CHAPTER 9

VISUAL WORD RECOGNITION

THE JOURNEY FROM FEATURES TO MEANING

DAVID A. BALOTA

I. INTRODUCTION

A. The Word

Ah yes, the word! The word is as central to psycholinguists as the cell is to biologists. In the present chapter, I review some of the major issues that have been addressed in visual word recognition research; other chapters in this volume are devoted to auditory word recognition. At the onset, I should note that the present review is from the perspective of a cognitive psychologist, not a linguist. Moreover, the goal of the present review is not to provide in-depth reviews of every area addressed by word recognition researchers. This would far exceed space limitations. Rather, I will attempt to acquaint the reader with the richness and diversity of the empirical and theoretical issues that have been uncovered in this literature.

The organization of the chapter is as follows: First, I briefly outline why word recognition research has been central to a number of quite distinct developments in both cognitive psychology and psycholinguistics. Second, I review the evidence regarding letter recognition, sublexical organization, and lexical-level influences on word recognition. Interspersed within each of these sections is a discussion of some of the current theoretical developments and controversies. Third, I review the literature on context effects in word recognition, again highlighting major theoretical developments and controversies. Fourth, I conclude by discussing limitations regarding inferences that are possible based on the available data and suggest some avenues for future research in this area. However, before turning to the review of the literature, I begin by discussing why lexical-level analyses have been so central to developments in both cognitive psychology and psycholinguistics.

COPYRIGHT © 1994 BY ACADEMIC PRESS, INC.
ALL RIGHTS OF REPRODUCTION IN ANY FORM RESERVED.

B. Why the Word?

In order to provide a framework for understanding the breadth of word recognition research, it is useful to list a few of the basic research issues that the word recognition literature has touched. For example, word recognition research has been central to notions regarding different levels/codes of analysis in language processing, attention, and memory (e.g., Craik & Lockhart, 1972; Posner, 1986). The lexical unit is ideally suited for such work because words can be analyzed at many different levels, e.g., features, letters, graphemes, phonemes, morphemes, semantics, among others. As we shall see below, much of the work in word recognition has been devoted to identifying functional roles of these different levels in the word recognition process.

A second domain where word recognition research has been a central player has been in the development of theories of automatic and attentional processes (e.g., Healy & Drewnowski, 1983; Laberge & Samuels, 1974; Neely, 1977; Posner & Snyder, 1975). Part of the reason for this emphasis is the natural relation between the development of reading skills and the development of automaticity. Here, one can see the extra impetus from education circles regarding the development of word recognition skills. Moreover, the notion that aspects of word recognition have been automatized and are no longer under conscious control of the reader has provided some of the major fuel for arguments regarding self-encapsulated linguistic processing modules (see Fodor, 1983).

Third, word recognition research has also been central to developments regarding basic pattern recognition processes. One of the most difficult problems in pattern recognition research has been in identifying the underlying subordinate critical features of a given pattern (e.g., Neisser, 1967). Written words are relatively well defined patterns. Historically, words have been the central unit of analysis in much of the verbal learning and memory research that dominated experimental psychology between the 1950s and 1960s. Because of this interest, there have been a number of important norming studies that are exemplified by Kucera and Francis's (1967) printed word frequency norms (counts of the frequency of occurrence of a given word in selected visual media), Noble's (1952) meaningfulness norms (a metric of the number of meanings generated to a given target), and Osgood, Suci, and Tannenbaum's (1957) semantic differential, among many others. In addition, there has been considerable recent interest in quantifying the characteristics of sublexical units, such as bigrams, trigrams, lexical neighbors, and others. Clearly, the centrality of the lexical unit in pattern recognition models is due in part to the efforts devoted to defining the stimulus.

Finally, because words are relatively well characterized patterns, they have been the focus of development of formal mathematical models of pattern recognition. For example, one of the first formal models in cognitive psychology was the Selfridge and Neisser (1960) Pandemonium model of letter recognition. Moreover, the interactive activation framework developed by McClelland and Rumelhart (1981) was central to nurturing the current widespread interest in formal connectionist models of cognitive performance. As we shall see below, word-level analyses appear to be ideally suited for the parallel distributed processing framework.

In sum, word recognition research has taken on so much importance because words are relatively well defined minimal units that carry many of the interesting codes of analysis (i.e., orthography, phonology, semantics, syntax) and processing distinctions (e.g., automatic vs. attentional) that have driven much of the work in cognitive psychology and psycholinguistics. Thus, although it would seem that the more important goal would be to pursue how individuals process language at higher levels such as clauses, sentences, and paragraphs, many researchers have pursued research at the level of the word because of its inherent tractability. As we shall see in the following review, although progress is being made, the ease of tracking the processes involved in word recognition may be more apparent than real.

II. CLASSIC VARIABLES OF INTEREST IN WORD RECOGNITION: WHAT DO WE KNOW?

In this section, I review some of the variables that have been pursued in word recognition research. First, I attempt to break the word down into smaller, more tractable bits. Second, I discuss the influence of variables that can be quantified at the whole word level, such as frequency, familiarity, meaningfulness, and contextual availability. Third, I provide an overview of the priming literature. Sprinkled within each of these sections is discussion of the major theoretical models and issues.

A. Breaking Down the Visual Word into Constituents

1. Features

One fundamental approach to pattern recognition is that a given pattern must first be broken down into features that are common to the set of patterns that one is interested in modeling. Some of the primary work in this area was developed by J. J. Gibson and Gibson (1955), who forcefully argued that feature-level analyses were an essential aspect of pattern recognition and, more generally, perceptual learning. The basic notion is that by identifying the primitive features, one has the building blocks for pattern recognition. This provided researchers with a well-specified problem: What are the primitive features used in letter recognition? The hunt was on!

Fortunately, it turns out that the feature analytic approach is ideally suited for letter recognition. Although there are differences across fonts, English orthography can be relatively well described by a limited set of features, such as horizontal lines, vertical lines, closed curves, open curves, intersections, cyclic redundancy, and others (see, for example, E. Gibson, Osser, Schiff, & Smith, 1963). Once researchers proposed such primitive features, both behavioral and neurological evidence began to accumulate that documented the role of such features in visual perception. On the behavioral side, there were studies of confusion matrices indicating that letters that shared features were more likely to be confused in degraded perceptual conditions, compared to letters that did not share many features (e.g., Kinney, Marsetta, & Showman, 1966). In addition, visual search studies by Neisser (1967), among others, indicated

that subjects were relatively faster to find a given target letter (e.g., Z) when it was embedded in a set of letters that did not share many features with the target (e.g., O, J, U, Z, D), compared to a set of letters that did share many features with the target (e.g., F, N, K, Z, X).

In addition to the behavioral evidence, there was also exciting evidence accumulating during the same period that appeared to identify neural substrates that might subserve feature-like detection processes. Consider, for example, the pioneering work by Hubel and Wiesel (1962, 1968). In this work, Hubel and Wiesel used single cell recording techniques to investigate neural activity in areas of the striate cortex in alert cats. When different stimuli were presented to the retina of the cat, there were increases in neural activity in specific cortical areas. In fact, Hubel and Wiesel found evidence that there were cells that appeared to be especially sensitive to visual stimuli that mapped onto such things as vertical lines, horizontal lines, angles, and even motion. The importance of this work is very simple: It provided the neurological evidence that converged with the notion that pattern recognition ultimately depends on primitive feature analytic processes. More recent work by Petersen, Fox, Snyder, and Raichle (1990) using positron emission tomography has extended this work to humans in demonstrating significant blood flow changes in specific areas of the striate cortex corresponding to feature-like detection systems in alert humans.

