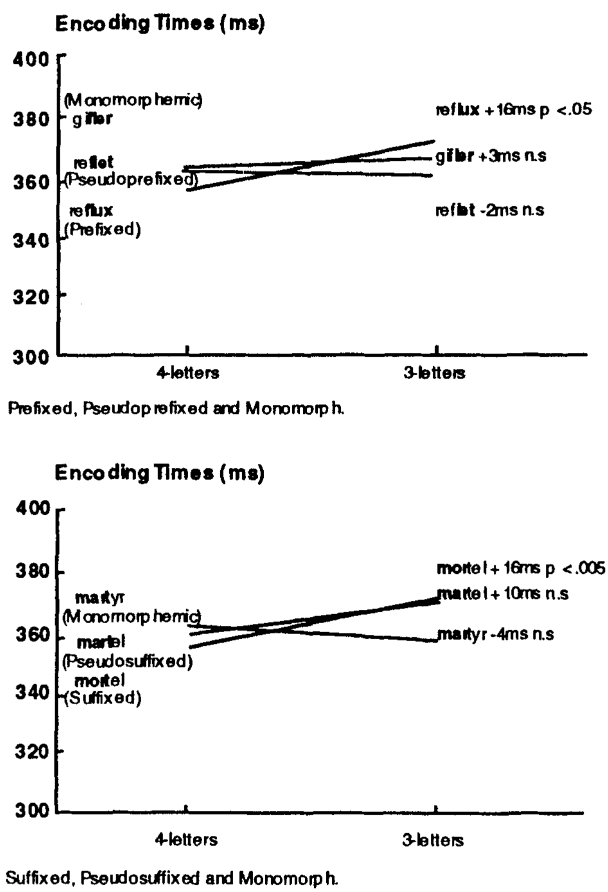

Figure 2. **Encoding Times as a function of the Number of High Contrast Letters and the Type of Word. For affixed words, the 4-letters High Contrast condition corresponds to the root.**

Figure 3 shows that this was not the case. When the prefix is highly contrasted, the performance is significantly better than when the root is highly contrasted. However, as expected, suffixed words showed a significant superiority of their beginning (root) over their ending (suffix). Interestingly, pseudo-affixed words of the same length did not show any significant beginning superiority effect. The simplest explanation for this absence of contrast location effect with pseudo-affixed words is that their beginning and ending segments do not constitute a morphemic unit.

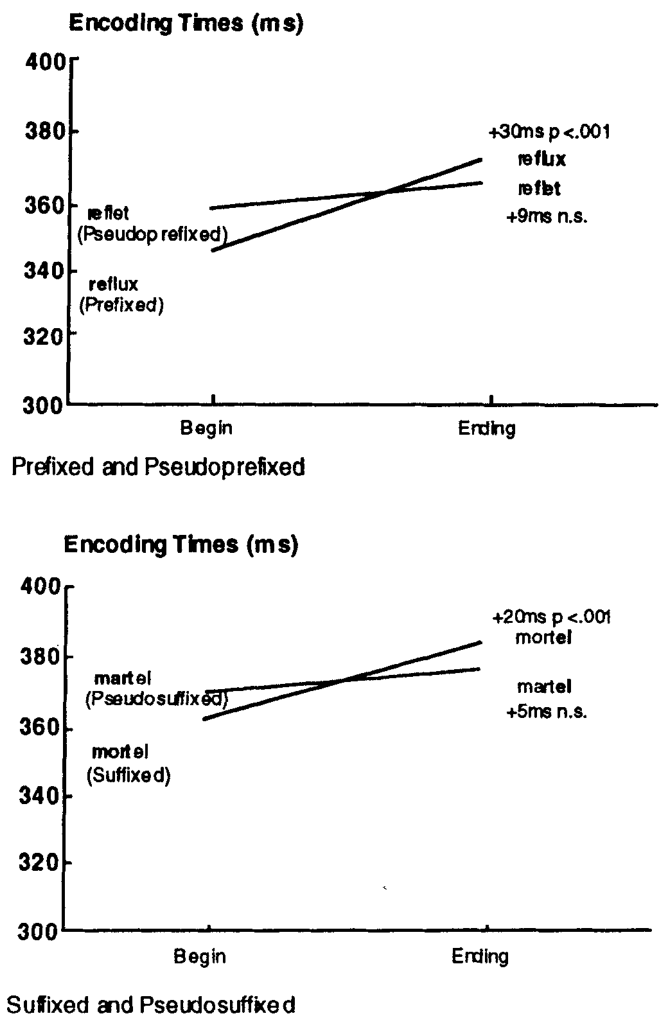

Figure 3. **Encoding Times as a function of the Location of High Contrast Letters (Beginning or Ending) and the Type of words. For affixed words the Beginning High Contrast Letters correspond to the first morphemic constituent (the prefix for prefixed and the root for suffixed words) and the Ending High Contrast Letters correspond to the second morphemic constituent (the root for prefixed and the suffix for suffixed words).**

Overall, these results suggest that when words are displayed in a contrast condition emphasizing a lexical unit such as a root, the dominant effect of the morphemic structure is accompanied by a left-to-right directionality in processing. Then, these recent findings replicate and extend the scope of earlier ones which showed that the exploitation of morphemic units within affixed words depends on a left-to-right process. The possibility should be considered that there is a sequential organization in the lexical representations that imposes a cost on any stimulus that enters the lexicon with the word ending.

Conclusion

This chapter has attempted to provide a review of the experimental evidence collected in our laboratory concerning lexical representations and access procedures of morphologically complex words. The work presented in the first section was formulated from the outset on a "representational" level in order to characterize the mode of representation of morphological information in the mental lexicon.

The position taken has been that lexical entries for derived words possess information concerning their whole word form. Information about the morphological structure of these words is reflected by the organization of the morphological family, i.e., in the way in which the different items sharing the same "head" are related in the affix frame. The results obtained with experiments using priming methodology showed that morphological relations must be distinguished from orthographic relations. Morphological priming effects may not be the result of the conjunction of orthographic and semantic effects.

In the final section, we examined the manner in which affixed words are recognized according to their morphemic structure. Contrary to Taft's hypothesis, none of the data from our experiments provides evidence in favor of an obligatory decomposition procedure. In particular, prefixed are not accessed on the basis of their root but rather on the basis of their whole word form. Thus, the different sequential organization of the morphemic constituents in affix + root (for prefixed words) and root + affix (for suffixed words) determines an access by the root only for suffixed words. Moreover, inducing parsing procedures, we have provided evidence that the morphemic structure of affixed words is accompanied by a left-to-right directionality in processing. Suffixed words are recognized faster when the root is identified before the affix, whereas prefixed words are recognized faster when the affix is identified before the root. We have suggested that such a left-to right directionality in processing should be a consequence of the properties of lexical representations themselves in that they should be accessed faster from left-to-right than from right-to-left. This obliges us to consider it is the properties of lexical representations that determine the nature of the procedures used to address the lexicon.

References

Beauvillain, C. (submitted). Directionality in visual word recognition: Evidence from prefixed and suffixed words.

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words, *Language and Cognitive Processes, 2,* 217-227.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed), *Language production* (Vol. 2). New York/London: Academic Press.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28,* 297-332.

Cole, P. (1987) *Morphologie derivationnelle et acces au lexique.* These de Doctorat, Universite de Paris V (unpublished).

Cole, P., Beauvillain, C., Pavard, B., & Segui, J. (1986). Organisation morphologique et acces au lexique. *L'Annee Psychologique, 86,* 349-365.
Cole, 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.
Cutler, A., Hawkins, J. A., & Gilligan, G. (1985). The suffixing preference: A processing explanation. *Linguistics, 23,* 723-758.
Forster, K. I. (1987). Form-priming with masked primes: The best match hypothesis. In M. Coltheart (Ed.), *Attention and performance XII.* Hillsdale, NJ: Erlbaum.
Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition, 10,* 680-690.
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.
Fowler, C. A., Napps, S. E., & Feldman, L. (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming, *Memory & Cognition, 13,* 241-255.
Grainger, J., Cole, P., & Segui, J. (1991). Masked Morphological Priming in Visual Word Recognition. *Memory and Language, 30,* 370-384.
Hawkins, J. A., & Cutler, A. (1988). Psycholinguistic factors in morphological asymmetry. In J. A. Hawkins (Ed.), *Explaining language universals.* Oxford, Basil Blackwell.
Henderson, L., Wallis, J., & Knight, D. (1984). Morphemic structure and lexical access. In H. Bouma & D. G. Bouwhuis (Eds), *Attention and performance X.* London, Erlbaum.
Henderson, L. (1985). Toward a psychology of morphemes. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol 1). London, Erlbaum.
Lukatela, G., Gligorijević, B., Kostić, A., & Turvey, M.T. (1980). Representation of inflected nouns in the internal lexicon. *Memory & Cognition, 8,* 415-423.
Manelis, L., & Tharp, D. A. (1977). The processing of affixed words. *Memory & Cognition, 5,* 690-695.
Murrell, G. A., & Morton, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology, 102,* 963-968.
Napps, S. E. (1989). Morphemic relationships in the lexicon: Are they distinct from semantic and formal relationships? *Memory & Cognition, 17,* 729-739.
Schriefers, H., Zwitserlood, P., & Roeloffs, A. (1991). The identification of morphologically complex words: Continuous processing or decomposition? *Journal of Memory and Language, 30,* 26-47.
Napps, S. E., Fowler, C. A. (1987). Formal relationships among words and the organization of the mental lexicon. *Journal of Psycholinguistic Research, 16,* 257-272.
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.
Segui, J., & Zubizarreta, M-L. (1985). Mental representation of morphologically complex words and lexical access. *Linguistics, 23,* 759-774.
Taft, M. (1985). The decoding of words in lexical access: A review of the morphographic approach. In D. Besner, T. G. Waller, & G. E. Mackinnon (Eds.), *Reading research: Advances in theory and practice* (Vol 5). London: Academic Press.
Tyler, L. K., Marslen-Wilson, W., Rentoul, J., & Henney, P. (1988). Continuous or discontinuous access in spoken word recognition: The role of derivational prefixes. *Journal of Memory and Language, 27,* 368-381.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)

CHAPTER 20

Bilingual Lexical Representation: A Closer Look at Conceptual Representations

Annette M. B. de Groot
Department of Psychology
University of Amsterdam, The Netherlands

Although usually extremely simple, the figures that illustrate bilingual memory organization in journal articles and book chapters often do a good job of accounting for the data. Yet, when studying them, one cannot help wondering every so often whether, rather than parsimoniously capturing its essence, these few strokes and dashes may do injustice to the complexity of reality. Take, as an example, Figure 1. It is based on my own work on between-language repetition and semantic priming (de Groot & Nas, 1991) and on word translation (de Groot, in press).

As is often done when depicting bilingual memory organization, two representational levels are distinguished. A whole word is represented in a single node at the lexical level; its meaning in a single node at the conceptual level. In other papers (e.g., Chen & Leung, 1989; Potter, So, Von Eckardt, & Feldman, 1984), instead of circles for individual words and their meanings, boxes are drawn to represent whole word and concept systems.

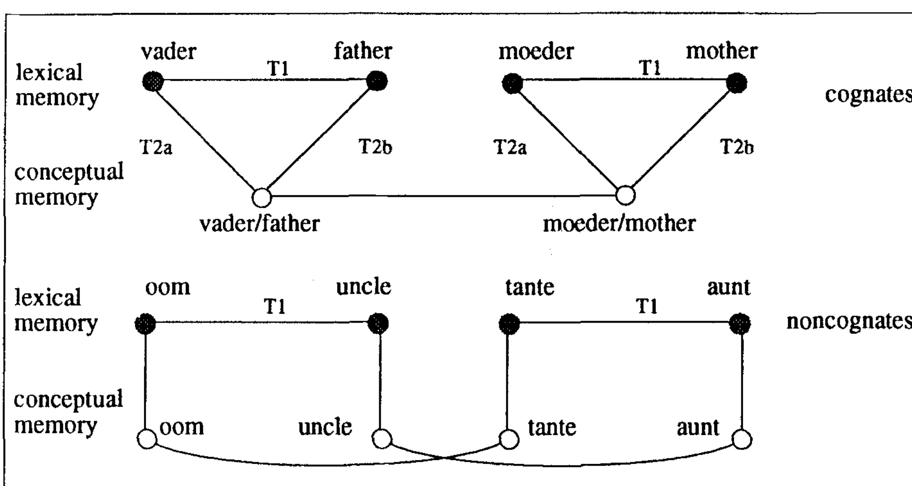

Figure 1. **Some representations in bilingual memory.**

What I primarily intend to do here is to zoom in on some of the circles of Figure 1 to see what can be discerned there. In so doing, the various parts of Figure 1 will be discussed, and standard accounts of a number of results from bilingual investigations will be reviewed. In addition, although they are not backed by new data, alternative explanations will be suggested.

Although their content was never explicated, the circles at the conceptual level were never deliberately intended to suggest indivisible entities. In Quillian's hierarchic network model of semantic memory (e.g., Collins & Quillian, 1969; Quillian, 1968), from which many of the views on monolingual and bilingual representation are derived, concepts were represented in nodes, relations between concepts in links between nodes, and the meaning of a concept by the pattern of relationships in which the concept node participates (see Rumelhart & Norman, 1985). So the meaning of *bird* would consist of 'is a subset of *animal,*' 'has as subset *canary*', 'has as subset *ostrich,*' 'has as parts *feathers,*' 'can fly,' etc. Or in the non-hierarchical successor of this model, the meaning of *red* would consist of its relation with *orange, yellow, green, fire, apples, roses,* etc. (Collins & Loftus, 1975). In the same vein, the concept nodes in Figure 1 can be seen as built up from a number of meaning elements. This is made explicit in Figure 2. The concept associated with the word *vader (father)* in Figure 1 is now spread out over six nodes, each of them representing one meaning element of the word *vader*. The number six is chosen arbitrarily. I will henceforth call these conceptual representations 'distributed' (see e.g., Hinton, McClelland, & Rumelhart, 1986). Instead of there being just one connection from the lexical node for *vader* to its conceptual node (Figure 1), the lexical node now has connections to each of the meaning elements of the conceptual representation. Upon presentation of the word *vader*, each of these elements receives excitatory activation via its connection with the lexical node.

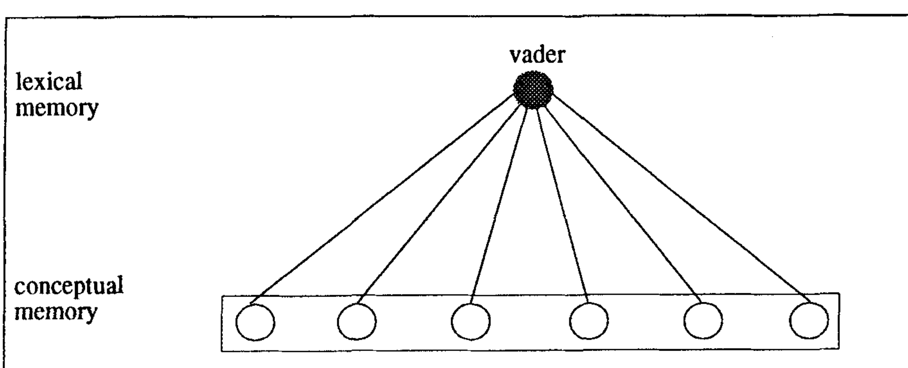

Figure 2. **A distributed conceptual representation in memory.**

In the next parts of this chapter I will gradually expand Figure 2 with the purpose of providing new accounts of some of the results obtained in a number of bilingual (and occasionally monolingual) processing tasks (the essence of some of these accounts was also suggested by Taylor and Taylor (1990). Only semantic memory tasks will be considered, that is, tasks that *could* be performed without consulting episodic knowledge (even though performance may well be influenced by such knowledge). The tasks being discussed are: word translation, primed lexical decision (that is, intra- and interlingual semantic priming and repetition priming, with lexical decision serving as the experimental task), word association, and semantic-relation assessment.

Word translation

In a number of studies the word translation task has been used as a means of obtaining information on the organization of knowledge in bilingual memory (e.g., Chen & Leung, 1989; de Groot, in press; Kroll & Curley, 1988; Potter et al., 1984). In its standard form the task simply involves presenting the bilingual subjects with words in one language, and asking them to produce their translation in a second language. In all but one of these studies (de Groot, in press) the task was used in conjunction with one or more other tasks, most often picture naming in a second language. A comparison of response times in word translation and picture naming in the second language was meant to solve the question of whether word translation takes place via a direct connection between the lexical representations of the translation equivalents (Route T1 in Figure 1), or indirectly, via an amodal conceptual representation shared by the two translation equivalents as well as by a picture of the referent of these words (Route T2 [T2a + T2b] in Figure 1; the node for the picture is not shown). If translation comes about by tracing T1, it is argued, translation should take less time than picture naming in the second language, because the route to the response would be shorter than in picture naming (in which access of the conceptual node cannot be circumvented). But if Route T2 is traced in translation, word translation and picture naming should take equally long. In the case of the latter outcome one might want to conclude, as Potter et al. (1984) did, that no direct connections exist between the representations of translation equivalents at the lexical representational level.