One of the first well-specified feature analytic models of letter recognition was developed by Selfridge (1959; Selfridge & Neisser, 1960). This model is displayed in Figure 1. The basic notion is that when a stimulus is displayed it briefly resides in an iconic representation (referred to metaphorically as an Image Demon). A set of 28 feature decoders that are sensitive to specific features (Feature Demons) begin to analyze the iconic representation. For example, as shown in Figure 1, one can see that the first circle is blackened indicating that the iconic image entails one vertical line. (An additional circle would be blackened with each additional feature along a given dimension.) These letter units then feed or activate a cognitive decision system. The more consistent the feature analyses are with a given letter, the greater the activation for that letter representation. The decision system simply picks the letter unit with the greatest activation (i.e., the Cognitive Decision Demon "listens" to determine which Letter Demon is shouting the loudest).

An important aspect of the Pandemonium model is that it entailed some capacity for learning. For example, during the learning phase, the system first maps which features correspond to which letters by storing the results of these feature tests. Moreover, if a particular feature is especially important for discriminating between letters, then the weights associated with (i.e., the importance of) this feature might be incremented. In this way, the Pandemonium model can extract the critical features that are most powerful in discriminating among letters. As we shall see, the Pandemonium model predates by some 20 years some of the important letter recognition and word recognition models that have more recently been developed. It is quite amazing that the Pandemonium model worked so well given the computational hardware limitations of the late 1950s and early 1960s.

Although most models of word recognition assume a first step of primitive feature identification, there are still many unresolved questions in this initial stage of processing. First, what is the glue that puts the features together? Specifically, once vertical lines, horizontal lines, and intersections have been

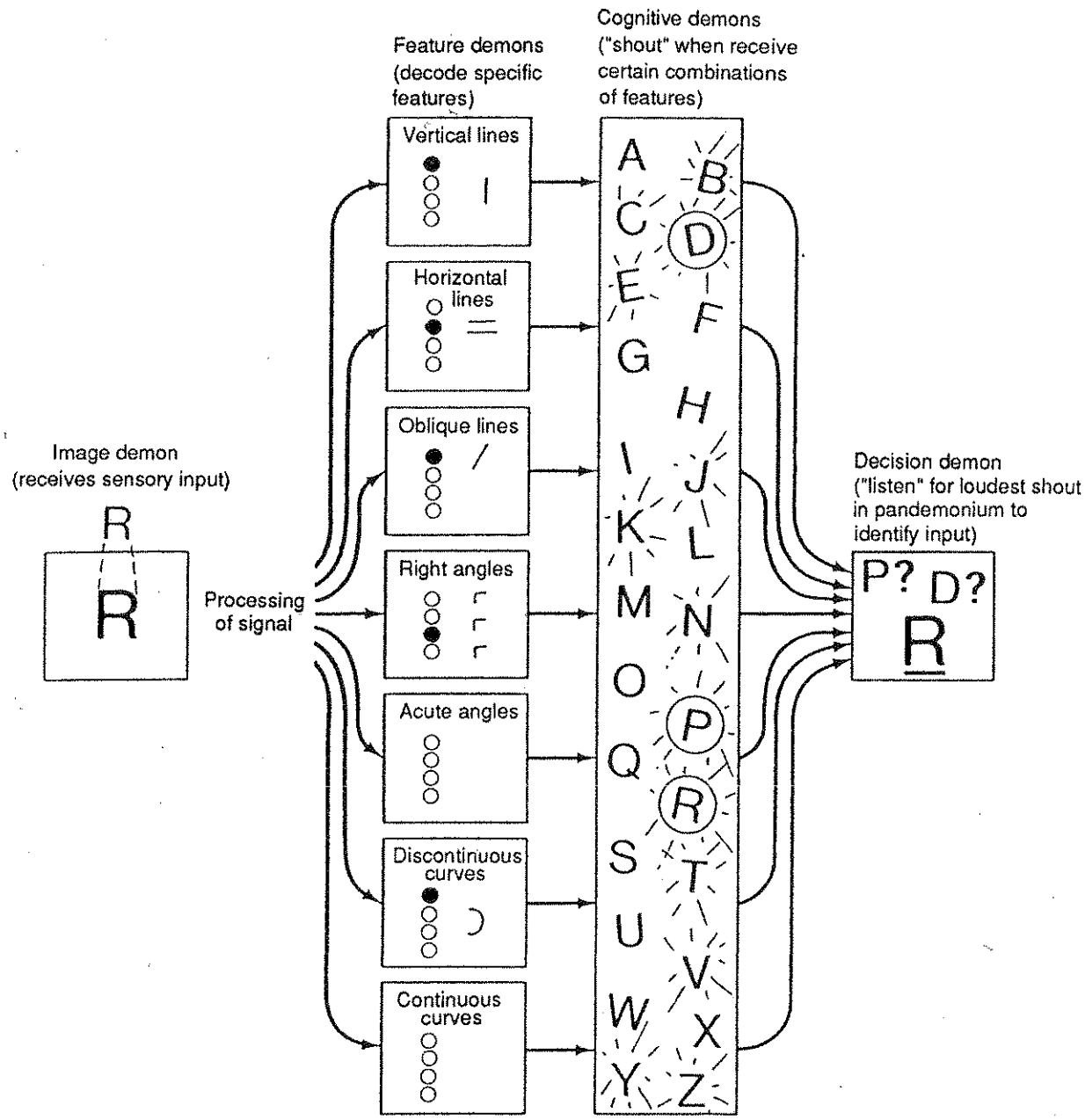

FIG. 1 Selfridge and Neisser's (1960) Pandemonium model of letter recognition. [From S. Coren, L. M. Ward, & J. T. Enns (1994). *Sensation and Perception, 4th Ed.* Ft. Worth: Harcourt Brace.

detected, how does one put the features together to identify a letter, e.g., the letter R? Obviously, we do not perceive free-floating features. Second, what happens in the feature analytic models when distortions occur that modify the feature (i.e., a 60° rotated vertical line is no longer a vertical line)? Third, and along the same lines, what are the critical features when the letters are distorted via different fonts or a novel style of handwriting? Reading still proceeds in an acceptable fashion even though there are considerable changes in the critical set of features (see Manso de Zuniga, Humphreys, & Evett, 1991, for a recent

discussion of reading handwritten text). Fourth, are features across letters coded serially in reading, e.g., from left to right in English orthography, or is there a parallel coding of features? Based on the work by Treisman (1986), one might expect that there is an early parallel coding of features that is followed by a more capacity-demanding conjunction process (however, see G. L. Shulman, 1990). Although there is little evidence that necessitates the application of parallel coding of features to reading, most models of word recognition appear to support a parallel registration of features in letter recognition. Finally, are features within letters the critical level of analysis in word recognition or are there supraletter or even word-level features (e.g., Purcell, Stanovich, & Spector, 1978)? Thus, although feature analyses are the first step in most of the models of word recognition, it should be clear that a considerable number of questions still need to be answered in mapping features onto letters. However, in lieu of being bogged down in some of the more important fundamental aspects of visual perception, let's assume we have made it to the letter. Surely, things must get a bit more tractable there.

2. *Letters*

Assuming that features are critical in letter recognition, and letters are crucial in word recognition, then one might ask what variables are important in letter recognition. For example, does the frequency of a given letter in print influence its perceptibility? Fortunately, there seems to be a relatively straightforward answer to this question. Appelman and Mayzner (1981) reviewed 58 studies that entailed 800,000 observations from a variety of paradigms that spanned 100 years of research. The conclusion from their review is very straightforward: Letter frequency does appear to influence speeded tasks such as letter matching, naming, and classification tasks (e.g., is the letter a vowel or a consonant?). However, letter frequency does not appear to influence accuracy in perceptual identification tasks. The results from the Appelman and Mayzner study are intriguing for three reasons: First, a priori, one would clearly expect that frequency of ANY operation (perceptual, cognitive, and motoric) should influence performance, and hence it is unclear why there is not a letter frequency effect in identification tasks. Second, as we shall see below, there is a consistent word-level frequency effect in both response latency tasks and perceptual identification tasks, and hence there at least appears to be a difference between frequency effects at different levels within the processing system, that is, letters versus words. Third, this is our first exposure of a general theme that runs across the word recognition literature, that is, different tasks that a priori should tap the same level of analyses often yield different patterns of data.

Another question that one might ask regarding letter-level influences is whether there is a word length effect in word recognition tasks, as measured by the total number of letters in a given word. Obviously, if the letter is a crucial player in word recognition, then one should find consistent effects of letter length in word recognition tasks. Interestingly, there has been some disagreement on this topic. Although there is clear evidence that longer words take more time in perceptual identification (McGinnies, Comer, & Lacey, 1952) and pronunciation (Forster & Chambers, 1973) and produce longer fixation durations in reading (see Just & Carpenter, 1980), there has been some conflict-

ing evidence regarding the lexical decision task (Henderson, 1982). In the lexical decision task, subjects are required to decide as quickly as possible whether a letter string is a word or nonword, with response latency and accuracy being the major dependent measures. Because the lexical decision task has been taken as a premier task for developing word recognition models, this is a troublesome finding. Interestingly, however, Chumbley and Balota (1984) reported relatively large length effects in the lexical decision task when the words and nonwords are equated on length and regularity. It is possible that inconsistent results with respect to past word length studies using the lexical decision task may have been due to subjects relying on any available dimensions that may be contaminated with length to bias their decisions in the lexical decision task. Thus, it appears safe to conclude at this point that there are consistent, albeit sometimes small, word length effects in virtually all word recognition tasks, when other variables are controlled.

More intriguing questions regarding letter recognition date back to questions that were originally posed by Cattell (1885). The interest here is basically, What is the perceptual unit in word recognition? A priori it would seem obvious that the letter should be the primary unit of analysis in visual word recognition, that is, words are made up of letters. However, Cattell (1885, 1886) reported a remarkable finding that was initially viewed as inconsistent with this notion. Cattell found that some words can be named more quickly than single letters. The problem this finding posed was very simple: How could the letter be the critical unit of analysis in word recognition, if words could be named more quickly than the letters that presumably make up the words? Along with the Cattell results, it was also reported that the exposure duration necessary to identify a word was in some cases less than the exposure duration necessary to identify a single letter. In fact, Erdmann and Dodge (1898) reported that the exposure duration necessary to identify four to five letters in a display was sufficient to read single words that could contain as many as 22 letters. Again, if words can be better perceived than letters, then how could letters be the basic unit of perception, since words are made up of letters?

Of course, an alternative account of this pattern of data is simply that subjects can use any available information regarding orthographic redundancy and lexical-level information to facilitate letter processing, and such information is unavailable when single letters are presented. This view was labeled the SOPHISTICATED GUESSING account of some of the initial findings. However, because of a seminal study by Reicher (1969), it appeared that there was more to this phenomenon than simply sophisticated guessing. In Reicher's study, on each trial, one of three stimuli was briefly flashed (e.g., a single letter, K, a word, WORK, or a nonword, OWRK), after which a patterned mask was presented. After the mask was presented, subjects were presented with two letters (e.g., D and K) adjacent to the position of the previous target letter for a forced-choice decision. The remarkable finding here is that subjects produced reliably higher accuracy when the first stimulus was a word than when it was a single letter or a nonword. Because both the letters D and K produce acceptable words within the WOR context, subjects could not rely on preexisting lexical knowledge to bias their response one way or the other (but see Krueger & Shapiro, 1979; Massaro, 1979, for an alternative view). This finding was termed the word superiority effect and was also reported in a study by Wheeler (1970).

There were two important subsequent findings that constrained the interpretation of the word superiority effect. First, the effect primarily appears under conditions of patterned masking (masks that involve letter-like features) and does not occur under conditions of energy masking (masks that involve high-luminous contrasts, e.g., Johnston & McClelland, 1973; Juola, Leavitt, & Choe, 1974). In fact, it appears that the interfering effect of the mask is primarily on performance in the letter alone condition and does not produce much of a breakdown in the word condition (Bjork & Estes, 1973). Second, letters are also better recognized when presented in pronounceable nonwords (e.g., MAVE), compared to unpronounceable nonwords or alone (e.g., Carr, Davidson, & Hawkins, 1978; McClelland & Johnston, 1977). Thus, the word superiority effect does not simply reflect a word-level effect.

The importance of the word superiority effect derives not only from the information that it provides about letter and word recognition, but also from its impact on the level of modeling that researchers began to use to influence their theory development. Specifically, this effect led to the development of a quantitative model of word and letter recognition developed by McClelland and Rumelhart (1981; Rumelhart & McClelland, 1982; also see Paap, Newsome, McDonald, & Schvaneveldt, 1982). As noted earlier, this type of modeling endeavor set the stage for the explosion of interest in connectionist models of cognitive processes (e.g., McClelland & Rumelhart, 1986; Rumelhart & McClelland, 1986; Seidenberg & McClelland, 1989).

Figure 2 provides an overview of the architecture of the McClelland and Rumelhart (1981) model. Here one can see three basic processing levels: feature detectors, letter detectors, and word detectors. These levels are attached by facilitatory (*arrowed lines*) and/or inhibitory (*knobbed lines*) pathways. As shown in Figure 2, there are inhibitory connections within the word level and within the letter level. Very simply, when a stimulus is presented, the flow of activation is from the feature level to the letter level and eventually onto the word level. As time passes, the letter-level representations can be reinforced, via the facilitatory pathways, by the word-level representations and vice versa. Also, as time passes, within both the letter- and word-level representations, inhibition from highly activated representations will decrease the activation at less activated representations, via the within-level inhibitory pathways.