Fluent bilinguals turn out to be as fast in second-language picture naming as in word translation (Chen & Leung, 1989; Kroll & Curley, 1988; Potter et al., 1984), but less proficient bilinguals (or, more precisely, less proficient *adult* bilinguals; Chen & Leung, 1989) translate faster than they name pictures in their second language (Chen & Leung, 1989; Kroll & Curley, 1988). The data thus indicate that fluent bilinguals use Route T2, whereas less proficient adult bilinguals take Route T1. This suggests that T1-connections do exist, but are bypassed by fluent bilinguals during word translation. From a study comparing naming in a native (Dutch) and a second (English) language, on the one hand, with translation between the two languages on the other, Kroll and Stewart (1990) indeed drew the conclusion that T1-connections exist (see also de Groot & Nas, 1991). At the same time they qualified this conclusion: There are T1-connections in both directions, from the stronger to the weaker language and vice versa, but they differ in strength. The link from the weaker (here English) to the stronger language (here Dutch) is the stronger of the two.

The process involved in what was called 'tracing' translation routes above is presumably 'spreading activation': When a stimulus word is presented, it first contacts its representation in lexical memory. The activation that originates in this memory node spreads out along the paths of the memory network and activates the representations it encounters en route. Consequently, the activation levels of the encountered representations are temporarily increased, and the corresponding words are readily available as responses. Which of these activated representations is eventually selected for responding will depend on the task at hand (and on the extent to which each single one of them is activated). In word translation by fluent bilinguals, the representation selected for responding will typically be that of the stimulus word's translation in the target language.

How can word translation by fluent bilinguals be depicted in light of the representations illustrated in Figure 2? Figure 3a shows how this can be done. It repeats Figure 2, but now a lexical node for the English translation equivalent of *vader* is added. Additionally, links between this representation and each of the meaning elements constituting the conceptual representation are shown. The implicit assumption in this example, right or wrong, is that the Dutch word and its English translation have exactly the same meaning. Translation again involves the tracing of links (spreading activation) from the lexical representation of a word in one language to that of this word in the second language via conceptual memory, but now the links to be traced are those connecting the lexical nodes with the individual meaning elements of the conceptual representation.

The meanings of translations often do not fully overlap. This may more often be the case with some types of words than with others. For instance, it has been suggested that the meanings of abstract words differ more across languages than those of concrete words (Taylor, 1976). Such a state of affairs can be captured in a very straightforward way by representational structures of the kind depicted in Figure 3a. Figure 3b shows a situation wherein an abstract word in Dutch shares four meaning elements with the corresponding word in English. In addition to the common meaning elements, each word has two not shared by the corresponding word in the other language. Thus, the conceptual representations now overlap only partly.

It is reasonable to assume that the situations depicted in Figures 3a and 3b lead to differences in translation performance: The more conceptual elements shared by a pair of translation equivalents, the more activation will spread from the lexical node of a word to that of its translation, and the better performance—as assessed by response speed, number of errors, and number of 'omissions' (where no translation is given)—will be. Staying with our examples, other things being equal, the concrete word *vader* should be translated faster and/or more often and/or more often correctly than the abstract word *idee.*

I indeed observed these differential effects for concrete and abstract words in a recent study on word translation (de Groot, in press), although I argued there that another word characteristic than word concreteness, namely, 'context availability' (see e.g., Schwanenflugel, Harnishfeger, & Stowe, 1988) may well underlie them (but this does not undermine the point I am trying to make here; I could have chosen another word characteristic than word concreteness to illustrate it). In this same study I also obtained differential effects for cognates (words with orthographically and phonologically similar translations) and noncognates (dissimilar translations). Performance was consistently better for the former type of words.

In sum, it appears that a memory with distributed conceptual representations could quite naturally cope with differential translation performance on different types of words.

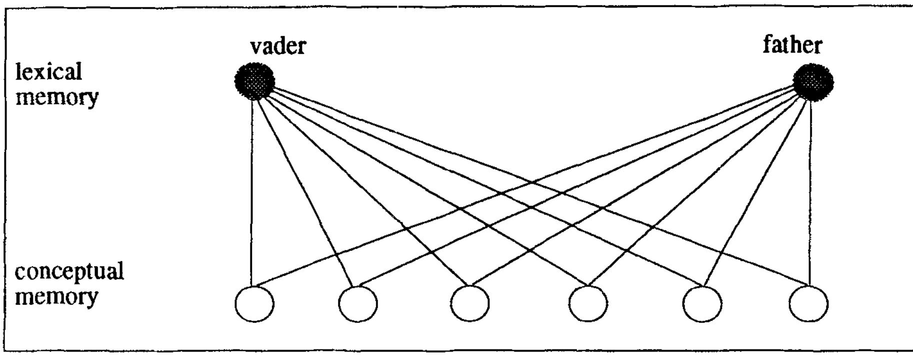

Figure 3a. **A distributed conceptual representation in memory. Translations have exactly the same meaning.**

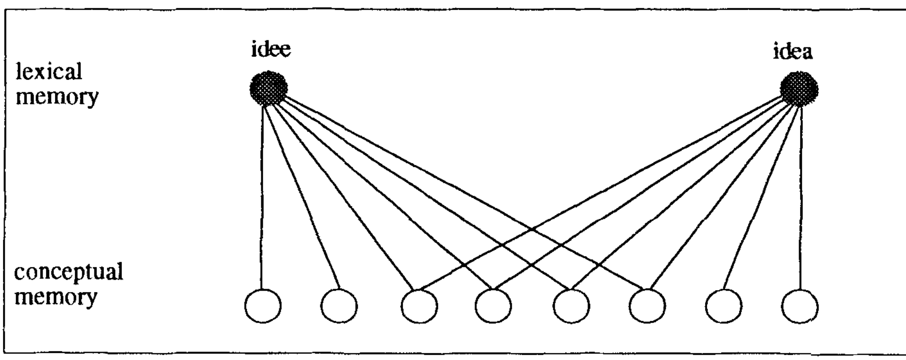

Figure 3b. **Distributed conceptual representations in memory. Translations differ in meaning.**

Note, however, that I am not suggesting that effects of *all* word characteristics on translation performance could be explained this way. For instance, word frequency effects (de Groot, in press) may be due to differences in the strength of connections (e.g., between the lexical and conceptual nodes in Figure 1, or between the lexical nodes and the nodes for the various meaning elements in Figures 3a and 3b) between words of high and low frequency, rather than to different numbers of shared meaning elements between the translations. The stronger the links, the more activation they transmit.

Errors in word translation

Words share aspects of meaning not only with their translations, but also with semantically related words in the same language. They may also share parts of their meaning with the translations of these semantically related words. Figures 4a and 4b serve to illustrate this point. At the risk of being too explicit I add that the amount of semantic overlap between the various words suggested in these figures is not based on empirical findings, but is contrived. At most, some sophisticated guessing is involved in places.

In Figure 4a, the words *vader* and *father* again (cf. Figure 3a) share all their meaning elements. The same is the case with the words *moeder* and its translation *mother*. In addition, the words *vader* and *moeder* have three elements in common. But because both *vader* and *moeder* share all their conceptual elements with their respective translations, *vader* also has the same three elements in common with *mother*, and, conversely, *father* has the same three in common with *moeder*.

Figure 4b depicts a situation in which Figure 3b is expanded. The words within both pairs of translations share four meaning elements. The semantically related words share two meaning elements both within and between languages.

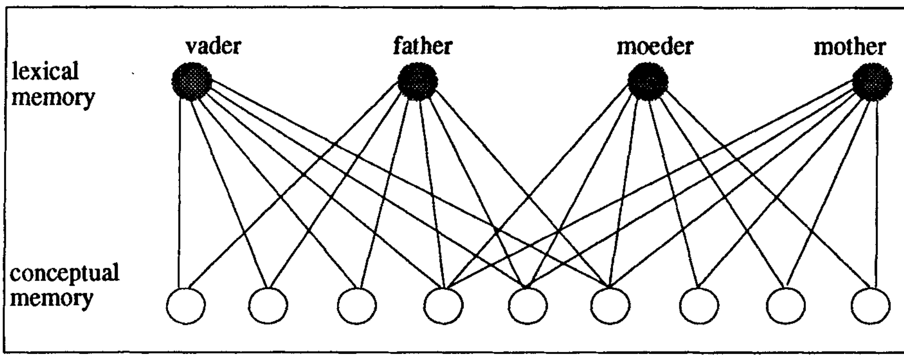

Figure 4a. **Distributed conceptual representations in memory. Translations share all meaning elements. Semantically related words share a few, both within and between languages.**

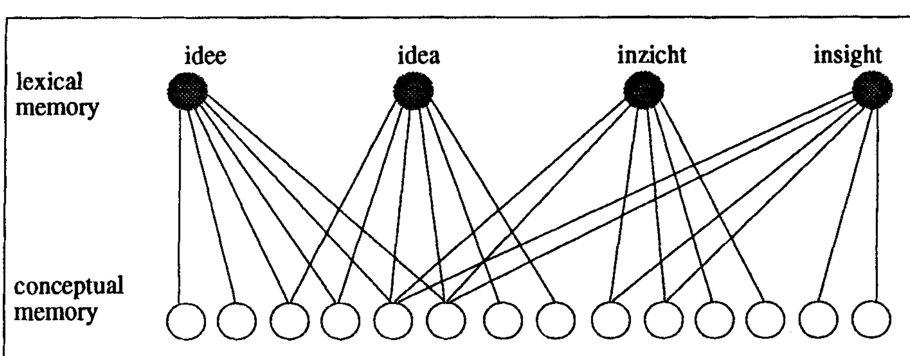

Figure 4b. **Distributed conceptual representations in memory. Translations differ in meaning. Semantically related words share a few meaning elements, both within and between languages.**

The representations depicted in Figures 4a and 4b could account for a very common translation error, namely, the production of a word semantically related to the stimulus word, but in the language of the translation. My work on word translation has abundantly provided me with such errors. Here are a few examples (the first word is the Dutch stimulus word to be translated. Its translation is given in between brackets. The last word is the response actually given): kaars (candle)—flame; speld (pin)—needle; cirkel (circle)—square; boomgaard (orchard)—vineyard; handdoek (towel)—blanket; bruid (bride)—wedding; bliksem (lightning)—thunder; aardappel (potato)—carrot; plafond (ceiling)—roof. These errors can be understood as arising from the activation of the lexical node of the response word (e.g., square) via the nodes representing the meaning elements it shares with the stimulus word (cirkel). The more elements shared between a stimulus word and a word in the other language that is not its translation, the larger the activation in the lexical node for the latter will be, and hence the larger the chance that the associated error will occur.

Translation recognition

In my work on word translation I explored a new version of this task, which I labeled 'translation recognition.' Bilingual subjects performing this task are presented with pairs of words, each consisting of a word in their first language and one in their second language. The task is to decide whether or not the words within each pair are translations of each other. In terms of Figure 1, translation recognition may involve the tracing of the links (a spread of activation) departing from the lexical nodes of the two presented words. Both the direct links at the lexical level and, if present, the indirect ones via the conceptual level, will be traced. If an intersection of activation occurs anywhere, a *yes* response is emitted. If not, the subject responds *no.* On negative trials (requiring a *no* response) the subjects in my experiment were always presented words that were not related in any obvious sense (phonologically, orthographically, or semantically). The searches starting from the lexical nodes of the presented words will thus intersect nowhere, and a *no* response will be given. According to the distributed view of representation discussed here, translation recognition again entails activation spreading from the lexical nodes of the words involved and the detection of intersecting activation waves, but now the routes involved are the links between the lexical nodes and the representations of the individual meaning elements at the conceptual level.

From the present conception of representation a prediction can be derived concerning an experimental condition I have not tested yet, namely one in which the negative trials consist of semantically related words (e.g., the word pair *vader-mother*). Because an intersection of activation will occur on one or more links (three in the example), the subjects will be biased towards a *yes* response. Consequently, the rejection of these words as a pair of translation equivalents should take relatively long, and relatively many errors should occur.

More tasks and more effects are hidden in Figures 4a and 4b, for instance, semantic priming effects within and between languages. They are the topic of the following section.

Semantic priming within and between languages

A robust effect in monolingual investigations is that words in, for instance, a lexical decision experiment are responded to faster when they follow a semantically related word (e.g., context stimulus or 'prime': *love;* test stimulus: *friendship*) than when preceded by a semantically unrelated word or some neutral context stimulus (Meyer & Schvaneveldt, 1971; see Neely, 1991, for a review). This 'semantic-priming' effect is often, again, attributed to activation spreading between memory nodes, for example, in a memory system of the type illustrated in Figure 1: If a word is presented that corresponds to one of the representations preactivated through activation spreading from the representation of an earlier prime, it is recognized, and hence responded to, relatively fast. In this view, recognition is effectuated as soon as the activation in the test word's lexical node exceeds a critical threshold value. In Figure 1, when *vader* is the prime and *moeder* the test word, responding to the latter comes about relatively fast because the lexical node for *moeder* has received preactivation from the lexical node of *vader* via the conceptual nodes for *vader* and *moeder*.

Priming effects of words presented in one language on semantically related words in a second language can be explained in the same way. If translation equivalents share a conceptual representation, when *vader* is presented as prime the lexical node for the test word *mother* should also receive preactivation, via the conceptual nodes *vader/father* and *moeder/mother*. In fact, under the scheme of the top half of Figure 1 (and given equally strong T2a and T2b connections), the lexical node *mother* should receive the same amount of preactivation as the node for *moeder*, and the between-language priming effect should thus be as large as the within-language effect. The bottom half of Figure 1 depicts a situation wherein within-language semantic priming should occur, but no between-language priming.

A number of studies have shown that a semantic-priming effect can occur between languages (Chen & Ng, 1989; de Groot & Nas, 1991; Jin & Fischler, 1987; Kerkman, 1984; Kirsner, Smith, Lockhart, King & Jain, 1984; Meyer & Ruddy, 1974; Schwanenflugel & Rey, 1986; Tzelgov & Henik, 1989). Some of them (Chen & Ng, 1989; Meyer & Ruddy, 1974; Schwanenflugel & Rey, 1986; Tzelgov & Henik, 1989) suggested that the interlingual effect is as large as the intralingual effect. Two studies indicated that the interlingual effect may be word-type specific. Jin and Fischler (1987) observed a semantic-priming effect across languages for concrete words but not for abstract words. Under conditions in which the prime was degraded so that it could not be identified by the subjects, Gerard Nas and I (de Groot & Nas, 1991) obtained a between-language effect for cognates but not for noncognates. This is in fact why in the bottom part of Figure 1, which shows the representation of noncognates, separate conceptual nodes for translation equivalents, and only within-language connections between semantically-related words at the conceptual representational level, are depicted. The assumption of separate representations directly follows from the absence of an interlingual effect for noncognates. Had there been a shared representation, such an effect should have occurred.

Figures 4a and 4b show how semantic priming, both within and between languages, comes about when conceptual representations are distributed across a number of nodes (see also Taylor & Taylor, 1990, and, for monolingual semantic priming, Masson, 1991). By the time the test word is presented, activation has accumulated in its lexical node via the representations of the meaning elements it has in common with the prime. The more elements the prime and test words share, the larger the amount of preactivation in the test word's lexical node, and the larger the priming effect should be. Thus, the effect should be larger in Figure 4a than in Figure 4b (three and two elements shared by the semantically related words, respectively). Differences in the number of conceptual elements that the prime and the test word have in common could explain the finding in the monolingual literature on semantic priming (e.g., de Groot, Thomassen & Hudson, 1982; Fischler & Goodman, 1978) that the size of the priming effect depends upon the 'strength' of the semantic relation between the prime and the test words. Of course, a representational structure of the kind depicted in Figure 1 could also explain such an effect, namely by assuming that the links between the conceptual representations of semantically related words differ in strength, the strength of each individual link reflecting the strength of the semantic relation between the two words represented in the nodes on both ends of the link. The stronger the link, the more activation it will transmit, and the larger the effect will be.

When one compares one of the accounts of word translation in terms of the representations depicted in Figure 1 (Route T2), on the one hand, with that in terms of the representations in Figures 4a and 4b, on the other, one might want to argue that they are in fact conceptually the same. But the two conceptions of semantic priming differ essentially. Given the representations in Figure 1, for semantic priming to arise, activation in one conceptual representation (of the prime word) has to traverse a link from this representation to another (of the test stimulus) in conceptual memory. According to the view depicted in Figures 4a and 4b, no such links between conceptual representations have to be traversed. They may even not exist. All links responsible for the effect directly connect nodes in lexical memory with nodes in conceptual memory. The priming effect is due to the fact that the prime, by activating its distributed conceptual representation, at the same time activates part of the conceptual representation of the test word.