How does the model account for the word superiority effect? The account rests heavily on the notion of cascadic processes in the information processing system (see Abrams & Balota, 1991; Ashby, 1982; McClelland, 1979). Specifically, a given representation does not necessarily need to reach some response threshold before activation patterns can influence other representations, but rather, there is a relatively continuous transferal of activation and inhibition across and within levels as the stimulus is processed. Consider the letter-alone condition in the Reicher paradigm, described earlier. When a letter is presented, it activates the set of features that are consistent with that letter. These featural detectors produce activation for the letter detectors that are consistent with those features and inhibition for the letter detectors that are inconsistent with those features. Although there is some activation for words that are consistent with the letter and some inhibition for words that are inconsistent with the letter, this effect is relatively small because there is little influence of a single letter producing activation at the word level. Now, consider the condition

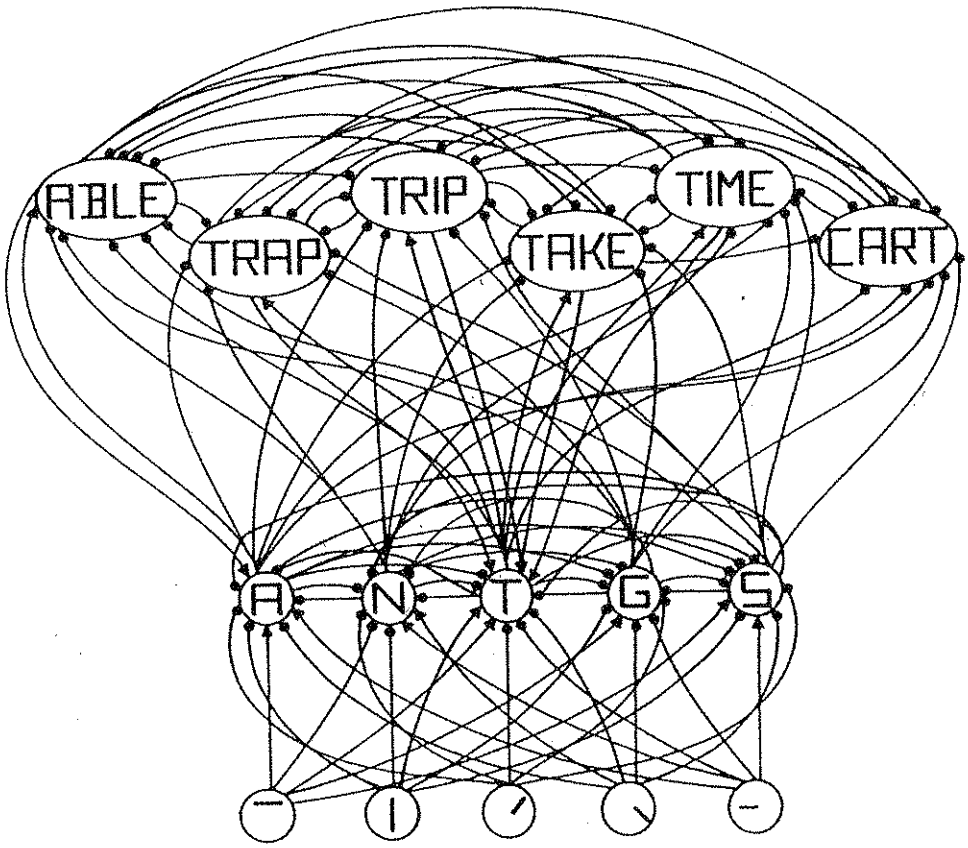

FIG. 2 McClelland and Rumelhart's (1981) interactive activation model of letter recognition. Copyright (1981) by the American Psychological Association. Reprinted by permission.

wherein the letter is embedded in a word context. In a word context there is now sufficient partial information from a set of letters to influence word-level activation patterns, and this will produce a significant top-down influence onto letter-level representations, that is, increase activation for consistent letters and decrease activation for the inconsistent letters. It is this higher level activation and inhibition that overrides the deleterious influence of the patterned mask.

In passing it is worth noting here that there is also evidence by Schendel and Shaw (1976) that suggests that features (e.g., lines) are better detected when the features are part of a letter than when presented alone. Hence, it is possible that there is also a letter superiority effect. This could easily be accommodated within the McClelland and Rumelhart model by assuming that there are also top-down influences from the letter level to the feature level.

Of course, one might ask at this point how this model could handle the pseudoword superiority effect. That is, letters are also better detected when embedded in pronounceable nonwords than when embedded in unpronounceable nonwords (Baron & Thurston, 1973; Carr et al., 1978) or presented in isolation (e.g., Carr et al., 1978; McClelland & Johnston, 1977). When letters are embedded in pronounceable nonwords, it is likely that there will be some overlap of spelling patterns between the pseudoword and acceptable lexical entries. For example, the pronounceable nonword MAVE activates 16 different four-letter words that share at least two letters within the McClelland and

Rumelhart network. Thus, the influence of orthographic regularity appears to naturally fall out of the interaction across multiple lexical entries that share similar spelling patterns within the language. As we shall see below, the influence of orthographic regularity on word recognition performance has been central to many of the more recent developments in word recognition research.

Although the influence of orthographic regularity appears to fall from this model, there are also some important limitations of orthographic regularity within the McClelland and Rumelhart model. Consider, for example, the impact of bigram frequency. For example, the vowel pair *ee* occurs in many more words than the cluster *oe*. The available evidence indicates that there is relatively little impact of bigram frequency on letter recognition within a Reicher-type paradigm (Manelis, 1974; McClelland & Johnston, 1977; Spoehr & Smith, 1975). McClelland and Rumelhart have successfully simulated this finding within their interactive activation framework. Although high-frequency letter clusters are more likely than low-frequency letter clusters to activate many word-level representations, this activation will be compensated by the fact that there will also be more word-level inhibition across those activated representations. Because, as noted above, there are influences of the number of lexical representations that share more than two letters, the lack of an influence of bigram frequency would appear to indicate that there may be a critical limit in the amount of overlap across lexical representations that is necessary to overcome the deleterious effects of within-level inhibition.

At this point, it is worth noting that there was some preliminary evidence that bigram frequency could influence word identification performance. For example, Broadbent and Gregory (1968) found that there was a clear influence of bigram frequency in word identification, but only for low-frequency words. In particular, for low-frequency words, low bigram frequency items produced an advantage over high bigram frequency items (also see Rice & Robinson, 1975). McClelland and Rumelhart suggested that this pattern might fall from their model if one assumes output from the word level is driving performance in the perceptual identification task. The presence of many neighbors (as in the high bigram frequency condition) produces considerable confusion for those words that are relatively difficult to identify to begin with, namely, the low-frequency words.

Unfortunately, the influence of bigram frequency on lexical-level performance has not produced consistent results in the literature. For example, Gernsbacher (1984) and Jastrzembski (1981) both failed to find any influence of bigram frequency in the lexical decision task, and Gernsbacher argued that previous findings were due to either response biases in the perceptual identification tasks or possible confoundings with familiarity of the letter string. Recently, Andrews (1992) again failed to find any influence of bigram frequency in either lexical decision or pronunciation performance. As we shall see below, one question that arises from this apparent lack of an influence of bigram frequency is, Why are there influences of neighbors primarily when the neighbors share more than two letters?

In addition to bigram frequency, one might ask whether positional frequency influences letter recognition. Positional frequency refers to the probability that a given letter(s) will occur in a given position within a word. Mayzner and Tresselt (1965) tabulated the summed positional frequency for single letters,

bigrams, trigrams, tetragrams, and pentagrams (Mayzner et al., 1965a, 1965b, 1965c) across a set of 20,000 words. This metric should reflect the orthographic structure across words within a given language. In fact, one might expect influences of such a metric to fall quite nicely out of the McClelland and Rumelhart model. In fact, Massaro, Venezky, and Taylor (1979) reported evidence of a large impact of summed positional frequency within a Reicher-type paradigm. Their results indicated that both summed positional frequency and a rule-based metric of orthographic regularity (see discussion below) were found to influence letter recognition performance. Thus, at least at the level of letter recognition, there does appear to be an influence of positional letter frequency in a Reicher-type paradigm. Because letter position obviously must be coded in the McClelland and Rumelhart model, one might expect this effect to naturally fall from the combined facilitatory and inhibitory influences across lexical-level representations.

In sum, the interactive activation model provides a cogent quantitative account of what appears to be evidence of multiple levels within the processing system working in concert to influence letter recognition. A particularly important aspect of this model is the fact that "other" similar lexical-level representations appear to have an influence on the ease to recognize a given letter within a word. It appears that letter- or word-level representations do not passively accumulate information, as in a logogen-type model (see Morton, 1969), but letters and words appear to be recognized in the context of similar representations that either reinforce or diminish the activation at a given representation. We now turn to some discussion of the dimensions that define similarity in such networks.

III. GETTING FROM LETTERS TO WORDS: INFLUENCES OF SUBLEXICAL LEVELS OF ORGANIZATION

The journey from letters to words has been a central concern in word recognition models. Although there are many issues that are addressed in this area, one of the major theoretical issues has been the specification of the RULES that are useful in translating an orthographic pattern into an acceptable lexical/phonological representation. Unfortunately, as we shall see, such a translation process is far from easy in English orthography.

A. Specifying the Rules of Translation

One of the most evasive goals encountered in the analysis of English orthography is the specification of the FUNCTIONAL UNIT(S) of sublexical organization. An obvious spelling-to-sound mapping might involve a simple one-to-one correspondence between graphemic units (single letters or letter clusters) and phonemes. Obviously, such an analysis fails relatively quickly in English because some graphemes, like *ph,* can serve as one phoneme in words like *grapheme* and *phoneme,* and two phonemes in a word like *uphill.* Likewise, even single letters are quite ambiguous, such as the *c* in the words *cat* and *cider.* English orthography simply does not allow a one-to-one mapping of spelling to sound.

Although a simple mapping of spelling to sound may not work for all words, it is still possible that one may gain considerable insight into the vast majority of words via an analysis of the regularities in the orthography. Such an enterprise was undertaken in a number of large-scale studies of English orthography in the late 1960s and early 1970s (e.g., Haas, 1970; Hanna, Hanna, Hodges, & Rudorf, 1966; Venezky, 1970; Wijk, 1966). For example, Venezky coded the grapheme-to-phoneme correspondences across a set of 20,000 medium- to high-frequency words. Through an in-depth analysis of the consistency of grapheme-to-phoneme patterns, Venezky distinguished between two large classes of grapheme-to-phoneme correspondences. PREDICTABLE patterns are those which can be based on the regular graphemic, morphemic (minimal meaningful units, e.g., *redistribution* = *re* + *distribute* + *tion*), or phonemic features of the words in which they occur, whereas UNPREDICTABLE patterns are simply those patterns that do not appear to fit within any predictable class. The important question here is how many patterns are predictable when one considers similarities across words within the language. For example, some correspondences appear to be relatively invariant (predictable invariant patterns), as when the grapheme *f* always corresponds to the sound /f/ with the only exception being in the word *of*. On the other hand, other graphemes have many variations, each of which appear to be relatively predictable (predictable variant patterns). For example, the letter *c* most typically corresponds to the phoneme /k/ but corresponds to the phoneme /s/ in many words when it is succeeded by the letter *i, y,* or *e*.

As Henderson (1982) points out, there are a number of sublexical constraints within the grapheme-to-phoneme system in English, which are called phonotactic constraints. For example, because certain stop consonant sequences are not permissible in English (e.g., /bp/ and /pb/), whenever one is confronted with such a sequence of letters (e.g., *pb* or *bp*), the correspondence is such that the first phoneme is silent (e.g., *subpoena*). Thus, in this case, the phonological constraints of the language drive the grapheme-to-phoneme conversion of the spelling patterns. There also appear to be predictable constraints on the grapheme-to-phoneme mapping that are derived at the morphemic and syllabic levels. For example, the graphemic sequence *mb* corresponds to two separate phonemes when it segments syllables as in *ambulance* and *amber,* but only one phoneme at word-ending positions, as in *tomb* and *bomb*. Unfortunately, as Henderson points out, the situation becomes somewhat more complex when one considers that *mb* also only corresponds to one phoneme when it precedes inflectional affixes (e.g., *bombing*), but not when it precedes other morphemes (*bombard*). Moreover, there appear to be other constraints that are simply based on allowable grapheme-to-phoneme correspondences in particular positions within words. For example, the *ck* spelling pattern corresponds to the phoneme /k/, but the *ck* pattern does not occur at the beginning of words; in these later cases the *c* to /k/ correspondence or the *k* to /k/ correspondence occurs.

For demonstrative purposes, I have only touched on some of the problems that one encounters in attempting to understand the regularity of spelling-to-sound correspondences in English orthography. Although ultimately it may be possible to specify such grapheme-to-phoneme rules in English, it is noteworthy that even with the relatively complex rule systems developed by Venezky and others, Coltheart (1978) estimates that 10–15% of the words will still be

unpredictable, that is, irregular. Likewise, Wijk notes that about 10% of the words will not fit his *Regularized Inglish*. This may be an underestimate, because, as Henderson points out, of the most common 3000 words, as many as 21% violate Wijk's regularization rules. More important, even if one had a fully developed rule-based system of spelling to sound in English, this would not necessarily indicate that such a rule-based system is represented in readers of English. In fact, even if such a rule-based system were represented, this would not be sufficient evidence to indicate that such rules play a role in fluent word recognition. Hence, instead of providing a detailed discussion of the enormously complex rule systems that have been developed, the present discussion focuses on the empirical evidence regarding how readers use sublexical information in word recognition. The interested reader is referred to Henderson (1982), Wijk (1966), and Venezky (1970) for excellent treatments of the search for rule-based translations of spelling-to-sound in English.

B. If Not Rules, Then What? The Controversy Regarding Dual-Route and Single-Route Models of Pronunciation

If it is unlikely that there will be a limited number of rules that specify the translation from spelling to sound in English (i.e., an ASSEMBLED route), it would appear likely that there is a second route (the LEXICAL or DIRECT route) to recognize words. In the second, lexical, route the reader may simply map the orthographic string onto a lexical representation and then access the programs necessary for pronouncing a given word aloud. Hence, we have the dual route model of word recognition (Coltheart, 1978; Humphreys & Evett, 1985), in which word pronunciation can be performed either by assembling phonology from orthography based on regularities within the language (e.g., as captured in the rules by Venezky, 1970) or by directly accessing a lexical representation via the whole-word orthographic input.

It is important to note here that because orthographies differ with respect to the regularity of spelling-to-sound correspondences, orthographies may also differ with respect to the weight placed on the assembled and lexical routes. For example, if the alphabetic system in a given language is unequivocal in mapping orthography to phonology, as in a language such as Serbo-Croatian, then one might find little or no influence of the lexical route in speeded pronunciation performance (Frost, Katz, & Bentin, 1987). The reader can rely totally on the assembled route, because it always produces the correct response. However, in English, and even to a greater extent in other languages such as Hebrew (e.g., Frost et al., 1987), the mapping between orthography and phonology is far less transparent. Hence, one should find increasing lexical effects in speeded pronunciation performance as one decreases the transparency of the spelling-to-sound correspondences (also referred to as orthographic depth). In support of this prediction, Frost et al. have reported larger frequency and lexicality effects in Hebrew compared to English, which in turn produced larger effects compared to Serbo-Croatian. Thus, comparisons across orthographies that differ with respect to the regularity of the spelling-to-sound correspondence support the notion that two routes are more likely in languages that have relatively deep orthographies.