Representations as in Figures 4a and 4b could also readily account for differences in the size of the semantic-priming effect within and between languages. I will not consider the hypothetical situation where larger between- than within-language effects are obtained. I do not know of any study in which such a finding was obtained, and it seems intuitively implausible. But priming effects may be smaller between than within languages. Models with non-distributed conceptual representations could explain such results in two ways: 1) There are between-language connections between the representations of semantically related words in conceptual memory for all words for which the corresponding within-language connections exist, but they are weaker than the latter. This option demands the existence of language-specific conceptual nodes. (In the case of shared representations, the within- and between-language connections between two nodes in conceptual memory would in fact be one and the same connection. It is hard to see how this one connection could be strong and weak at the same time.) Thus, for example, there would exist a conceptual node for *liefde* and one for its translation *love;* one for *vriendschap* and one for its translation *friendship*. Additionally, there would be relatively strong links between the nodes for *liefde* and *vriendschap,* and between those for *love* and *friendship,* and there would be weaker links between the nodes for *liefde* and *friendship,* and between those for *love* and *vriendschap*. 2) The between-language connections between the representations of semantically related words in conceptual memory are as strong as the corresponding within-language connections, but they do not exist for all of the semantically related words that are connected *within* a language. The situation depicted in Figure 1 is one way to instantiate this second option. Here a subset of the translation equivalents in the bilingual lexicon (cognates) shares a conceptual representation. The reason the between-language connections are as strong as the corresponding within-language connections is that they are in fact the same connections. The remaining words (noncognates) are represented in language-specific conceptual nodes that are only connected to conceptual representations of semantically related words of the same language. But also compatible with Option 2 would be a situation in which *all* translation equivalents would be represented in language-specific nodes. So there would be, for instance, two word quartets, *vader-father-moeder-mother* (cognate translations), and *oom-uncle-tante-aunt* (noncognate translations). Each of these eight words would individually be represented in a conceptual node. All intralingual connections between the conceptual representations of semantically related words, irrespective of the cognate status of the words involved, could be equally strong (*vader-moeder*; *father-mother*; *oom-tante*, and *uncle-aunt*). The interlingual connections between the representations of semantically related cognates (*vader-mother and father-moeder*) could also be this strong, but no connections would exist between *oom* and *aunt*, or between *uncle* and *tante* (noncognates). Options 1 and 2 could be distinguished by item analyses on the data, because the interlingual priming effect should be significant by items if Option 1 were true (all or the majority of the interlingually semantically related word pairs would show the effect), but not if Option 2 were true (only a subset of these word pairs would show the effect).

These solutions are relatively complex and may even appear contrived. They are certainly more complex than the one the distributed view has to offer: It does not seem far-fetched to assume that semantically related words of the same language often share more meaning elements than semantically related words of different languages. The larger the overlap, the more activation will accumulate in the lexical node of the test word, and the larger the priming effect will be. Hence, the effect will be larger within a language than between languages.

In this framework, when for particular types of words (abstract words, Jin & Fischler, 1987; noncognates, de Groot & Nas, 1991) a within-language but no between-language priming effect is obtained, one is not compelled to conclude that the conceptual representations of those words are strictly separated by language, as one is when conceptual representations are regarded as indivisible entities. The translations of such words may still have a large part of their conceptual representations in common, but these words would not share any of their conceptual elements with semantically related words in the other language.

Word association

Word association has also been used as a tool to investigate bilingual memory (Kolers, 1963; Taylor, 1971; 1976). There are two common versions of this task: discrete word association and continued word association. In the former the associative response to a stimulus word has to consist of a single word that is the first word that comes to the subject's mind when reading or hearing the stimulus word. In the latter version, the subject generates as many word associates to the stimulus word as possible within a prespecified amount of time (often 30 or 60 seconds). In bilingual word-association studies stimulus words are typically presented in one or both of the bilingual's two languages, and responses have to be given either in the language of the stimulus word, or in the other language. The issue at stake is to what extent the responses in the various experimental conditions are or are not the same (responses that are translations of those given in other conditions are considered 'same' responses). Same responses are regarded as supporting the view of conceptual representations being shared between languages. Different responses are seen as evidencing language-specific conceptual representations.

Kolers (1963) collected discrete word associations within and between languages. His subjects all had English as their second language, and German, Spanish, or Thai as their native language. Each individual subject produced associations in each of the four within- and between-language conditions. His main finding was that within all three groups of bilinguals over half of all responses in the cross-language conditions were unique, that is, not the same as or a translation of the response word this particular subject gave in either one of the two within-language conditions. He concluded that 'experiences and memories of various kinds are not stored in common in some supralinguistic form but are tagged and stored separately in the language S used to define the experience to himself' (Kolers, 1963, p. 300). This conclusion may be too strong, given the fact that at least a number of responses were shared between languages: On average, just over 20% of the responses of an individual subject were the same as or a translation of those she or he produced in all remaining conditions (e.g., *king-queen; king-reina; rey-reina; rey-queen.* Examples are taken from Kolers, 1963). Furthermore, about 30% of the interlingual responses were the same as or a translation of the response word this subject gave in either the native or nonnative intralingual condition (e.g., *boy-girl; boy-nina; muchacho-hombre; muchacho-trousers.* In this example the subject's response in the nonnative-to-native condition *[boy-nina]* was the same as his response in the nonnative-to-nonnative condition *[boy-girl]*). Note that in the examples above the response words are always in some sense semantically related to the stimulus words. This reflects a fact that always immediately strikes any student of word association: Although the task instructions never explicitly demand this, by far the majority of word association responses indeed are words semantically related to the corresponding stimulus words.

A further interesting finding of Kolers (1963) is that concrete words more often generated the same responses within and across languages than abstract words and emotion words did. The former result was also obtained by Taylor (1976), who tested French-English bilinguals in intra- and interlingual continued word association. Additionally, she observed that stimulus words with cognate translations more often gave rise to the same response words in the intra- and interlingual conditions than noncognates did.

Before drawing conclusions on the basis of these data on the organization of bilingual knowledge in memory, one would first want to know about the chances that a subject will respond with the same word when associating to a word twice within the same language. But for the time being the data suggest, first, that words and their translations in bilingual memory neither fully share their conceptual representations nor are represented in a totally segregated way, and, second, that the degree of separation between languages varies with word type.

Like word translation and semantic priming discussed before, the word association task, both within and across languages, can also be detected in Figures 4a and 4b. In within-language association, viewed in terms of the memory structures suggested here, the same paths are involved as in within-language semantic priming (recall that the response words in word association are typically semantically related to the stimulus words). In between-language association either a laborious process may take place, or a simpler one. Kolers (1963) assumed that in cross-language word association, bilinguals either first translate the stimulus word and then associate to the translation, or they first associate to the stimulus word in the language of the stimulus and subsequently translate the association. Both of these indirect routes are visible in Figures 4a and 4b (e.g., from *vader* via conceptual memory to *father*, and from there, again via conceptual memory, to *mother;* or from *vader* to *moeder* to *mother,* again both via conceptual memory), but a direct route (as direct as the within-language word association route in terms of this type of representation would be) can also be discerned, from *vader* into conceptual memory and from there straight to the lexical node for *mother* (cf. interlingual semantic priming).

On any trial a number of lexical nodes for words that would all constitute appropriate responses will be activated. The one activated most will generally determine the response. So if, in an intralingual condition, after presentation of the stimulus *father* the lexical node for *mother* receives more activation than the nodes for any of the other same-language words plausibly being activated (for instance, *son* and *child*), the corresponding word *mother* will be produced as response. If in an interlingual condition following the presentation of this same stimulus word *father* the lexical node for *moeder* is activated more than any of the other lexical nodes of words in the target language, this node will determine the response. In this situation, the within- and between-language responses will thus be the same. But if the lexical node *zoon (son)* is activated more than is the node for *moeder*, for instance, because the conceptual representation of *zoon* shares more elements with that of *father* than the conceptual representation of *moeder* does, the intra- and interlingual conditions will give rise to different responses. The association data suggest that for some types of words (concrete words; cognates) the maximum activation in the intra- and interlingual conditions relatively often (as compared to abstract words, noncognates, and emotion words) occurs in the lexical nodes of translation equivalents.

The tasks discussed so far may be classified into three groups: production tasks, priming tasks, and relation-assessment tasks. In the *production* tasks a stimulus word is presented from which the subject has to generate a particular type of response. All such tasks implicit in Figures 4a and 4b have been explored so far. These tasks were: word translation (in Figure 4a, from *vader* to *father* and vice versa, and from *moeder* to *mother* and vice versa); within-language word association (from *vader* to *moeder* and vice versa, and from *father* to *mother* and vice versa); and between-language word association (from *vader* to *mother* and vice versa, and from *father* to *moeder* and vice versa).

In the *priming* tasks the subjects have to respond to target stimuli preceded by a prime. The required response could be lexical decision, but other responses may be requested instead (e.g., pronouncing the targets, or performing some semantic classification of them). The influence of the prime on target processing is assessed. Unlike in the relation-assessment tasks to be discussed below, the prime may be ignored by the subjects. The tasks of this type hidden in Figures 4a and 4b are: intralingual semantic priming (prime: *vader*, target: *moeder*, and vice versa, and prime: *father*, target: *mother*, and vice versa), interlingual semantic priming (prime: *vader*, target: *mother*, and vice versa, and prime: *father*, target: *moeder*, and vice versa), and interlingual repetition priming (or 'translation' priming; prime: *vader*, target: *father*, and vice versa, and prime: *moeder*, target: *mother*, and vice versa). The first two of these have already been discussed, but translation priming has been ignored so far. It is the topic of the next section. A characterization of the third group of tasks, the *relation-assessment* tasks, is postponed until later, when a few examples of this class of tasks will be discussed.

Translation priming

Translation priming or between-language repetition priming has been looked at in a large number of studies (Altarriba, 1992; Chen & Ng, 1989; Cristoffanini, Kirsner, & Milech, 1986; de Groot & Nas, 1991; Gerard & Scarborough, 1989; Jin & Fischler, 1987; Kerkman, 1984; Kirsner, Brown, Abrol, Chadha, & Sharma, 1980; Kirsner et al., 1984;

Scarborough, Gerard & Cortese, 1984). In all but four of them the 'classical' interlingual repetition-priming paradigm has been used. In this paradigm, the inter-stimulus-interval between a word and its translation is typically long, several minutes or more, and the subjects produce some response to both the word and its translation. In the four remaining investigations (Altarriba; Chen & Ng; de Groot & Nas; Jin & Fischler), as in studies on semantic priming, a word and its translation (or some other test stimulus) followed one another immediately (across the studies, the stimulus-onset-asynchrony between prime and test stimulus varied between 60 ms and 1000 ms), and the subjects only responded to the latter. In all four of these studies translation priming occurred. The effect occurred not only when the prime was clearly visible (Altarriba, 1992; Chen & Ng, 1989; Jin & Fischler, 1987), but also when it was masked so that it could not be identified by the subjects (de Groot & Nas, 1991). We thought masking the prime to be relevant because, when the experimental task is lexical decision (true for all four studies) and when both the prime word and the test stimulus are clearly visible, a post-lexical integration process may also cause a priming effect. This post-lexical process searches for any relation, for instance, a translation relation, between prime and test stimulus. If it finds one before the subject executes his or her response to the test stimulus, it speeds up this response (see de Groot & Nas, 1991, for details). Whenever the primes are not masked it is thus not clear to what extent the effect may be attributed to the actual priming process.

In de Groot and Nas (1991, Experiments 3 and 4) the cognate status of the translation equivalents was varied. Considering the masked-prime condition only, the effect was always larger for cognates than for noncognates (in one condition the difference in effect size was substantial: 53 ms), although statistically the effect was always equally large for the two types of words. The language combination studied by Jin and Fischler (1987) was Korean-English; that studied by Chen and Ng (1989) was Chinese-English. Unlike in the English script, the units in both Chinese and Korean script are characters. Consequently, Korean-English and Chinese-English translations will always be orthographically dissimilar. They will also generally be distinct phonologically (except that words imported from English into Chinese and Korean or vice versa may retain aspects of the pronunciation of the imported words). In short, the stimulus materials of Jin and Fischler and of Chen and Ng consisted of noncognates. Altarriba's (1992) subjects were Spanish-English bilinguals. The languages involved are both alphabetic, but belong to different language families (Romance and Germanic, respectively). Therefore, her translations probably also consisted primarily of noncognates. Despite the use of noncognates as stimulus materials, translation priming was obtained in all three studies.

In the studies using the classical paradigm, the interlingual effect is less robust, but there is a pattern: Translation priming occurs for cognates (Cristoffanini et al., 1984; Gerard & Scarborough, 1989; Kerkman, 1984), but not for noncognates (Kirsner et al., 1980; Kirsner et al., 1984; Scarborough et al., 1984). However, there are grounds to doubt that the effect under the conditions of these experiments is attributable to spreading activation in bilingual lexical memory, which is our concern here. Instead, it may be an episodic effect (see de Groot & Nas, 1991, for a discussion).

If the representations in Figure 1 are the building blocks of bilingual memory, translation priming for cognates (at least in studies where the non-classical paradigm, the one modeled on semantic-priming studies, is used) could come about through activation spreading directly, via Route T1, or indirectly, via Route T2, from the lexical representation of the prime word to that of the test word, preactivating it prior to its presentation (cf. word translation). For noncognates preactivation could only come about via Route T1, because no indirect connections via conceptual memory exist. Recall that we (de Groot & Nas, 1991) assumed language-specific conceptual representations for noncognate translations because for noncognates no interlingual semantic-priming effect was obtained. However, a translation-priming effect *did* occur. The combination of these two findings forced us to conclude that direct links exist between the lexical nodes of translation equivalents. If indeed no indirect connections between these translations via conceptual memory exist, how else could translation priming for noncognates be explained?

Unlike the view of representation illustrated in Figure 1, the present view does not require the conclusion that direct (T1) connections exist between the lexical representations of translations. They may exist (indeed others have proposed their existence for different reasons; see the section on word translation), but they do not have to. The data summarized above can no longer be regarded as conclusive about this. If the conceptual representation is divided over a number of different nodes, it is perfectly plausible that for a particular type of word (presently noncognates) translation priming occurs, and does so via conceptual memory, whereas at the same time no interlingual semantic-priming effect for this type of words comes about. What would be required is (at least partially) overlapping conceptual representations for a pair of noncognate translations, while at the same time none of the nodes representing the various meaning elements in these conceptual representations is linked to the lexical node of the relevant target word in an interlingual semantic-priming condition.

Relation assessment

One of the tasks discussed so far, translation recognition, may be considered an instance of a class of tasks in which the subjects have to decide whether or not a particular relation between two stimuli exists. These tasks necessarily involve the processing of both stimuli on a trial. In this respect they differ from the above priming tasks (excluding the 'classical' repetition-priming studies), in which the subjects may ignore the first stimulus within each pair of stimuli. Other instances of this group of tasks implicit in Figures 4a and 4b would be intra- and interlingual semantic-relation-assessment tasks, which would require subjects to categorize word pairs according to the presence or absence of *any* semantic relation between the words in these pairs. If such relation is detected, as with the pairs *vader-moeder* (intralingually) and *vader-mother* (interlingually), the subject should respond *yes*. If not (*vader-boom*, or *vader-tree*), *no* should be the response.

Analogous to the conception of translation recognition, semantic-relation assessment may be conceived of as involving activation spreading from the lexical nodes of the two presented words. If an intersection occurs, a *yes* response can be emitted. If not, a *no* response may be executed. I do not know of any study in which it is the subjects' task to categorize the presented word pairs on the presence or absence of *any* semantic relation between the words of a pair, but this hypothetical task is strongly reminiscent of the more specific 'semantic-verification' task that has been used in a very large number of studies (e.g., Collins & Quillian, 1969; Smith, Shoben, & Rips, 1974). In semantic verification as well a relation between the two words on a trial has to be discovered, but the relation to be detected has to be of a specific kind. Other than in the above task, if the words on a trial are semantically related, but not in the prespecified way, such a trial demands a *no* response. In one study of this type (Caramazza & Brones, 1980), semantic verification was investigated both intra- and interlingually.

Caramazza and Brones presented word pairs on a screen, the first word referring to a semantic category, and the second to an instance of this or another category. Subjects had to press one key if the second word belonged to the category referred to by the first, and to press another key if such was not the case. Three categories and six instances of each of them constituted the experimental materials. The categories were 'furniture,' 'fruit', and 'vegetables.' Hence, all stimuli were concrete words. The category and instance names were in the same or in different (English and Spanish) languages. The finding most relevant here was that response time was not influenced by whether or not the names of category and instance were in the same language. Two robust findings in semantic verification studies were replicated by Caramazza and Brones in their cross-language condition: (1) correct *yes* responses took less time when the instance was typical of the corresponding category *(fruit-apple)* than when it was atypical *(fruit-melon)*, and (2) correct *no* responses took longer when the instance was drawn from a category semantically related to the category mentioned on the trial *(fruit-carrot)* than when drawn from a semantically unrelated category *(fruit-chair)*.