If the inadequacy of a rule-based system demands a lexical route in English orthography, then one might ask what evidence there is for a role of an assem-

bled route. Why would subjects ever use an assembled route to name a word aloud if, by necessity, there must be a lexical route? One piece of evidence that researchers originally identified as being consistent with an assembled route is the relative ease with which individuals can name nonwords aloud. Because nonwords do not have a direct lexical representation, it would appear that a nonlexical route is necessary for naming nonwords. However, this piece of evidence was soon disabled by evidence from activation synthesis approaches (e.g., Glushko, 1979; Kay & Marcel, 1981; Marcel, 1980), in which the pronunciation of a nonword could be generated by the activation of similarly spelled words. In fact, activation synthesis theorists have argued that pronunciation performance is always generated via analogies to words represented in the lexicon (mental dictionary), thus denying any important role for the assembled route.

However, there is a second, and more powerful, line of support for the role of an assembled route in English. This evidence is provided by studies of acquired dyslexics, who produce a double dissociation between the two routes. Specifically, one class of dyslexics, SURFACE DYSLEXICS, appear to have a selective breakdown in the lexical route but have an intact assembled route. These individuals are likely to regularize irregular words and exception words, e.g., they might pronounce *broad* such that it rhymes with *brode* (e.g., Marshall & Newcombe, 1980; McCarthy & Warrington, 1986; Shallice, Warrington, & McCarthy, 1983). A second class of acquired dyslexics, DEEP (phonological) DYSLEXICS, appear to have an intact lexical route but an impaired phonological route. These individuals can pronounce irregular words and other familiar words that have lexical representations; however, when presented a nonword that does not have a lexical representation, there is considerable breakdown in performance (Patterson, 1982; Shallice & Warrington, 1980). The argument here is that phonological dyslexics have a selective breakdown in the assembled route.

Although it would appear that the basic tenets of dual-route models are well established in the literature, an intriguing alternative single-route connectionist model has been developed by Seidenberg and McClelland (1989) that does an excellent job of handling some of the major findings that were originally viewed as strong support for the dual route model. This model could be viewed as a second generation of the original McClelland and Rumelhart (1981) model of letter recognition described above. One of the major differences between the two classes of models is that the later Seidenberg and McClelland model was specifically developed to account for lexical tasks such as word pronunciation and the lexical decision task, whereas the McClelland and Rumelhart model was developed in large part to account for letter recognition performance. A second major difference between the two models is that the McClelland and Rumelhart model involves localized representations for the major processing codes (i.e., features, letters, and words), whereas the Seidenberg and McClelland model involves distributed representations; that is, there is not a single representation that reflects the word *dog*. A third difference is that the McClelland and Rumelhart model assumes the existence of a specific architecture (i.e., sets of features, letters, and words along with the necessary connections), whereas the Seidenberg and McClelland model attempts to capture the development of the lexical processing system via the influence of a training regime. However, given these differences, both models account for performance

by assuming a flow of activation across a set of relatively simple processing units and have been detailed sufficiently to allow for mathematical tractability. We now turn to a brief discussion of the Seidenberg and McClelland model.

As shown in Figure 3, the Seidenberg and McClelland model involves a set of input units that code the orthography of the stimulus and a set of output units that represent the phonology entailed in pronunciation. All the input units are connected to a set of hidden units (units whose only inputs and outputs are within the system being modeled, i.e., no direct contact to external systems, see McClelland & Rumelhart, 1986, p. 48), and all the hidden units are connected to a set of output units. The weights in the connections between the input and hidden units and the weights in the connections between the hidden units and phonological units do not involve any organized mapping before training begins. During training, the model is given an orthographic string, which produces some phonological output. The weights connecting the input and output strings are adjusted according to the back-propagation rule, such that the weights are adjusted to reduce the difference between the correct pronunciation and the model's output. During training, Seidenberg and McClelland presented the model with 2897 English monosyllabic words at a rate that is proportional to their natural frequency in English. The exciting result of this endeavor is that the model does a rather good job of producing the phonology that corresponds to regular words, high-frequency exception words, and even nonwords that were never presented. Although there is clearly some controversy regarding the degree to which the model actually captures aspects of the data (see e.g., Besner, 1990; Besner, Twilley, McCann, & Seergobin, 1990), the fact that it provides a quantitative account of aspects of simple pronunciation performance (without either explicit Venezky-type rules or even a lexicon) is quite intriguing and presents a powerful challenge to available word-recognition models.

One of the more important results of the Seidenberg and McClelland model is its ability to capture the frequency by regularity interaction. This finding was initially viewed as rather strong support for a dual-route model (cf. Andrews, 1982; Monsell, Patterson, Graham, Hughes, & Milroy, 1992; Paap & Noel, 1991; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). The interaction is as follows: For high-frequency words, there is very little impact of the correspondence between orthography and phonology, whereas for low-frequency words there is a relatively large impact of such a correspondence. The dual-route

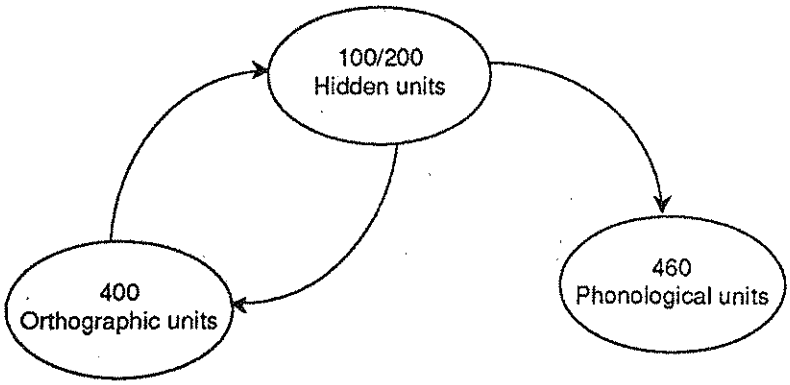

FIG. 3 Seidenberg and McClelland's (1989) implemented connectionist architecture. Copyright (1989) by the American Psychological Association. Reprinted by permission.

model accommodates this finding by assuming that for high-frequency words the lexical route is faster than the assembled route, and hence any inconsistent information from the assembled route does not arrive in time to compete with the pronunciation that is derived from the lexical route. For example, the incorrect assembled pronunciation for the high-frequency word *have* (such that it rhymes with *gave*) should not arrive in time to compete with the fast and correct lexical pronunciation. However, if one slows up the lexical route by presenting a low-frequency word, then one finds that the assembled output has time to interfere with the lexically mediated route, and hence response latency is slowed down. The important point for the dual route model is that the output of a low-frequency lexically mediated response can be inhibited by the availability of phonological information that is produced via the assembled route.

Although the dual route model would appear to provide a reasonable account for this interaction, this pattern also nicely falls from the Seidenberg and McClelland single route model. That is, the error scores (a metric that is mapped onto response latencies) for high-frequency regular words and exception words are quite comparable; however, for low-frequency words the error scores are worse for exception words than for regular words. Thus, one does not have to assume separate routes (or even a lexicon) to handle the frequency by regularity interaction, because this pattern naturally falls from the correspondences between the frequency of a particular spelling-to-sound correspondence even in a relatively opaque alphabetic system such as English. The interaction between frequency and regularity and the predictions from Seidenberg and McClelland's model are displayed in Figure 4.