These findings can readily be understood in terms of the distributed conceptual representations proposed here, by assuming that the critical variable in the decision process is the number of conceptual elements a category shares with the instance presented on the same trial (cf. the interpretation of Smith et al., 1974, in terms of the number of shared features). Three specific assumptions need to be made: (1) A typical instance shares more conceptual elements with its category than an atypical instance. (2) Not only does an instance share conceptual elements with the category it belongs to, but it also shares some with a related category. (3) An instance and a semantically unrelated category do not have any of their conceptual elements in common. When there are many common elements (typical instance) and, hence, a large amount of activation at the intersection of the activation waves spreading out from the two presented words, the subject assumes the instance belongs to the specified category, and immediately responds *yes*. When there are no shared elements (unrelated non-instance) and, hence, no area of intersecting activation in conceptual memory, the subject assumes the instance does not belong to the specified category and responds *no* relatively fast. In the case of a few shared elements (atypical instance; related non-instance) and, therefore, some activation at the intersection, the subject has to be on guard, because either *yes* or *no* may be the correct response. He or she must somehow evaluate the links of the intersection, a process taking additional time. Consequently, the response times are relatively long on these trials. The decision process in the semantic-verification task is thus more complex than in the above general semantic-relation-assessment task, because in the latter the evaluation stage is redundant: Any intersection of activation indicates a relation, so whenever an intersection is detected, a *yes* response is appropriate (even on trials of the related non-instance type).

The fact that the response pattern in Caramazza and Brones' study was independent of the language of the stimulus materials suggests that the English and Spanish words for the (concrete) categories and instances used in their study shared the same set of conceptual elements in these bilinguals' memories.

A second bilingual investigation that belongs in this section is an unpublished study by Colletta, reported by McCormack (1977). It resembles that of Caramazza and Brones (1980) in that not *any* but a specific type of relation had to be searched for. Colletta's subjects were English-French bilinguals. They were presented with word pairs and had to decide for each individual pair whether or not it consisted of synonyms. The words within a pair were presented either in the same language or in different languages. Response times in the intra- and interlingual presentation conditions were equally long. This finding was seen as support for the view that translations share a representation in bilingual memory. In the present terms it again suggests that the corresponding words in the two languages share the same set of conceptual elements.

Integration or segregation?

Many an opening paragraph of writings on bilingual lexical organization states that the lexical knowledge of the bilingual may be represented in two language-specific memory stores, one for each of this bilingual's languages, or may instead be integrated in a single language-independent store. A tenet of the foregoing has been that the truth may lie somewhere in between these two extreme positions. Some words may have all of their conceptual representation, others relatively little or maybe even nothing in common with their closest translation. Another suggestion made in this chapter is that individual words may or may not share part of their conceptual representation with a semantically related word in the other language (and in the same language, but that is of less interest here). The data reported in this chapter suggest that the emerging representational form is likely to depend on word type (Is the word abstract or concrete? Does it evoke particular emotions or is it emotionally neutral? Is it a cognate or a noncognate?). But the degree of overlap between the meaning of a particular word and that of its closest translation may vary with word type, so it is possible that ultimately not word type *per se*, but the extent to which the meanings of the translations overlap, is the critical factor that determines how the two are stored in memory.

At various points in this chapter word concreteness was mentioned as a determinant of bilingual task performance. If amount of meaning overlap indeed underlies the effects of this variable (meaning overlap determining the amount of sharing between the conceptual representations, and the latter, in turn, determining the effects), concrete words and their translations must be more similar in meaning than abstract words and their translations. Although empirical data will have to be collected to substantiate it, the view that the meanings of concrete words are more similar across languages than those of abstract words is intuitively very plausible. The function of the entities referred to by concrete words will generally be the same in different language communities. Wherever we come across them, chairs are to sit on, and apples to eat. The appearance of these entities will also generally be the same across different language communities. That of man-made objects like chairs will to a large extent be imposed on them by their function, and hence be similar across different communities. That of natural objects like apples will generally be the same everywhere by virtue of the fact that they are natural categories. The end-product of learning a concrete word will thus be a representation of which the content varies relatively little across languages. Abstract words have no external referents that could be looked at, handled, utilized, and thus guarantee similarity of the content of the developing representations across languages. Their meanings have to be acquired by looking up these words' definitions (but see the next section) in a dictionary (or asking others to provide them), and, more importantly, by deducing them from the various contexts in which these words are used. To the extent that these contexts differ between languages (cultures), the meanings of these words will also differ. In sum, there are good grounds to assume that concrete words and their translations have very similar meanings, whereas abstract words and their translations have meanings that differ more substantially. Consequently, the chances that abstract words are represented language-specifically are larger than for concrete words.

For one group of abstract words an abundance of literature exists which bears on the present issue: Many ethnographic studies have been concerned with the meaning of particular emotion words in the community under investigation. Russell (in press) reviews the relevant literature and provides a wealth of examples suggesting that the reference of these words often differs between languages (cultures). A first indication for this is that languages differ considerably in terms of the number of words they possess to categorize emotions. The number of emotion words in different languages may vary between over two thousand at one extreme (in English, although only a minority of these may be in the vocabularies of individual speakers of English) and only seven at the other (in the Chewong language; Russell, in press). This may be taken to indicate that there are large cross-cultural differences in the extent to which people experience emotions (the fewer emotions, the fewer emotion words, or/and vice versa), but it may also indicate that the meaning of an emotion word and that of its closest translation in another language differs (of course, both may be the case). For instance, each of the emotion words in languages that contain only a few of them may cover more than the corresponding words in languages with a richer emotion vocabulary. But even when two languages have an equally large emotion vocabulary, the reference of corresponding words in the two languages may differ. Given two languages L1 and L2, some of the emotion words in L1 may have a broader, others a more narrow reference than the corresponding words in L2. Also, L1 words may exist for concepts that cannot be expressed in a single word in L2 or that do not exist as, or cannot even be conceived as, concepts in L2, and vice versa (the reader is referred to Russell, in press, for a thorough documentation of all these situations). All these words are likely candidates for language-specific representation in the memory of a bilingual whose two languages are L1 and L2. Kolers' (1963) word-association data discussed earlier support the present view that emotion words are relatively often represented language-specifically.

Some of the studies mentioned in this chapter suggested that, besides word concreteness and the emotional content of words, cognate status of the translation equivalents is yet another determinant of bilingual performance. If, again, the degree of meaning overlap between the translations is the critical factor underlying the observed effects (with representational form as mediator), cognate translations must have more similar meanings than noncognate translations. A reason for this could be the differential origin of cognate and noncognate translations: Cognate translations will generally derive from the same root in a common parent language. If they have both preserved the meaning of this root, or at least a large part of it, over time, they will have ended up having (about) the same meaning. In contrast, noncognate translations will generally not derive from the same root, and there is thus a relatively large chance that their meanings will differ more between the languages. But it may also be that not (or not only) differences in the degree of meaning similarity but (also) in perceptual similarity cause cognate translations to come to share more of their conceptual representations than noncognate translations: L2-learners, noticing the orthographic and phonological similarity between a cognate word and its translation, may simply assume the two have the same or about the same meaning, and thus conveniently link the new L2-word onto the conceptual representation of the corresponding L1-word. An interesting consequence of this may be that orthographic and phonological similarity of translations thus blind the learner to differences, if any, between the meanings of these two words.

There are other factors that may affect the way translations are stored in bilingual memory, but that I will only touch upon here. One is the circumstances under which the languages in question are acquired. Ervin and Osgood (1954) suggest that bilinguals who learn their two languages in different environments ('coordinate' bilinguals) develop a memory structure with separate representations for word translations in the two languages, whereas those who learn their languages by using them interchangeably ('compound' bilinguals) develop a memory structure with representations that are shared by the two translations. For instance, the common practice in foreign language classrooms where an L2 word is taught by directly associating it with its translation in the native language is one way to create compound bilinguals (see Keatley, 1992, for a longer discussion of this and related distinctions). Yet another critical factor may be whether or not the person's two languages belong to the same family: The chances that translations share a representation in memory (or share a relatively large part of their representations) may be larger when the languages of the bilingual are related than when unrelated. But here, again, the degree of meaning similarity between the translations may ultimately be the critical factor. Translations of words belonging to related languages may be more similar in meaning than those of words in unrelated languages. Not the presence or absence of a relationship between the languages *per se*, but this meaning similarity at the level of individual words, may determine how they will eventually be represented in memory.

The contents of the conceptual representations

The starting-point of the view on bilingual lexical memory set forth in this chapter was that a conceptual representation is composed of a number of conceptual elements. Until now nothing has been said about the nature of these elements. In this last section some of the relevant literature will be discussed.

According to what is known as the 'classical' or 'traditional' view of concept representation, concepts are represented by a fixed list of features that together *define* the concept (e.g., Katz, 1972, but dating back much longer, to the Greek philosopher Aristotle); that is, the features are individually necessary and jointly sufficient for membership of the category (I am using the terms *concept* and *category* interchangeably here, as is done more often). Assuming that concepts can be defined implies that the boundaries between concepts are clear-cut and stable. For instance, it should be clear where exactly cups turn into bowls, and where bowls turn into plates. Another implication of the classical view is that members of a category have equal status, that is, each member should be as good a member of the category as any other member. Yet another is that a given concept does not vary within the same individual or across individuals.

Many empirical findings and thought-experiments have cast doubt on this view of concept representation. I will mention just a few here. Wittgenstein (1953), to name an illustrious opponent, took the concept 'game' as an example with which he challenged it. He argued that this concept (and most others) cannot be captured in terms of a set of features that holds for all instances of this category. Rosch (e.g., 1973) collected experimental evidence suggesting that individual members of a category do not have equal status. Instead, many categories appear to have a 'graded' structure, with some members being more typical of the category than other members (a chair is a more typical instance of the category 'furniture' than a clock is). Typical members are those that can be captured in a set of 'prototypical' or 'characteristic' features that are listed for many (not *all*) members of this category, whereas atypical members can be described with a list that contains less of these features common to the category, but instead contains relatively many features not shared by most of the other members.

Experimental studies showing an effect of context on categorization constitute yet another serious challenge to the classical view. In an early study Labov (1973) collected data suggesting that concept boundaries are not clear-cut and static, but vary with context. His subjects had to categorize pictures of objects like cups, bowls and plates by naming the depicted object. Prior to naming them, they were instructed to, for instance, imagine someone holding the object and drinking coffee from it or, in a second condition, to imagine the object filled with mashed potatoes and sitting on the dinner table. It turned out that the objects were classified differently in different contexts: One and the same object was classified relatively often as a cup in the first of the above contexts, and relatively often as a bowl in the second. In related work, Barsalou (1987; Barsalou & Medin, 1986) provides a wealth of experimental data indicating that concepts vary with context, both 'long-term' context (people's experiences) and current context (e.g., linguistic context or point of view). For instance, two individuals' concept of 'bird' may differ because of different experiences of these individuals with birds, but the concept of one and the same individual may also differ at two different points in time because of this individual's new experiences with birds in the intervening period. In other words, people's representation of categories reflects their experiences (Barsalou & Medin, 1986). That concepts also vary with *current* context can be concluded from a study by Barsalou and Sewell (in Barsalou, 1987). Subjects judged instance typicality from one of several international points of view (for instance, from the American and Chinese points of view). Groups of subjects (sampled from the same population) taking different points of view produced different graded structures for the same category (for instance, *robin* and *eagle* were judged to be typical instances of the category 'bird' from the American point of view, whereas *swan* and *peacock* were considered typical from the Chinese point of view). Discussing the variability of concepts, Aitchison (1987, p. 40) uses some lively metaphors. She likens concepts to elusive butterflies and slippery fish: 'Word meanings cannot be pinned down, as if they were dead insects. Instead, they flutter around elusively like live butterflies. Or perhaps they should be likened to fish which slither out of one's grasp.'

All these studies thus indicate that it is not a fixed set of defining features that conceptual representations typically consist of (although *some* concepts may be represented that way). The studies showing an effect of current context on the content of concepts (Barsalou and Medin's point of view experiment, but also a study by Roth and Shoben, 1983, showing an effect of current linguistic context) suggest that people *construct* representations that suit the context (Barsalou & Medin, 1986). The clearest demonstration of this is by Barsalou (1983). He showed that people often construct new categories to achieve a current goal. These 'ad hoc' categories differ from common categories in that they are not well established in memory. Examples from his study are 'ways to make friends', 'things that could fall on your head,' and 'ways to escape being killed by the Mafia'. It could be argued that this study does not bear on the representation and processing of *common* concepts, but the data suggest otherwise: Ad hoc categories possess the same graded structures as common categories.

In short, conceptual representations appear to be constructed when needed. Does this imply that *all* of the concept is built up on-line, in other words, that there is no permanent representation in memory to be accessed as a whole each time the corresponding word is encountered, irrespective of context? If so, the view of representation set forth in this chapter would run into trouble, because it assumes the existence of such static part in the concept representations. Fortunately, it seems that a relatively stable *core* representation may still be assumed. These cores are generally not definitional (because word definitions exist for few words) but experientially-based (Barsalou & Medin, 1986). Barsalou (1982) distinguishes between context-independent and context-dependent properties (features) in concepts. Context-independent properties are activated each time the corresponding word is encountered (e.g., the property 'smells unpleasantly' when encountering the word *skunk*). Context-dependent properties are activated only by a relevant context in which the word occurs (e.g., the property 'floats' of the concept 'basketball' in the following sentence: *Chris used a basketball as a life preserver when the boat sank*; examples are taken from Barsalou, 1982). A concept's set of context-independent properties may constitute a relatively stable core representation in memory. *Relatively* stable, not just stable, because these cores are experientially-based. With new experiences the cores may change somewhat. But many of their properties will be immune to changes: Most birds will go on flying forever and chairs will always be for sitting. Furthermore, many of these properties will hold across languages. They are thus plausible candidates for the language-independent conceptual elements I have assumed in the preceding sections.

An interesting possibility to ponder is that the size of these cores varies with word concreteness (and maybe with other word characteristics as well), the cores of abstract words containing fewer elements than those of concrete words. One reason to consider this is the casual observation that, when asked to define a word, our response often is particularly clumsy in the case of abstract words. We often do not fare well with concrete words either (which is not surprising, given the fact that for most words definitions do not exist), but at least we can come up with some information on the associated concept (e.g., its characteristic properties; its function, if any; the superordinate; a number of subordinates). In the case of abstract words it seems that we can often only think of a number of contexts in which the word can occur. Another reason is based on a study in which I collected continued word associations to concrete and abstract words (de Groot, 1989). The words to be associated to were presented out of context, so I assumed the response words only to reflect context-independent information in the corresponding concepts. More responses were produced to concrete words than to abstract words, which I took to indicate that concepts corresponding to concrete words contain more context-independent information than those of abstract words do. Of course, the fewer context-independent properties in the representations, the fewer there are to be shared interlingually. This could explain the concreteness effects in bilingual processing tasks discussed in the preceding sections.

Conclusion

In the preceding sections a unitary account of performance in a number of bilingual processing tasks was suggested. Its starting point was a very simple one, namely, that conceptual representations consist of a set of meaning elements of which larger or smaller numbers may be shared by a word and its translation in another language. But it may be that things will eventually turn out to be far more complex than suggested here. Although many findings have been discussed, many others have been ignored or hardly attended to, for instance, those from studies investigating bilingual memory with episodic memory tasks. It remains to be seen whether the present framework could also account for those. Also, the present view assumes processes the details of which appear somewhat mysterious for the time being. How, for instance, are areas of (intersecting) activation in memory detected and, if necessary, evaluated? In spite of these questions, I hope to have succeeded in convincing the reader that the present view is a plausible alternative to the more established conceptions of bilingual memory and performance.

Acknowledgment
I would like to thank Lucia Dannenburg for composing the figures in this text, and Batja Mesquita for drawing my attention to a number of interesting studies on the emotion vocabulary.

References

Aitchison, J. (1987). Words in the mind: An introduction to the mental lexicon. Oxford: Basil Blackwell.

Altarriba, J. (1992). The representation of translation equivalents in bilingual memory. In R. J. Harris (Ed.), *Cognitive processing in bilinguals.* Amsterdam: Elsevier.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10,* 82-93.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition, 11,* 211-227.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization.* Cambridge: Cambridge University Press.

Barsalou, L. W., & Medin, D. L. (1986). Concepts: Static definitions or context-dependent representations? *Cahiers de Psychologie Cognitive, 6,* 187-202.

Caramazza, A., & Brones, I. (1980). Semantic classification by bilinguals. *Canadian Journal of Psychology, 34,* 77 - 81.

Chen, H.-C., & Leung, Y.-S. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 316-325.

Chen, H.-C., & Ng, M.-L. (1989). Semantic facilitation and translation priming effects in Chinese-English bilinguals. *Memory & Cognition, 17,* 454-462.