Interestingly, the impact of the consistency of a given spelling-to-sound correspondence in English is also influenced by the characteristics of a given word's neighborhood. For example, Jared, McRae, and Seidenberg (1990) provide evidence that there are consistency effects in the pronunciation of regularly pronounced words (e.g., lint) primarily under conditions when the neighbors that have consistent spelling patterns (e.g., friends or mint) are higher in frequency than the neighbors that have inconsistent spelling patterns (e.g., enemies or pint). Such neighborhood frequency effects would appear to fall quite nicely from a connectionist model such as the Seidenberg and McClelland (1989) model. Alternatively, a rule-based model might suggest that the consistency of the neighbors defines the rules of translation from orthography to phonology. However, because of the difficulties noted above in specifying such rules, it is at the very least appealing that the Seidenberg and McClelland model is ideally suited to capture such neighborhood effects.

Although the Seidenberg and McClelland model does appear to provide an interesting alternative account to the dual-route model, there are still some important issues that need to be resolved. First, it is unclear how such a model might handle the fact that some acquired dyslexics appear to have only an intact assembled route while others appear to have only an intact lexical route (see Patterson, Seidenberg, & McClelland, 1989, for some discussion of this issue). Second, Monsell et al. (1992) have recently reported evidence that suggests that subjects can be biased by the presence of nonwords and exception words to rely on a more assembled route or a more lexical route. Third, there is intriguing evidence from Paap and Noel (1991) that provides evidence for differential attentional demands of the direct and assembled route, by indicating

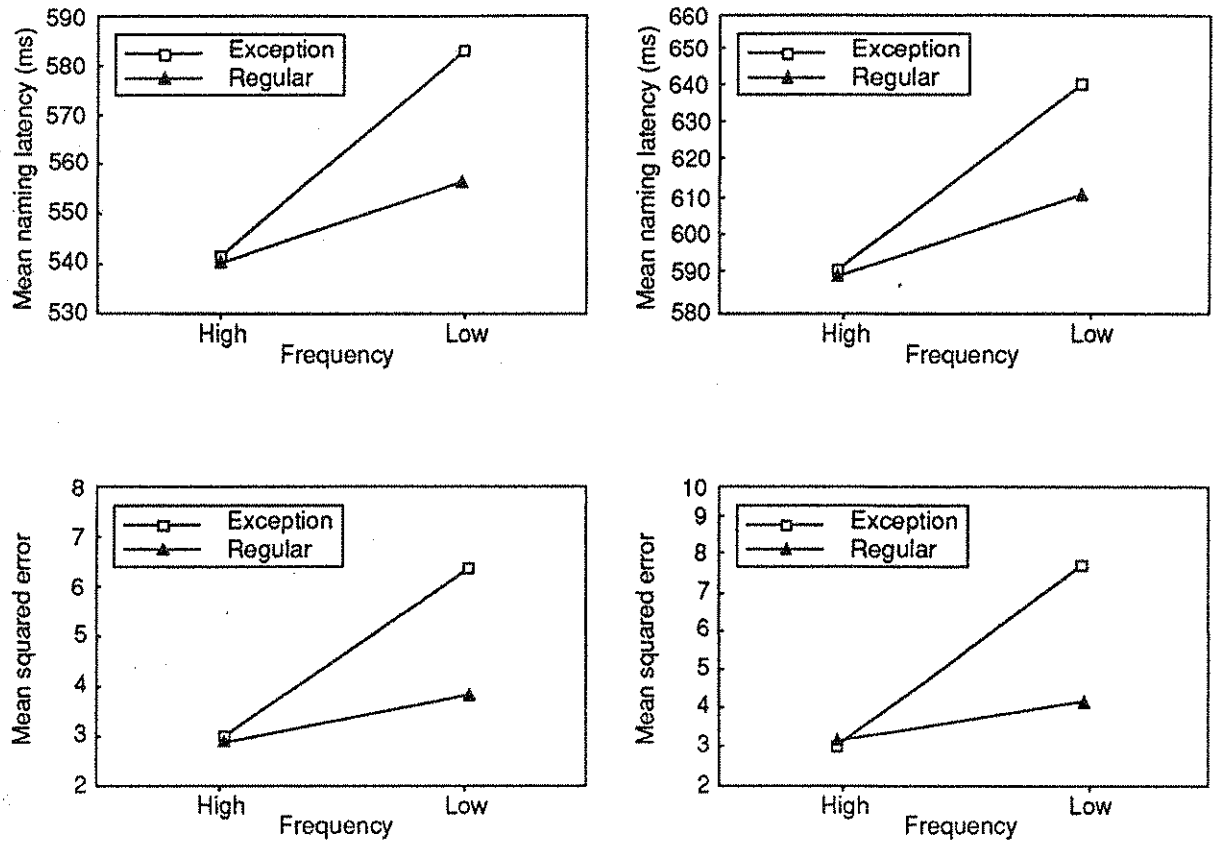

FIG. 4 Results and simulations of the (*left*) Seidenberg (1985) and (*right*) Seidenberg, Waters, Barnes, and Tanenhaus (1984, Experiment 3) studies: experimental results (*upper*) and simulations from the Seidenberg and McClelland (1989) model (*lower*). Copyright (1989) by the American Psychological Association. Reprinted by permission.

that a secondary attention-demanding task provides more interference with the assembled route than the direct route (also see Balota & Ferraro, 1993). Fourth, as described below, it appears that in some tasks, meaning-level representations can influence pronunciation and lexical decision performance. Thus, without some level of semantic input, it is unclear how an unembellished Seidenberg and McClelland model could account for such effects. Fifth, Besner (1990) and Besner, Twilley et al. (1990) have documented that the phonological error scores and the orthographic error scores do a rather poor job of simulating some characteristics of nonword performance. Of course, it is quite possible that future generations of the Seidenberg and McClelland type model (see, for example, Seidenberg & McClelland, 1990) will be able to handle these deficiencies. At this point, it seems most reasonable to acknowledge the accomplishments of this model and note the apparent problems with the current implementation. (The reader is also referred to Coltheart, Curtis, Atkins, & Haller, 1993, for a recent computational version of the dual route model.)

C. Types of Sublexical Stages of Analysis

At this point, it should be noted that I have yet to discuss specific types of sublexical but supraletter influences on word recognition. I have grouped to-

gether a set of effects under the "regularity" umbrella, even though there are a number of distinct classes of words that follow differing degrees of spelling-to-sound correspondences, e.g., words like *aisle* appear to have a unique spelling-to-sound correspondences, whereas, words like *gave* are pronounced consistently with the exception of a single high-frequency neighbor (*have*). I now turn to a brief discussion of three distinct levels of sublexical representation that have been at the center of this area of research: onsets and rhymes, morphology, and syllables.

1. Onsets and Rhymes

Treiman and her colleagues (e.g., Treiman & Chafetz, 1987; Treiman & Danis, 1988; Treiman & Zukowski, 1988) have argued that there is an intermediate level of representation in lexical processing between graphemes and syllables (also see Kay & Bishop, 1987; Patterson & Morton, 1985). They argue that syllables are not simply strings of phonemes, but there is a level of subsyllabic organization that is used in both speech production and recognition of visual strings. This subsyllabic distinction is between the onset and rhyme of a syllable. The onset of a syllable can be identified as the initial consonant or consonant cluster in a word. For example, /s/ is the ONSET for *sip*, /sl/ is the onset for *slip*, and /str/ is the onset for *strip*. The RHYME of a word involves the following vowel and any subsequent consonants. For example, in *sip*, *slip*, and *strip*, /ip/ would be the rhyme. Thus, syllables have a subsyllabic organization in that each syllable is composed of an onset and a rhyme.