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

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 240-247.

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *The Quarterly Journal of Experimental Psychology, 38A,* 367-393.

Ervin, S. M., & Osgood, C. E. (1954). Second language learning and bilingualism. *Journal of Abnormal Social Psychology, Supplement,* 139-146.

Fischler, I., & Goodman, G. O. (1978). Latency of associative activation in memory. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 455-470.

Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 305-315.

Groot, A. M. B. de (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 824-845.

Groot, A. M. B. de (in press). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Groot, A. M. B. de, & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30,* 90-123.

Groot, A. M. B. de, Thomassen, A. J. W. M., & Hudson, P. T. W. (1982). Associative facilitation of word recognition as measured from a neutral prime. *Memory & Cognition, 10,* 358-370.

Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. (1986). Distributed representations. In D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 1: Foundations.* Cambridge: MIT Press.

Jin, Y.-S., & Fischler, I. (1987). *Effects of concreteness on cross-language priming of lexical decision.* Paper presented at the Southeastern Psychological Association Meeting, Atlanta, Georgia.

Katz, J. J. (1972). *Semantic memory.* New York: Harper & Row.

Keatley, C. W. (1992). History of bilingualism research in cognitive psychology. In R. J. Harris (Ed.), *Cognitive processing in bilinguals.* Amsterdam: Elsevier.

Kerkman, J. P. M. (1984). Woordherkenning in twee talen (word recognition in two languages). In A. J. W. M. Thomassen, L. G. M. Noordman, & P. A. T. M. Eling (Eds.), *Het leesproces* (The Reading Process). Lisse, The Netherlands: Swets & Zeitlinger.

Kirsner, K., Brown, H. L., Abrol, S., Chadha, N. N., & Sharma, N. K. (1980). Bilingualism and lexical representation. *The Quarterly Journal of Experimental Psychology, 4,* 585-594.

Kirsner, K., Smith, M. C., Lockhart, R. S., King, M. L., & Jain, M. (1984). The bilingual lexicon: Language-specific units in an integrated network. *Journal of Verbal Learning and Verbal Behavior, 23,* 519-539.

Kolers, P. A. (1963). Interlingual word associations. *Journal of Verbal Learning and Verbal Behavior, 2,* 291-300.

Kroll, J. F., & Curley, J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. In M. Gruneberg, P. Morris, & R. Sykes (Eds.), *Practical aspects of memory, Vol. 2.* London: John Wiley & Sons.

Kroll, J. F., & Stewart, E. (1990). *Concept mediation in bilingual translation.* Paper presented at the 31st Annual Meeting of the Psychonomic Society, New Orleans, November, 1990.

Labov, W. (1973). The boundaries of words and their meanings. In C.-J. N. Bailey & R. W. Shuy (Eds.), *New ways of analyzing variation in English.* Washington, DC: Georgetown University Press.

Masson, M. E. J. (1991). A distributed memory model of context effects in word identification. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition.* Hillsdale, NJ: Lawrence Erlbaum Associates.

McCormack, P. D. (1977). Bilingual linguistic memory: The independence-interdependence issue revisited. In P. A. Hornby (Ed.), *Bilingualism: Psychological, social, and educational implications.* New York: Academic Press.

Meyer, D. E., & Ruddy, M. G. (1974). *Bilingual word-recognition: Organization and retrieval of alternative lexical codes.* Paper presented at the Eastern Psychological Association meeting, Philadelphia, Pennsylvania, April.

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.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition.* Hillsdale NJ: Lawrence Erlbaum.

Potter, M. C., So, K-F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23,* 23-38.

Quillian, M. R. (1968). Semantic memory. In M. Minsky (Ed.), *Semantic information processing.* Cambridge, Mass: MIT press.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language.* New York: Academic Press.

Roth, E. M., & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology, 15,* 346-378.

Rumelhart, D. E., & Norman, D. A. (1985). Representation of knowledge. In A. M. Aitkenhead & J. M. Slack (Eds.), *Issues in cognitive modeling.* London-Hillsdale, NJ: Lawrence Erlbaum.

Russell, J. A. (in press). Culture and the categorization of emotions. *Psychological Bulletin.*

Scarborough, D. L., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning and Verbal Behavior, 23,* 84-99.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language, 27,* 499-520.

Schwanenflugel, P. J., & Rey, M. (1986). Interlingual semantic facilitation: Evidence for a common representational system in the bilingual lexicon. *Journal of Memory and Language, 25,* 605-618.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81,* 214-241.

Taylor, I. (1971). How are words from two languages organized in bilinguals' memory? *Canadian Journal of Psychology, 25,* 228-240.

Taylor, I. (1976). Similarity between French and English words—A factor to be considered in bilingual language behavior? *Journal of Psycholinguistic Research, 5,* 85-94.

Taylor, I., & Taylor. M. M. (1990). *Psycholinguistics: Learning and using language.* Englewood Cliffs, N. J: Prentice-Hall.

Tzelgov, J., & Henik, A. (1989). *The insensitivity of the semantic relatedness effect to surface differences and its implications.* Paper presented at the First European Congress of Psychology, Amsterdam, 2-7 July.

Wittgenstein, L. (1953). *Philosophical investigations.* New York: MacMillan.

Orthography, Phonology, Morphology, and Meaning
R. Frost and L. Katz (Editors)
1992 Elsevier Science Publishers B.V.

CHAPTER 21

Memory-addressing Mechanisms and Lexical Access

Kenneth Forster
Department of Psychology
University of Arizona, Tucson

The study of lexical access is important for two reasons. Not only does it deal with an integral component of the perception of language, but more generally, it raises a fundamental question about the functioning of the brain: how is previously stored information about an input pattern retrieved? The lexical domain is well suited to an experimental study of this problem, since words form a well-structured and easily manipulated set of patterns. Furthermore, due to the pioneering work of Herbert Rubenstein and his colleagues in developing the lexical decision task (Rubenstein, Garfield, & Millikan, 1971), we have access to a rich set of findings concerning the time it takes to recognize a word.

The central concept that integrates much of the theoretical and empirical work in this area is the concept of *content-addressable memory*. In this paper, we review the arguments for content-addressability, and consider the similarities and differences among the various models that have been proposed. In particular, we discuss the proposal that the mental lexicon is only approximately content-addressable, and that serial search mechanisms are inevitably involved in lexical access. We will discuss some of the evidence in favor of this claim, and deal with some of the objections to the notion of serial search.

Content-addressable retrieval

In a conventional computer memory, each memory location is assigned a number, which represents its address. Storage or retrieval of data from a particular memory location requires that the bit pattern corresponding to its address be first loaded into the address decoder, a circuit which selects the memory location designated by that address. Once selected (i.e., enabled), the data within this memory cell can be retrieved or modified. This retrieval function, *R*, can be expressed as follows:

$$R(address) = contents$$

For this reason, this type of memory is referred to as location-addressable memory. With content-addressable memory, however, what is retrieved is the address of the memory cell that has a specified content. That is,

$$R(contents) = address$$

Clearly, the process of word recognition must involve the latter kind of retrieval. The input is either an orthographic or a phonological description of a word, and the task is to find where in lexical memory this description is located. The retrieved address can then provide a pointer to additional information about the word, such as its semantic and syntactic properties.

Content-addressability can be achieved in two ways, depending on the computing hardware available. The first involves building a totally different kind of memory, commonly referred to as an *associative* memory. Neural nets are a good example of this type of device. In this case, the "key" that unlocks memory is a partial description of the pattern that we want to retrieve, rather than its address. However, if the available memory is not associative, but is strictly location addressable, then content-addressability must be achieved through software rather than hardware. Such programming techniques are referred to as *hash-coding* (see Kohonen, 1977). The essential part of hash-coding is a function that is designed to translate the input pattern into a unique address. Whenever data about a particular pattern is to be stored, the hash-code is used to decide where to store this information. Similarly, whenever data about a particular pattern is to be retrieved, the hash-code can be used to determine where that information has been stored.

Figure 1 illustrates how these two procedures could retrieve the pattern of semantic features associated with a given letter pattern. The device using associative memory is a familiar three-layer feedforward network, in which the input pattern representing the spelling of the word is first mapped onto a set of hidden units, and this representation is then mapped onto the output units, which represent semantic features. This particular network uses distributed representations in the sense that no word activates just a single hidden unit. The device using location-addressable memory looks similar, except that there is an additional level. At the lowest level, the input pattern is mapped onto a set of units which represent the address of the memory location that contains information about the input pattern. This mapping operation is carried out by the hash-coding function.

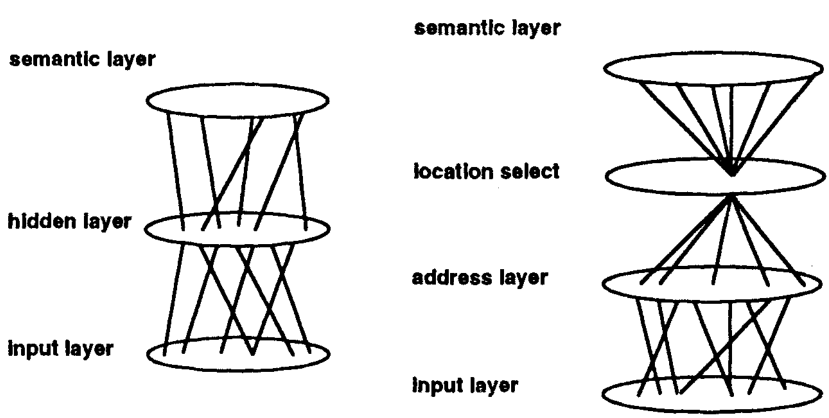

Figure 1. **Retrieval of semantic features associated with a printed word. The device on the left uses an associative memory, and consists of a conventional three-layer network. The device on the right uses a location-addressable memory, in which the input is mapped to an address by a hash-coding function.**

This representation is distributed in the sense that no input pattern activates a single unit, but the pattern of activation at this level is simply the binary code for the hash coded address. At the next level, the pattern representing the address is mapped onto a single unit, which represents the read/write select line for the memory location at that address. This is the conventional address decoder circuitry, which associates a single select line with every possible address pattern. The final stage maps from the read/write select unit to the semantic features, these being contained in the memory cell activated by the select line.

Although the functional design of these two systems appears to be similar, there are important differences in their mode of operation. For example, associative memory models allow for competitive processes within any layer, so that incompatible units are mutually inhibitory. This makes it possible to remove some noise from the input pattern. Another difference is that the activation passed from one unit to the next varies continuously as a function of the strength of the connection between them, whereas in the location-addressable model, everything is strictly digital. Another important difference is that the information that generates the final semantic representation is not represented in any single location in the associative memory model, as it is in the location-addressable model. Instead, it is effectively stored in the pattern of connections across the entire network.

Defects in the location-addressable model

In a hash-coding system, a function is defined that maps the properties of the target onto an address in memory. Any function will do, provided that it scatters the targets evenly throughout the address space. As a very simple example, assume that we assign a number to each letter determined by its position in the alphabet, and then simply sum the values of each letter in the word, regardless of position. Thus the word *BAT* generates the address 23, i.e., 2+1+20. When information about this word is initially stored, the hash-coding function specifies an address for that word (i.e., 23), and a pointer to the detailed information about the word (another address) is stored at that location. When information about the word is retrieved at a later date, the hash-coding function is again used to recover the pointer.

A problem arises in a hash-coding system when two different words yield the same address. In the example above, the word *TAB* would also generate the address 23, as would the words *KEG, AGO,* etc. These *collisions* arise when the hash-coding function fails to achieve an even scatter. There are many different algorithms for coping with collisions (Kohonen, 1977), and each of them involves some kind of serial search. For example, if there were three words involved in the collision, then we might simply store these three words as a special list somewhere else in memory, and store a pointer to this list at the address computed by the hash-code function. Recognizing the word would then involve comparing the target input with each of these words in turn.

Collisions can be avoided if the address space is suitably large compared with the number of words that are stored within it. The example above could be improved if we added a weighting factor that took position into account, say by multiplying the letter value by 10 for the first position, 20 for the second, 30 for the third, and so forth. This would yield an address of 640 for *BAT* (20+20+600), but 280 for *TAB* (200+20+60). This function will spread the words out further apart, and hence reduce the possibility of collision, although not eliminating it entirely. In order to guarantee no collisions at all, an enormously large address space would be required.

However, there is a more serious problem with hash-coding. As pointed out by Kohonen (1977), the location-addressable system is particularly vulnerable to ill-formed input patterns. If the pattern presented at retrieval differs slightly from the original stored pattern, or if part of the pattern is missing, then the hash-code will generate the wrong address, and any error in the address means that retrieval will fail. Since the most notable feature of human pattern recognition is its capacity to tolerate ill-formed inputs, this failure is particularly serious for modeling purposes.

These problems suggest that a location-addressable system can achieve only approximate content-addressability. The best that can be hoped for is that the hash-code will indicate the general location of the desired information, with the precise location being determined by a serial search procedure within this general area. This argument was the original motivation for the introduction of the "bin" theory of lexical access (Forster, 1976). In this model, the lexicon is subdivided into bins, each bin containing many patterns. All the entries within a bin are assigned the same address by the hash-code function. Location of the actual pattern that matches the input requires a serial search within the bin. In order to optimize this search, the patterns are ordered according to their frequency of occurrence, which minimizes search time. This frequency-ordering gives rise to the very strong frequency effects that are observed in a variety of word recognition tasks (e.g., Forster & Chambers, 1973; Monsell, Doyle, & Haggard, 1989; Rayner & Duffy, 1986). Ill-formed inputs do not create special problems for the bin model, provided that the ill-formed input generates the same bin number as the well-formed input. If this is the case, then the serial search will still have a chance of discovering the entry that is the best match to the input. But if this is not the case, then the pattern will not be recognized.

So to summarize, a location addressable memory model of the lexicon appears to have only approximate content-addressability. This forces the assumption of a serial search process. As implemented in the bin model, this search is frequency-ordered, which provides a way of explaining why it is that perfectly familiar, but relatively infrequent words such as *similarity* take longer to recognize than more frequently occurring, but no more familiar words such as *punishment.*

Problems with associative memory models

With an associative memory, the addressing circuitry is designed so that all memory locations essentially compare themselves with the input pattern simultaneously. In engineering terms, this increases the complexity of the memory considerably. Each memory location must be able to detect matching properties, and to keep track of the degree to which it matches the input. Associative memories of this type are a feature of both the logogen model (Morton, 1970), and the interactive activation *(IA)* model (McClelland & Rumelhart, 1981). Corresponding to each word in the lexicon there is a computational unit that is tuned so that it will be strongly activated by one letter sequence only, but necessarily will be activated by similar strings as well, although to a lesser degree. If the input stimulus is allowed to activate all word-units simultaneously, then the task of word recognition merely involves finding which unit is most strongly activated. The major difference between the logogen model and the *IA* model is in the procedure for locating the most strongly activated unit. In the logogen model, a criterion threshold is defined for each logogen, so that if the activation level exceeds this threshold, the logogen fires and inhibits the activation in all other units. However, in the interactive interaction model, a competitive process is set up between the output units, such that each unit inhibits all other units to a degree proportional to its activation level. This method of selection can be described as "survival-of-the-fittest," in contrast to the logogen model's "first-past-the-post" principle.

The resulting output is a pattern of activation across all output units. For the logogen model and the *IA* model, this pattern will be maximal activation in one output unit, and zero (or near-zero) activation in all the rest. This pattern corresponds to what is called a *localized* lexical representation. For each word, there is a single output unit. For a model using *distributed* lexical representations, the pattern might consist of activation in a number of different output units, so that no single unit is identified with any single word.

It is important to realize that the operational effectiveness of these memory models depends to a great extent on how well the post-activation selection process works. The reason for this is that the parallel nature of the activation process achieves nothing if non-parallel methods of finding the most strongly activated output unit must be used. So, obviously it will not do to suggest that the activation levels of all units are scanned in order to find the largest value, since this introduces a sequential search procedure. Either of the parallel selection procedures mentioned above would work perfectly well if activation levels were strictly a function of the degree of match to the input. However, in an effort to explain phenomena such as the word frequency effect, or the semantic priming effect, it is usually assumed that activation levels are also influenced by the frequency of occurrence of the word, or by the context in which it occurs. This undermines the effectiveness of the selection process, since it can no longer be assumed that the most active unit is the best match to the input. A good illustration of the problems that this causes is given by considering a pair of words with very similar form, but with markedly different frequencies of occurrence, such as *bright* and *blight*. In the logogen model, the logogen for *bright* would have a lower threshold than the logogen for *blight*, so the problem is to make sure that when we present *blight*, the activation level in the logogen for *bright* does not reach threshold first. In the *IA* model, the problem is to make sure that frequency does not have such a large effect that the activation level in the *bright* unit is not greater than the level in the *blight* unit (for discussion of these issues, see Forster, 1976, 1989a; also, see Norris, 1986 for an alternative view). Very similar arguments can be made for context effects. For example, how do we make sure that *blight* is correctly recognized in the context of a word such as *dim*?

The point being made here is simply that the failure to distinguish between sources of activation such as frequency, goodness of fit or contextual appropriateness can create problems for the selection process, and that there is a danger in attempting to explain all phenomena in terms of the one mechanism, namely, strength of activation. This is not a problem in a lexical search model, since different mechanisms are responsible for the effects. Frequency merely controls the order in which lexical entries are compared with the input stimulus, and hence the effect of frequency is quite independent of other influences. Admittedly, this involves postulating more than one mechanism to handle the effects, but if it can be shown that one mechanism is not enough to handle all effects, then there is little point in raising parsimony as an issue.

Is serial search fast enough?

The standard argument against a serial search model is that it could not possibly be fast enough to serve normal reading speeds. Assuming a rather modest limit of 50,000 entries to the lexicon, it might be supposed that at least half of those entries would have to be searched on average in order to find any word. Given reading speeds of around 200 words per minute, we would have to postulate search speeds in excess of 80,000 entries per second.

This argument overstates the case against a search process. First, the frequency ordering of the search set means that the average number of entries required to be scanned is well under half the total search set. This figure applies only to a random search. Further, the search is limited to just one bin, so the average number of entries searched will be less than half the number of entries in a single bin. The time taken to search an entire bin is estimated by subtracting the average time to process a very high frequency word (which is likely to be near the top of the bin) from the average time to process a very low frequency word (which is likely to be near the bottom of the bin). In a task such as lexical decision, this figure is probably somewhere around 100 ms. If we believe that no more than *N* entries could be scanned within 100 ms, then *N* will be our best estimate of the bin size. So, the argument about speed of search turns out to be an argument about the likely size of a bin, and this we have no way of knowing. All we know is that the bin size must be large enough to produce a graded frequency function.

The *XECAUSE* problem and parallel search

Although we can reduce the speed problem by restricting the search set to a single bin, there is still a serious problem remaining. It concerns the time taken to decide that a letter string is not a word. According to the bin model, this should be equal to the average time taken to search a bin, all other things being equal (e.g., the similarity of the nonword to an actual word). This correctly predicts that on average, nonwords will take longer to classify than words, since the entire contents of the relevant bin must be scanned before it can be established that the test item is not present. It also predicts that decision times for nonwords should be quite close to the decision times for very low frequency words, which also seems to be correct (e.g., Forster & Chambers, 1973; Stanners, Jastrzembski, & Westbrook, 1975; O'Connor & Forster, 1981). This much is quite straightforward. The problem arises when we consider nonwords that closely resemble an actual word, such as *xecause*. Such a nonword takes longer to classify than normal, which is presumably due to the fact that the entry for *because* is marked as a close match for the stimulus, and this needs to be evaluated further before a decision is reached. But this implies that the bin number generated for *xecause* is the same as the bin number generated for *because*. If it wasn't, then the entry for *because* would not have been encountered during the search for *xecause*, and no interference would have occurred.

Now consider other nonwords that resemble the same underlying word, such as *fecause, hecause, pecause, decause*, etc. It seems likely that these nonwords would all show an interference effect as well, which means that each of these nonwords must generate the same bin number as *because*. This implies that the hash code must be indifferent to the letter in first position. This is a puzzling conclusion, since it seems that the hash code could not possibly ignore such an important letter position. But this is also an absurd conclusion, since the argument can be extended to show that *none* of the letters are taken into account. Consider what happens when other letter positions are changed, producing *bxcause, bexause, becxuse, becaxse, becauxe*, and *becausx*. If each of these nonwords suffers from interference, then it follows that the bin number is unaffected by *any* of the letter positions. This result creates severe problems for the hash-coding hypothesis.

Are we in fact certain that an interference effect is generated no matter which letter is changed? One of the very first projects completed in our laboratory dealt specifically with this topic (Amey, 1973). Our original concern was to see whether the position of the changed letter altered the amount of interference obtained. It did not. Strong interference was generated no matter which letter position was changed. It is perhaps conceivable that this conclusion may apply only to group data and that a more careful examination of the data would reveal different patterns of sensitivity within subjects which cancel out when group data are used, but this seems very unlikely. For the moment, it seems more sensible to assume that the facts are as they seem, and so we are left with the problem of explaining how it is that the hash code for a given letter string can remain constant, regardless of changes in any one of its letters.

It seems most unlikely that it will be possible to design a hash coding scheme that has the desired constancy. As we pointed out earlier, it is generally agreed that hash coding is particularly vulnerable to errors in the input, and that is the problem we face here. Somehow we have to explain how the entry for *because* is always encountered no matter which bin is selected as the initial search set. This means that the only way out is to assume that classifying an item as a nonword involves an exhaustive search of *all* bins. That is, when the bin specified by the hash code is found not to contain the target, a much wider search is immediately triggered which eventually covers all bins, and hence the entry for *because* will always be encountered, no matter where we started.

This explains the interference effect neatly, but immediately raises the question of search speed, for we now must assume that the entire lexicon is searched before a "No" decision is made. This assumption predicts that there should be a very substantial difference in response latency between a "Yes" response to a very low frequency word (which will lie at the bottom of the initial bin) and a "No" response to a nonword. This difference would represent the time to search all the remaining bins. But as we have mentioned, this is generally not the case. Typically, the word-nonword difference in a lexical decision experiment is only marginally greater than the frequency effect. Translated, this implies that the time taken to search *all* bins is roughly comparable to the time taken to search a single bin, since that is what we assume that the frequency effect estimates.

The implication is clear. Bins must be searched in parallel. Instead of a single comparator unit that successively compares the members of a bin with the input, there are multiple comparator units, one for each bin, as shown in Figure 2. This parallel search scheme enables the entire lexicon to be scanned in a reasonable time, and explains the *xecause* effect.

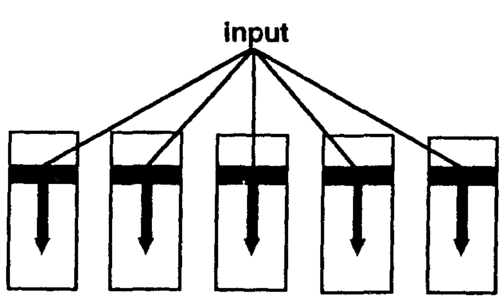

Figure 2. **Parallel search scheme in a bin model. For each bin, there is a separate comparator that scans the contents of its bin, comparing the entries to the input.**

But one difficulty remains. If the entire lexicon can be scanned in the same time as a single bin, why have an initial search of just one bin? A much simpler idea is to postulate that the *initial* search is in fact a parallel search of all bins. This has a surprising consequence, namely, there is no longer any need for a hash coding scheme. This means that newly acquired words can be assigned to bins at random, regardless of their orthographic or phonological form.

Parallel search and network models

If we decide to abolish hash coding altogether, we need to ask whether the new parallel search system differs at all from a network model. It appears that it does. In a network model, the activation mechanism makes it possible to compare the input with every word in the lexicon simultaneously. In the new parallel search model, the input is compared with *M* words simultaneously, where *M* is the number of bins. Further, these *M* words will all occupy the same *rank* position within their bins. That is, their frequency relative to other words in the bin will be the same. We can think of the search as first taking the most frequent members of each bin, and comparing them to the input stimulus in parallel. Then it proceeds to take all the next most frequent words, i.e., those with rank 2, and compares them with the input.

The next question to consider is whether there is any advantage to limiting the parallelism to just *M* words. Why not extend the parallelism to the entire lexicon? The principal advantage seems to be that from a hardware point of view, there is an enormous savings in the number of comparator units required. Also, monitoring the output of a small number of units may be a lot simpler than trying to monitor a very large number. Of course, there is a strong additional theoretical motivation for limiting the parallelism, and that is that it provides an account of the frequency effect. High-frequency words will still be accessed before low-frequency words in a parallel search, but only if *M* is small relative to the number of words in the lexicon. If this ratio approaches 1, as in a fully parallel network model, then there will be no frequency effect at all.

It may be that network models would also benefit from this type of design. In a three-layer feedforward network, this would involve a partitioning of the hidden unit layer into multiple sets, each set of connections being responsible for a different frequency band. In the example shown in Figure 3, there are five sets of hidden units, each set being fully connected to the input and output layers. Each set of hidden units is responsible for handling a separate set of words. Thus, there are five distinct sets of connections between the input and output layers. The critical part of the network is a control mechanism which selectively enables just one set of connections at a time (i.e., it inhibits all sets except one). This selector mechanism would first enable the connections for the highest frequency band, and then the next highest, and so forth. This will generate a frequency effect in much the same way as the serial search model.

This gating function could also provide an advantage in recognizing very similar words of different frequencies, e.g., *bright-blight.* These words will be in different frequency bands, and hence will be handled by different partitions of the hidden unit layer. This means that their output units will never be simultaneously active, and hence they will not compete. That is, when the output unit for *blight* is activated, the unit for *bright* will be inactive, since its hidden units will now be disabled. Similar gating functions have been proposed by Jacobs, Jordan, and Barto (1991).

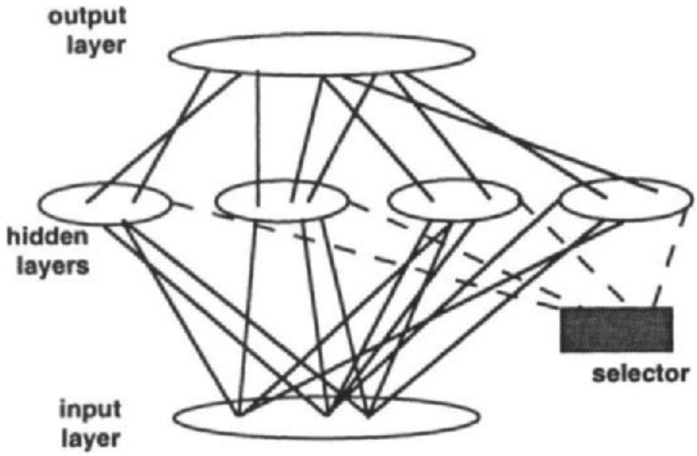

Figure 3. **Access with partitioned layers of hidden units. Each partition is responsible for a different subset of words. The selector mechanism is a gating function that selectively enables one partition at a time.**

The role of the frequency effect

Much of the case for a serial search mechanism rests on the frequency effect, which is taken to be diagnostic of the fact that full content-addressability is not feasible, at least as far as word recognition is concerned. Therefore, in order to make a strong case for full content-addressability, the problem of explaining the frequency effect must be confronted squarely. This can be done in two ways. Either it is shown that the dynamics of an associative memory also produce frequency effects with the right properties, or a case is made that the frequency effect is to some extent spurious. That is, it plays no role at all in the content-addressable part of lexical access.

In the logogen model, frequency effects were modeled purely in terms of criterion bias. That is, high-frequency words required less evidence to reach threshold. A similar mechanism is used in the interactive activation model proposed by McClelland and Rumelhart (1981), where the units for high frequency words have higher resting levels of activation. The other obvious mechanism for explaining the frequency effect within a network model is to propose that the more frequently a given pattern is presented during training, the stronger the connection weights become. This is the case in the simulation of naming performance carried out by Seidenberg and McClelland (1990). Each of these approaches essentially treats the effects of frequency as an increase in signal *strength.* As we argued earlier, this undermines the reliability of activation levels as measures of goodness of fit, and this can lead to complications. Where the models differ is in their assumptions about the activation gradients. In the interactive activation model, activation accumulates in the units for high and low frequency words at the same rate (i.e., the activation functions over time are parallel), whereas in models that use different connection strengths, the activation functions are not parallel. These assumptions have implications for the ways in which frequency ought to interact with other variables, such as stimulus quality. With non-parallel activation functions, the effects of stimulus degradation should be greater for low-frequency words than for high-frequency words, but with parallel functions there should be equal effects.

The critical question to ask here is whether stimulus probabilities play a role in perception. Is the fact that a test item begins with the letter *b* better evidence for the hypothesis that the item is the word *bright* than for the hypothesis that it is the word *blight*? One way or another, activation models generally answer this question in the affirmative, and adopt a Bayesian approach to detection, in which the a priori probability of the truth of the hypothesis is taken into account when evaluating the evidence. This is an entirely sensible approach if it is restricted to the *interpretation* of the evidence, but it is dangerous if it is also extended to the *collection* of evidence. The problem for many network models is that this distinction can be very difficult to draw. Indeed, it might be said that in some cases, this is a quite deliberate design feature.

Is the frequency effect an access effect?

It might be that frequency has no impact at all on the content-addressable part of the retrieval process, but rather, it affects some aspect of performance more related to task-demands. There are many possibilities to choose from. It could be that we are more confident in our response to high-frequency words, either because we are more confident of the spelling, or because we are simply more confident that we have made a correct discrimination. Or it could be that frequency affects the time it takes to become aware of the fact that an entry exists, perhaps because semantic information can be extracted more quickly from the entries of high-frequency words. These types of explanations are termed *post-access* because they take place after the correct entry has been isolated.

Given the pivotal role that frequency plays, it is important to evaluate carefully the assumption that frequency is a major determinant of access time. Most recently, this assumption has been called into question by the work of Balota and Chumbley (1984, 1985). Basically, their argument is that response times in both the lexical decision task and the naming task are multiply determined, and therefore we cannot assume without question that they reflect access time in any simple or direct way. Consequently, it is possible that frequency may be having its effects at more than one site. This point is indisputable, and need not be discussed any further. However, this research has sometimes been interpreted as having stronger implications, namely, that frequency has *no* effect on access at all, or that its effects have been vastly exaggerated. This claim is based mainly on their finding that there is no frequency effect for semantic categorization (e.g., "is it an animal?"), a task which clearly requires lexical access. A potential explanation of this result could be that it is only tasks that are sensitive to the *familiarity* of the stimulus that show a frequency effect. Lexical decisions are influenced in this way because a familiar letter string is far more likely to be a word than a nonword.[1] However, the familiarity of a word says nothing about whether it belongs to some semantic category or not, and is therefore irrelevant to this type of decision. Hence no frequency effect is expected.

This is an important argument, which focusses attention on the question of task analysis, and which is also relevant to the interpretation of other effects such as the repetition priming effect. The main weakness in the argument is the assumption that any task involving the retrieval of lexical information must automatically show a frequency

[1] We leave aside the problems of stating more exactly what familiarity is and how it is determined. We take it to be a state of mind induced by contacting an episodic memory trace of the stimulus.

effect. This assumption is too strong. Under some circumstances, the categorization task might be performed without a frequency-ordered search, as proposed in category-search models, where a pointer to the lexical entry might be found by searching through a list of common exemplars of the category (see Forster, 1989b). Another possibility is that some tasks may involve additional processes that may delay the eventual decision, thus obscuring the frequency effect (Bradley & Forster, 1987). These possibilities are made more reasonable by the fact that quite strong frequency effects *can* be obtained with the categorization task under certain conditions (Forster, 1989b; Monsell, Doyle, & Haggard, 1989). These effects cannot be attributed to familiarity, as we have already argued, and hence some additional mechanism must be postulated. We would then have three separate mechanisms to explain frequency effects: familiarity in the case of lexical decision, response output factors in the case of naming, and some third mechanism in the case of semantic categorization. The alternative view is that the frequency effect has the same explanation in each case (frequency-ordered access), but that additional task-dependent factors modulate this effect in various ways. This seems a more parsimonious approach.

But this is not to say that variables such as familiarity might not play some role in tasks such as lexical decision. Indeed, the repetition priming effect might be explained in just this way. Hence it is important to realize that the *size* of the frequency effect may be overestimated in the lexical decision task. In fact, the large difference in magnitude between the effects observed in lexical decision and naming is often taken as an indication of this overestimation. However, this could equally well mean that the size of the effect is *underestimated* in the naming task. This latter interpretation is supported by the fact that naming produces a much larger frequency effect when steps are taken to make sure that the naming response is controlled by lexically-driven processes (Monsell et al., 1989; Paap, McDonald, Schvaneveldt, & Noel, 1987).

Finally, we should point out that this discussion makes sense only if we believe that there is a sharp distinction between processes that occur during access and those that take place after access has been completed. This is not an empirical issue: the distinction can only be made *vis-a-vis* some particular theoretical model. If we use a model in which there is no clear distinction between the retrieval of stored information and the inferential processes involved in selecting an appropriate response, then it is meaningless to raise this issue. But it is not meaningless if we are considering a model in which such a distinction is made. This is not to deny that it may be unclear whether some process should be classified as an access process, or as a post-access process. For example, where should we classify processes such as checking routines, in which the spelling of the accessed entry is checked against the stimulus? Is this a post-access process, or should it be treated as part of the process of isolating the correct entry? This appears to be a simple matter of terminology. Access could be taken to mean all processes that are involved in selecting the final entry, in which case the spelling check would be an access effect. But if access is taken to cover just the processes involved in locating the entry (i.e., the search process itself), then it would be a post-access effect. In other words, there is no theory-free way to categorize effects.

The shape of the frequency function

It is well-known, but seldom discussed, that frequency of occurrence is related to performance in a logarithmic fashion. That is, variation in frequency has little impact on recognition performance at the high end of the frequency spectrum (e.g., comparing words with frequencies of 430 vs. 400), but the same difference would have a strong effect at the low end (e.g., comparing words with frequencies of 40 vs. 10). Such a function is generally taken to be a relatively uninteresting consequence of some psychophysical process, and there has been little interest in attempting to explain why the function should take this form. However, it turns out that the serial search assumption actually specifies the form precisely.

For a serial search model, frequency has no direct effect on access time at all. The frequency of a word merely determines the *rank position* of the word within the bin, and rank is related in a *linear* fashion to access time. That is, each increment in rank leads to a constant increase in access time. So the shape of the function relating frequency to access time will depend on the shape of the function relating frequency to rank. As can be seen from the example of a hypothetical bin in Figure 4, the resulting function is likely to have something like the desired property. The differences in frequencies between adjacent words at the top of the bin are likely to be very large, whereas at the bottom of the bin, they are very small.

By way of illustration, imagine a single bin containing all the words listed in the Kučera-Francis norms (Kučera & Francis, 1967). The order of the entries within this bin is given in the rank listing section of the norms. To search from the word at the top of the bin down to the word at the end of the first column involves going from a frequency of 69971 down to a frequency of 937, a drop of 69034. But going from the top of the next column to the bottom (the same amount of search) involves going from a frequency of 923 to 442, a drop of only 481. If we plot the frequency change for the first 25 columns, we get the function shown in Figure 5, which clearly is logarithmic.

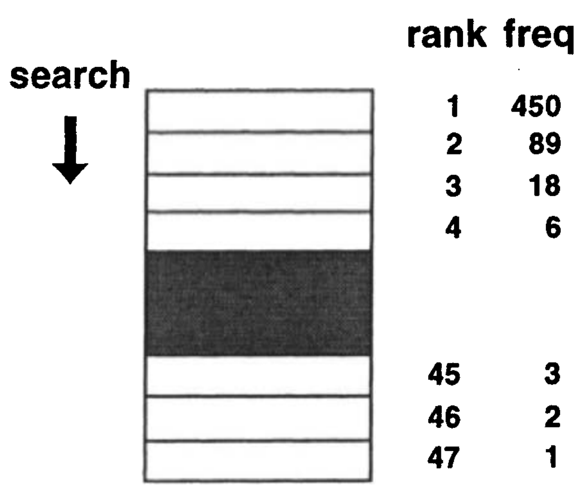

Figure 4. **Contents of a hypothetical bin, showing frequencies associated with entries at various rank positions.**

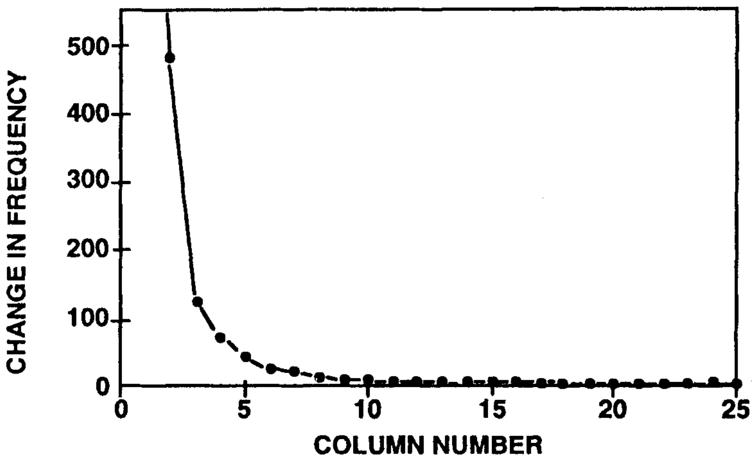

Figure 5. **Change in frequency for first and last entries in successive columns of the Kučera and Francis (1968) frequency norms. The change for the first column is off-scale (69,034).**

In an ongoing project, Wayne Murray and I have been attempting to test the bin model's account of the shape of the frequency effect. Our argument is that search time is a linear function of rank, and therefore to predict search time, we need to estimate the rank of a word within its bin. Since the details of the bin structure are unknown, we must assume that the rank of a word in a lexicon containing only one bin is a satisfactory estimate of its rank in a lexicon containing more than one bin. The absolute rank values will be incorrect, but relative ranks will not. That is, a word half-way between the top and the bottom of a single bin lexicon will still be somewhere near the midpoint of its bin in a multiple bin lexicon (on average). We then calculate the rank of each word, attempting to correct the ranks for errors introduced by spurious entries such as *vieux, Weigel's, C.A.I.P.*, which tend to become more common at the lower frequency ranges.[2] We then tested to see whether there was a linear relationship between these ranks and lexical decision times, as predicted by the serial search model. The procedure involved taking samples of 36 words at each of 16 frequency intervals. For example, we included 36 words in the frequency range of 315 to 197, a further 36 words in the range 100 to 85, and so forth, for 16 frequency ranges, the final group all having a frequency of 1. A total of 45 subjects were tested, each subject receiving only a sub-sample of the total set of items. The task was lexical decision, and the nonwords were orthographically legal, but not neighbors of words. The results of this experiment are shown in Figure 6. Clearly, the prediction of a linear relationship between average rank and lexical decision time is impressively supported.

Of course, we do not claim that other types of word recognition models could not make similar predictions. But it should be noted that in order for the search model to generate this prediction, no special assumptions were necessary. Many other models might

[2]Each spurious entry increases the error in estimation of the ranks of all words below it. Correcting for this error involves recalculating all ranks after pruning. Our procedure involved estimating the amount of change required by taking samples of words at various frequency ranges, and determining the proportion of spurious entries.

generate a similar prediction, but only at the cost of making special assumptions about how frequency is related to some mediating variable, such as familiarity, confidence, or strength of association. For example, if we argue that lexical decision time is related to familiarity, not search time, then we might simply assume that familiarity increases according to the log of frequency. This assumption seems quite plausible, but it makes the explanation completely vacuous nevertheless.

Subsequent research in this project is aimed at investigating whether these results can be duplicated for individual subjects. Figure 6 gives the impression of a smooth, graded function, but the true nature of the frequency effect might be a simple step function. For example, Balota and Chumbley (1984) suggest a decision procedure in which all words that exceed a particular familiarity value are immediately classified as words in the same time. Any item falling below this value has to be checked for lexical status, which increases decision time. However, the time taken to check the lexical status is not influenced by frequency. This model divides items in a lexical decision task into three categories: (1) highly familiar words that do not need access, (2) less familiar words that do require access, and (3) nonwords. Items within a category are all accessed in the same time. This model predicts a discontinuous step function relating frequency to decision time. The discontinuous nature of this function is normally masked in group data, since there is variation in the familiarity value of each item across subjects, and in the criterion value adopted by each subject. The only way to test for discontinuity is to examine the data for individual subjects. Figure 7 shows the functions for three individual subjects, each of whom completed the entire set of items used in the previous experiment on three separate occasions.[3] Once again, impressive evidence for a linear relationship between rank and lexical decision time is obtained for each individual subject.

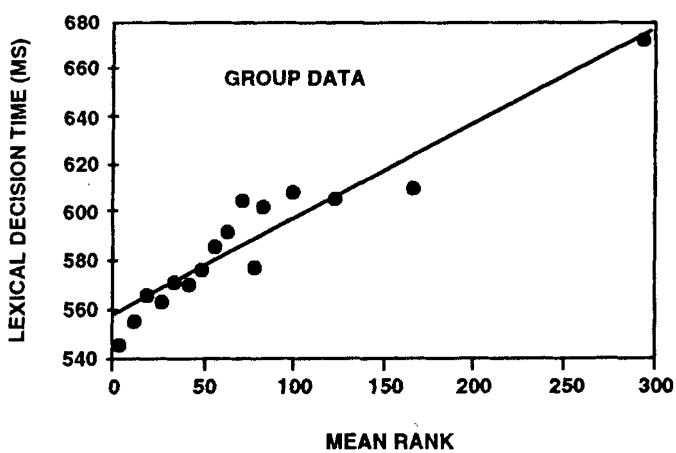

Figure 6. **Mean lexical decision times for 16 samples of words as a function of mean rank. Each data point is based on 540 observations.**

[3]The reason for repeating the materials three times was simply to increase the reliability of measurement, especially at low frequency ranges, where errors are more common. This repetition has the disadvantage of attenuating the size of the frequency effect, since low-frequency words benefit more from repetition than do high frequency words (see for example, Forster and Davis, 1984).

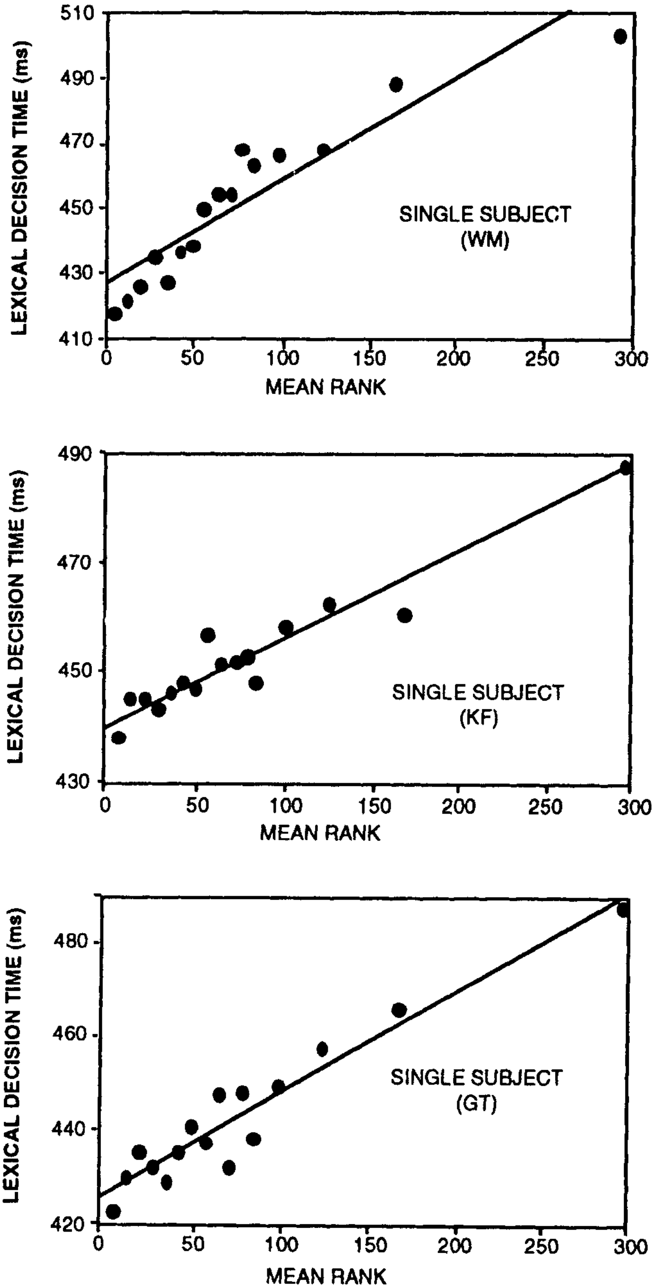

Figure 7. **Mean lexical decision times as a function of rank for three individual subjects. Each data point represents 108 observations.**

Frequency updating mechanisms

An important question so far unaddressed within lexical search theory is how the entries within a bin organize themselves into a frequency-ordered list. Several algorithms can be suggested. For example, each time a word occurs, a counter attached to its lexical entry could be incremented. The value of this counter is an index of the absolute frequency of occurrence of that word. The search is then organized within a bin so that the entries are scanned in order from the highest values of this index down to the lowest values. Unfortunately, this account begs a critical question: how is the entry with the next highest value found? Just to find the highest value would require a search of the entire bin, and then there would be another search required to find the second highest, and so forth.

This involves a great deal of preliminary scanning, as Grossberg and Stone (1986) point out in their evaluation of the serial verification model (Paap, Newsome, McDonald, & Schvaneveldt, 1982). This is clearly an unworkable proposal if this preliminary scanning has to be carried out "on the fly," each time a word is encountered. This appears to be the case for the verification model, as currently formulated. This model differs from the bin theory in that the search set changes for every different input stimulus. Paap et al. use a parallel activation procedure to define a unique set of words that have similar orthographic properties to the current input, and then this set is subjected to a frequency-ordered search. However, it is not explained how those entries get ranked according to frequency.

The bin model escapes this problem because its bin structure is *static*. That is, the contents of each bin remain constant no matter what the stimulus may be. Hence it is possible for the rank-ordering to be *pre-compiled*. The most obvious way to implement this would be to use a linked-list structure, in which each entry contains a pointer to the next entry in the list, so that the order of search is defined by pointers, not the physical address of the entry, as in a location-addressable memory. So, one could postulate some entirely independent procedure that is responsible for adjusting the pointers so that the first member of the list defined by the pointer system is always the entry with the highest frequency index, and this entry contains the address of the next-highest entry, and so forth.

Of course, the bin structure cannot be *entirely* static, since the relative frequencies of words within a bin will change over time. Hence the pointer system must undergo modification, if it is to accurately reflect current circumstances. What is needed, then, is a theory of how pointers get changed as a function of experience. The most economical assumption is to assume that the bin is essentially like a *stack*. Whenever a word is accessed, its entry is temporarily removed from the list, and then reattached at the *top* of the bin, rather than at its original position. This will produce a clear frequency effect, since words that are seldom used will get pushed down to the bottom of the stack, while words that are frequently used will never be far from the top of the stack. This account makes it quite unnecessary to have a frequency counter for each entry. However, the ordering is in terms of *recency*, not frequency. This is not unreasonable, since recency is really the property that underlies relative frequency. The more recently a word has occurred, the more likely it is have a high frequency of occurrence.

The weakness in the stack mechanism is that it predicts that there should be no frequency effect at all for recently presented words, since recency is the only effective variable. However, from studies of repetition priming, we know that there is a residual effect of frequency for recently presented words (Scarborough, Cortese, & Scarborough, 1977), so this arrangement will not work. An improvement can be gained if, instead of promoting an entry right to the top of its bin when it is used, we use a *partial* promotion scheme. For example, an entry could be promoted to a position half-way between its current position and the top of the bin. This scheme has a number of useful properties. First, it explains why a frequency effect is still obtained for recently presented words. Both high- and low-frequency words benefit from recency. However, the low-frequency words can never entirely catch up to the high-frequency words. Even in the limiting case where the high-frequency word is at the top of its bin, and hence cannot be promoted, the low-frequency word will still only move half the distance between its position and the position of the high-frequency word. The second feature of interest is that low-frequency words benefit more from repetition than do high-frequency words, which corresponds well with the findings in repetition priming studies (Forster & Davis, 1984; Scarborough et al., 1979).

The findings from repetition studies must be evaluated with care, however, since it can be argued that at least some of the improvement in performance must be due to episodic influences (Feustel, Shiffrin, & Salasoo, 1983; Forster & Davis, 1984). That is, the first presentation of a repeated word lays down some type of memory trace, which may then influence the subject's decision-making about the second occurrence of this word. Such an effect could be described as an implicit memory effect, since the benefit occurs whether the subject attempts to recall the earlier presentation or not, and even if the subject would have been unable to recognize the word as a repeated item (Schacter, 1987). So the fact that the partial promotion mechanism fits the repetition data so well might be misleading. On the other hand, if we have independent grounds for believing that a partial promotion mechanism actually exists, then it must be acknowledged that the repetition effect has exactly the right properties.

Of course, there are alternatives to partial promotion that would predict quite different effects of repetition. For example, one way to produce a rank order would be to use a "bubble-sort" algorithm. Each entry has a frequency counter, and whenever its counter registers a higher value than that of its upstairs neighbor, it exchanges places with it. In this type of system, changes in rank order would take place only very gradually, and there probably would be no detectable effects of a single repetition. The reason is simply that the greatest change that could occur as a result of a single presentation is that the word moves up one position, which ought not to have very marked effects if bins contain many entries. Another possibility is that the frequency counter is replaced by a time-tag which indexes how recently the entry was last accessed. This time tag generates a strong signal immediately after access, and the strength of this signal gradually decays with time. Promotion might then depend on whether the time tag was active when the word was accessed. This provides a crude index of the rate at which a word occurs.

Further work is required to establish which aspects of the effects of repetition reflect internal changes in the bins, and which aspects reflect the external influence of a memory trace on decision processes. The kind of effect that would fall into the latter category is one where the target word activates the memory trace of the first presentation, producing a "feeling of familiarity," which in turn produces a response bias to respond "Yes" in a lexical decision task. What is needed to remove this effect is a task that is sensitive to frequency effects, but is not subject to bias effects. If such a task fails to show any repetition effects, then this will suggest that something like the bubble-sort algorithm is closer to the truth than the partial promotion algorithm.

Age of acquisition and frequency

There is one curious feature of the partial promotion algorithm that leads to interesting predictions. Since a word can only move some proportion of the distance remaining to the top of the bin, it follows that a word could never displace another word at the top of the bin (much like the paradox of the frog that could only jump half the distance to the goal). This would mean that the words at the top of each bin could never be dislodged, which raises the question of how they got there in the first place. The answer would have to be that these words must *always* have been at the top of their bins.

How could this be? Perhaps during early lexical acquisition, there is only a single bin, and words are added to this bin in order of their acquisition. When this bin gets too large, a new bin is added, and further words are added to this new bin. So, the initial order laid down is purely in terms of the order of acquisition, but this is subsequently revised as the relative frequencies of the words change -- except for the words that were placed at the top of their bins, which will remain there forever!

Such an argument may seem farfetched, until it is recalled that there is a serious question as to whether access is controlled by the frequency of occurrence of words, or the order in which they were acquired, i.e., age of acquisition. Carroll and White (1973) argued that it is age of acquisition alone that determines access time. They suggested that the apparent correlation of access time with frequency is due to the fact that frequency and age of acquisition are themselves highly correlated variables. When the effects of age of acquisition were removed statistically, there was no effect of frequency.[4] But this conclusion seems too strong. Some words must surely become high-frequency only in adulthood (e.g., words describing recently invented technology), and it is difficult to imagine that these words permanently function as if they were low-frequency words. Strong support for this argument comes from the work of Gardner, Rothkopf, Lapan, and Lafferty (1987), who found differences in lexical decision times as a function of the occupational background of the subjects. Nurses were faster in responding to medical words than engineering words, while engineers showed the reverse effect.

This issue has been explored in an unpublished study carried out by Linda Cupples in our laboratory. Instead of attempting to control age-of-acquisition (AOA) by regression techniques, Cupples manipulated AOA and frequency factorially in a 2 × 2 design. This involved finding words that were acquired early (as indexed by the rating norms provided by Gilhooly and Logie, 1980), but which varied in frequency of occurrence (as indexed by Kučera-Francis frequency). High-frequency examples are plentiful (e.g., *money, circle, uncle*), but low-frequency examples are more difficult to find (e.g., *peach, rattle, spoon*). For words acquired relatively late in life, high frequency examples are *novel, belief, election,* while low frequency examples are *baron, receipt, peasant.* The experimental task was lexical decision, and college students served as subjects. The results are shown in Figure 8. The surprising outcome is that *both* frequency and age of acquisition exert a reliable effect of equal magnitude, and there is an interaction between them: words acquired late show a normal frequency effect of about 80 ms, whereas words acquired early show a much smaller frequency effect of only 20 ms.

[4]The opposite result was obtained by Gilhooly and Logie (1981), who found no independent effects of age of acquisition when frequency and familiarity were controlled.

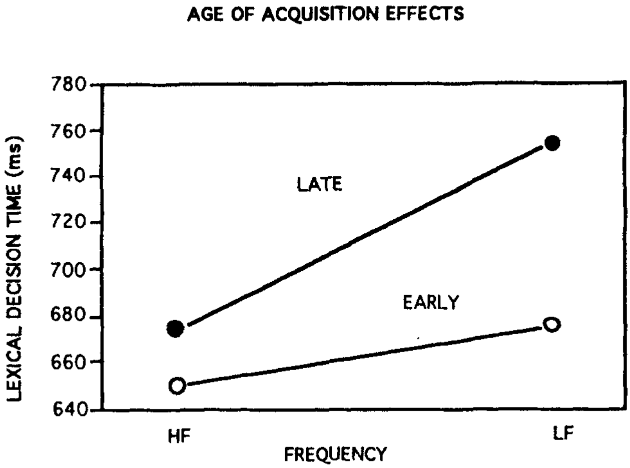

Figure 8. **Age-of-acquisition effects for high and low frequency words.**

Obviously, there are many possible reasons for this interaction. Because frequency and AOA are so highly correlated, the set of words that permits these factors to be manipulated orthogonally must be rather special, and it is entirely possible that some other variable has been accidentally manipulated (e.g., concreteness). Nevertheless, it is interesting to observe that partial promotion can explain some aspects of these results. For example, it predicts an AOA effect for high-frequency words, since early-acquired words will have a better chance of getting to the top position, and hence are less likely to be subsequently dislodged by late-acquired words. This appears to be true even when the subsequently acquired words are much higher in frequency, since the early-acquired words that are now low-frequency in adulthood are accessed in the same time as much higher-frequency words acquired late. Partial promotion also predicts that frequency effects should be greater for late acquisitions, since frequency is the only variable that controls the position of these words. That is, these words are free to move, whereas the early-acquired words are less so.

The most important feature of these results is that there appears to be both a *primacy* effect, and a *recency* effect. Words acquired early retain an advantage over later words, despite large variations in frequency. Obviously, words acquired early in childhood must have been high frequency words *at that time*, and so it is possible to describe the results purely in frequency terms. That is, words that have been high-frequency at *any* time in the life of a particular individual will retain a partial advantage over other words, regardless of their current frequency. In effect, once a high-frequency word, always a high-frequency word. The same is not true for low-frequency words, however.

Finally, it should be stressed that frequency updating is not just an issue for the bin model. Some kind of frequency updating mechanism must be specified in *all* models of word recognition, regardless of whether they use a serial or parallel search mechanism. What we have seen here is that the notion of rank position in the serial model offers some interesting perspectives on the nature and consequences of frequency updating, which are worth exploring in more detail. In particular, we need to explore whether movement from higher to lower positions is indeed more sluggish than movement in the reverse direction, as the AOA effect suggests. Also, we need to consider whether this creates problems for the bin model. As currently stated, the "frog" model restricts this effect to just to the topmost entry of each bin. But it seems more plausible to imagine this as a more general property, so that entries close to the top also share this property, only to a lesser degree. This implies that the higher you are, the less distance you can fall. But will this work? Surely there has to be some kind of *balance* between the number of entries moving up and the number moving down?

Note that this problem does not arise in other types of models. For example, if the frequency effect is to be explained in terms of a higher resting level of activation, then the frequency updating rule would involve a rapid increase in the resting level immediately after activation of the word, with a slow, gradual decline to the original baseline. However, there is nothing to correspond to the notion of *balance*. There is no requirement, for instance, that the average activation level over all entries should remain constant. There is no theoretical reason why *all* words could not have high resting levels at some given moment. Or, if frequency is assumed to affect the threshold at which a word detector fires, then it is possible that *all* word detectors could eventually acquire low thresholds. But this state of affairs is impossible in the bin model. It is impossible for *all* words to be at the top of their bin simultaneously.[5] If one word is promoted, then some other words must be demoted. This notion of balance is a core assumption of the bin model. Whether this principle of balance can be directly tested remains an interesting theoretical challenge.

Acknowledgment
This research was supported in part by a grant from the McDonnell-Pew Cognitive Neuroscience Program, and also by grant DC-01409, a research and training grant funded by the National Institute on Deafness and Other Communicative Disorders (NIDCD).

References

Amey, T. (1973). Word recognition: An investigation of feature analysis and lexical access time. Unpublished honors thesis, Monash University.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language, 24,* 89-106.

Bradley, D. C., & Forster, K. I. (1987). A reader's view of listening. *Cognition, 25,* 103-134.

Feustel, T. C., Shiffrin, R. M., & Salasoo, M. A. (1983). Episodic and lexical contributions to the repetition effect in word identification. *Journal of Experimental Psychology: General, 112,* 309-346.

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

Forster, K. I. (1989a). Basic issues in lexical processing. In W. Marslen-Wilson (Ed.), *Lexical representation and process.* Cambridge, MA: M.I.T. Press.

[5] Unless there are as many bins as words, in which case the bin model is really a fully content-addressable model (see Forster, 1989a).

Forster, K. I . (1989b). On knowing how many entries. In D. S. Gorfein (Ed.), *Resolving semantic ambiguity* (pp. 126-145). New York: Springer-Verlag.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12,* 627-635.

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.

Gardner, M. K., Rothkopf, E. Z., Lapan, R., & Lafferty, T. (1987). The word frequency effect in lexical decision: Finding a frequency-based component. *Memory & Cognition, 15,* 24-28.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation, 12,* 395-427.

Gilhooly, K. J., & Logie, R. H. (1982). Word age-of-acquisition and lexical decision making. *Acta Psychologica, 50,* 21-34.

Grossberg, S., & Stone, G. (1986). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review, 93,* 46-74.

Jacobs, R. A., Jordan, M. I., & Barto, A. G. (1991). Task decomposition through competition in a modular connectionist architecture: The What and Where vision tasks. *Cognitive Science, 15,* 219-250.

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

Kohonen, T. (1977). *Associative memory: A system theoretical approach.* New York: Springer-Verlag.

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.

McClelland, J. L., & Rumelhart, D. E. (1986). *Parallel distributed processing: Explorations in the microstructures of cognition. Volume 2: Psychological and biological models.* Bradford: M.I.T. Press.

Monsell, S. (1985). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language; Volume 2.* London: Lawrence Erlbaum Associates.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on word recognition tasks: Where are they? *Journal of Experimental Psychology: General, 118,* 43-71.

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

Norris, D. (1986). Word recognition: Context effects without priming. *Cognition, 22,* 93-136.

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

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.* London: Lawrence Erlbaum Associates.

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition. *Psychological Review, 89,* 573-594.

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.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9,* 487-494.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructures of cognition. Volume 1: Foundations.* Bradford: M.I.T. Press.

Scarborough, D. L., Cortese, C., & Scarborough H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 1-17.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13,* 501-518.

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

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

Index of Authors

Andjelković, D., 7, 343, 345, 347, 351, 353, 355, 359
Beauvillain, C., 7, 366, 376, 377, 381, 384, 387, 388
Bentin, S., 2, 4, 5, 27, 30-33, 35- 38, 40, 42, 43, 49, 64, 69, 73, 76, 80, 83, 86, 114, 115, 193-195, 197, 201-205, 207, 213, 221, 223, 224, 242, 243, 246, 324, 337, 343, 350, 351, 358, 362, 365, 367, 375
Berent, I., 5, 227
Besner, D., 2, 45, 46, 49, 51-53, 55, 56, 58, 59, 61, 63-65, 66, 78-80, 83, 96, 98, 113-117, 140, 144-146, 148, 156, 162, 211, 216, 224, 230-232, 241, 243, 244, 246, 248, 253, 264, 279, 280-282, 284, 285, 312, 315-317, 320, 328, 335-338, 367, 376, 388, 410, 411
Brichetto, T., 5, 249
Burani, C., 7, 361, 363, 364, 366, 367, 371, 373-375, 378, 387
Carello, C., 5, 52, 53, 64, 87, 116, 211, 215, 223, 225, 247, 324, 336, 337
Colombo, L., 6, 319, 321-323, 325, 326, 328, 329, 333, 335, 336
Feldman, L. B., 7, 40,-43, 49, 50, 51, 53, 58, 64, 66, 73, 75, 76, 83, 85, 86, 87, 89, 95, 96, 115-117, 181, 191, 215, 217, 224, 225, 232, 243, 246-248, 324, 337, 343, 344,--346, 347, 349-351, 353, 355, 358, 359, 361, 362, 365, 367, 375, 376, 379, 388, 389, 411
Forster, K., 7, 49, 64, 89, 115, 131, 132, 134, 137, 138, 140, 145, 220, 224, 250, 253, 283, 317, 319, 324, 336, 347, 358, 359, 361, 367, 376, 380, 388, 413, 416, 417, 418, 423, 429, 432, 433
Frost, R., 1, 2, 27, 28, 30, 33, 35-40, 42, 43, 49, 50, 51, 53, 58, 61, 64, 67, 69, 73, 76, 77, 79, 80, 81, 83, 86, 88, 92-95, 102, 115, 181, 189, 191, 195, 204, 205, 209, 213, 221, 224, 225, 242, 243, 246, 324, 335, 337
Garlington, K. L., 5, 249, 282
Grainger, J., 3, 131, 132-137, 138, 139, 140-146, 148, 161-163, 235, 236, 246, 348, 357, 359, 360, 379, 380, 388
Groot, A. M. B. de, 7, 324, 336, 389, 391-393, 396, 397, 398, 400-402, 408
Hung, D. L., 3, 85, 115, 119, 120, 127, 129, 130, 238, 248
Johansen, L. S., 5, 293
Johnson, N. F., 3, 147, 148, 152, 156-158, 159, 160, 161, 163
Katz, L., 1, 2, 25, 27, 28, 30, 33, 38, 42, 43, 49-51, 53, 58, 64, 67, 72, 73, 75, 76, 77, 79, 80, 83, 86, 87, 95, 115, 116, 157, 164, 181, 189, 190, 195, 208, 213, 215, 221, 223-225, 242, 243, 246, 247, 324, 337, 358, 360, 406, 410
Laudanna, A., 7, 349, 360, 361, 363, 367, 368, 370-372, 375, 376, 378, 387
Liberman, A. M., 4, 15, 22, 23, 25, 26, 30, 31, 71, 72, 84, 167, 171, 172, 176, 178, 180, 181, 184, 188, 189, 191, 193-196, 201, 203, 204, 208, 209, 211-213, 223, 225, 227, 229, 247, 301, 318
Lukatela, G., 5, 52, 55, 64, 66, 73, 84-87, 89, 116, 117, 131, 146, 181, 191, 211, 213-220, 222-226, 232, 235, 236, 243, 247, 248, 278, 284, 324, 336, 337, 358, 360, 361, 378, 388
Lundquist, E., 4, 179
Markson, L. R., 5, 249, 282
Mattingly, I. G., 2, 11, 14, 15, 17, 18, 23, 26, 42, 71, 83, 84, 171, 172, 178, 179, 191-193, 207, 209, 211, 212, 225, 227, 247
Noel, R. W., 5, 48, 61, 65, 218, 226, 230, 235, 247, 250, 284, 293, 295, 297, 298, 301, 309-316, 318-321, 335, 338, 423, 433
Paap, K. R., 5, 6, 48, 61, 65, 81, 87, 131, 132, 134, 137, 138, 146, 160, 163, 218, 226, 230, 235, 247, 250, 284, 293, 295, 297, 298, 301, 309-316, 318, 319, 320, 335, 338, 423, 428, 433
Perfetti, C. A., 5, 30, 43, 73, 74, 83, 84, 114, 115, 120, 127, 129, 184, 191, 204, 210, 217, 220, 226, 227, 229, 232-235, 238, 239, 240, 245-248, 265, 278, 284
Pinnt, G. S., 5, 249, 282
Segui, J., 7, 132, 133, 135-141, 143-146, 148, 161, 163, 347, 348, 359, 360, 366, 376-381, 388
Seidenberg, M. S., 3, 6, 48, 49, 53, 61, 66, 74, 78, 81, 82, 84, 85, 87-93, 95, 97-103, 105, 107, 108, 112, 113, 115, 117-119, 127, 129, 131, 134, 144, 146, 148, 164, 230, 231, 235, 244, 248, 259, 260, 278, 280, 281, 283, 284, 297, 315, 316, 318, 320, 324, 329, 333-335, 338, 344, 360, 421, 434
Shankweiler, D., 4, 22, 23, 26, 30, 31, 43, 44, 71, 72, 83, 84, 176, 178-182, 184, 186-192, 194, 195, 204, 208, 209, 211, 225, 227, 229, 247, 301, 318
Simonfy, C. M., 5, 249
Smith, M. C., 2, 45, 59, 63, 322, 335
Stone, G. O., 5, 61, 66, 73, 74, 84, 97, 118, 211, 223, 224, 226, 230, 248, 249, 259, 264, 265, 272, 278, 285, 295, 317-319, 338, 428, 433
Tabossi, P., 6, 53, 66, 79, 80, 84, 105, 117, 319, 324, 328, 335, 338
Turvey, M. T., 5, 49, 52, 55, 64, 66, 73, 83-87, 89, 98, 100, 102, 115-117, 131, 146, 181, 191, 211, 213, 215-220, 222-226, 232, 235, 236, 243, 246-248, 278, 284, 324, 336, 337, 378, 388
Tzeng, A. K. Y., 119
Tzeng, O. J. L., 3, 26, 85, 115, 119, 120, 121, 127, 129, 130, 178, 163, 226, 238, 248, 282, 358
Van Orden, G. C., 5, 6, 61, 66, 73, 74, 83, 84, 88, 96- 99, 101, 106, 107, 110, 112, 118, 211, 214, 218, 219, 220, 222, 223, 226, 230, 232, 248-251, 253-255, 258, 259, 260, 262, 264-268, 272, 277, 278, 279, 280, 282-285, 295, 301, 305, 310, 317-319, 338
Zhang, S., 5, 14, 26, 127, 129, 199, 227, 232, 240, 248

Printed in the United Kingdom
by Lightning Source UK Ltd.
115458UKS00001BD/3