Although our primary interest is in visual word processing, it is interesting to note that there has been evidence from a number of quite varied research domains that supports the distinction between onsets and rhymes in English. For example, there is evidence from the types of speech errors that speakers produce (MacKay, 1972), the distributional characteristics of phonemes within syllables (Selkirk, 1980), along with the types of errors that subjects produce in short-term memory tasks (Treiman & Danis, 1988).

In one of the studies addressing visual word recognition, Treiman and Chafetz (1987) presented strings like FL OST ANK TR to subjects with the task being to determine whether two of the strings in these four strings of letters could be combined to form a real word. In this case, one can see that FL and ANK can be combined to produce FLANK, with FL corresponding to the onset of the word *flank* and ANK corresponding to the rhyme. Now, consider performance on conditions where the strings again correspond to words but they are not broken at onsets and rhymes. For example, a subject might be presented FLA ST NK TRO. For these items, the correct answer is again FLANK, but now the FLA and NK do not correspond to onsets and rhymes. The results of the Treiman and Chafetz experiments indicated that anagram solutions were better when the breaks corresponded to onset-rhyme divisions compared to when the breaks did not. A similar pattern was found in a lexical decision task. In this study, the items were again presented such that there was either a break that matched the onset-rhyme division (e.g., CR//ISP, TH//ING) or a break that did not match the onset-rhyme division (e.g., CRI//SP and THI//NG). The results indicated that lexical decisions were reliably faster when the break matched the onset–rhyme division. Thus, Treiman

and Chafetz argued that onset and rhyme units play a role in visual word recognition.

Interestingly, Patterson and Morton (1985) and Kay and Bishop (1987) argue for a functional role of a variable, WORD BODY, that clearly resembles the rhymes. The word body involves the vowel and the following consonant. For example, Kay and Bishop report that there is an effect of spelling-to-sound correspondence of the word body especially for low-frequency words. For example, in the word *spook,* the phonological correspondence to the *oo* is actually the most popular correspondence to this grapheme. However, the sound corresponding to the higher level word body *ook* is actually quite unpopular among the neighbors (e.g., *book, cook, hook,* etc.). Kay and Bishop provided evidence that in speeded word pronunciation, consistencies in higher level word bodies have more of an impact than consistency in lower level grapheme-to-phoneme correspondences, primarily for low-frequency words. Of course, it would be useful to determine whether this pattern would also fall from the Seidenberg and McClelland (1989) model.

2. Syllables

If the distinction between onsets and rhymes plays a functional role en route to word recognition, then one would also expect a functional role for the syllable. At this level it is quite surprising that there has been considerable disagreement regarding the role of the syllable in visual word recognition. For example, Spoehr and Smith (1973) argued for a central role of the syllable, whereas Jared and Seidenberg (1990) have recently questioned the role of the syllable as a sublexical unit. In fact, as Seidenberg (1987) points out, there is even some disagreement regarding where syllabic boundaries exist. For example, according to Howard's (1972) rules that emphasize intrasyllabic consonant strings surrounding a stressed vowel, *camel* would be parsed as (*cam*) + (*el*), whereas, according to Selkirk's (1980) more linguistically based view that emphasizes the maximal syllable onset principle, *camel* would be parsed (*ca*) + (*mel*). Obviously, before one can address the functional role of the syllable in visual word recognition, one must have some agreement on how to parse words into syllables. Fortunately, for the vast majority of words there is agreement on how words are parsed into syllables.

The question here of course is whether a word like *anvil* is parsed as (*an*) + (*vil*) en route to word recognition. It should again be emphasized here that the concern is not whether subjects have access to syllabic information, surely they must; that is, most subjects can accurately decompose most words into syllables. The more important question is whether this information is used in accessing the lexicon for visually presented words.

Prinzmetal, Treiman, and Rho (1986) reported an important set of experiments that investigated the impact of syllabic structure on early-level perceptual operations in word recognition. These researchers used a paradigm developed by Treisman and Schmidt (1982) in which feature integration errors are used to examine perceptual groupings. The notion is that if a set of strings (e.g., letters or digits) forms a perceptual group, then one should find migration of features (colors) toward that group. In the Prinzmetal et al. study, subjects were presented with words such as *anvil* and *vodka*. At the beginning of each trial, subjects were given a target letter with the task being to report the color

of a target letter that would appear in the upcoming display. After the target letter was designated, subjects were presented a letter string with letters in different colors. The data of interest in such studies are the types of errors that subjects make as a function of syllabic structure. Consider the third letter position in the words *anvil* and *vodka*. In the word *anvil*, the third letter is part of the second syllable, whereas in the case of *vodka* the third letter is part of the first syllable. Now, if the syllable produces a perceptual grouping, then one might expect errors in reporting the colors such that the *d* in *vodka* might be more likely to be reported as the color of the *o*, compared to the *k*, whereas, the *v* in *anvil* might be more likely to be reported as the color of the *i*, compared to the *n*. This is precisely the pattern obtained in the Prinzmetal et al. study.

Seidenberg (1987) has questioned the conclusion that syllables are an access unit in visual word perception. He pointed out that there is a special level of orthographic redundancy at bigrams that breaks words into syllabic units. Specifically, Adams (1981) has noted that the letter patterns that often flank syllable patterns have relatively low bigram frequencies. In fact, the *nv* and *dk* are the lowest bigram frequencies in the words *anvil* and *vodka*. In general, if one considers relatively high frequency bisyllabic words, there appears to be a decrease in frequency of the bigrams that occur at syllabic boundaries. This bigram TROUGH may actually increase the likelihood of feature errors, due to the orthographic neighbors of the target instead of an actual subsyllabic parsing en route to word recognition.

In order to address this possibility, Seidenberg (1987, Experiment 3) looked at the probability of featural errors in the Prinzmetal et al. paradigm across words that have similar orthographic patterns and redundancy. Consider for example, the bisyllabic word *naive* and the monosyllabic word *waive*. If the syllable is an access unit that produces migration errors, then subjects should be more likely to misreport the color of the letter *i* as the color of the letter *v* than as the color of the letter *a* for *naive*, but this pattern should not occur for the monosyllabic word *waive*. Seidenberg found that similar feature migration errors were found for both bisyllabic words and monosyllabic words. This would suggest that the syllable was not used in access.

Interestingly, however, Rapp (1992) has recently reported an experiment that appears to question Seidenberg's conclusion regarding the influence of bigram troughs in the production of syllable effects. Rapp reported that illusory conjunctions followed syllabic constraints both for words that had a syllabic trough present (e.g., *anvil*) and words that do not have a trough present (e.g., *ignore*). Rapp argued from these data that merely the presence of a bigram trough is not sufficient to account for syllabic effects found in word recognition.

3. Morphemes

Another sublexical unit that has received considerable attention in the literature is the morpheme. One of the most compelling reasons that morphemes might play a functional role in word recognition is the generative nature of language. Rapp (1992) provides *chummily* as an interesting example. Although we may have never encountered the word *chummily*, we may assume that it means something like 'in a chummy way' or 'friendly' because it appears to have the morphological form *chummy* + *ly*. In fact, most linguistic models of lexical representation assume that there is some base form of representation and a set of rules that can be used to construct other forms of that item. The

present question is whether a given form of a word such as *jumped* is parsed as (*JUMP*) + (*ED*) in route to word recognition. As in the case of syllables, we are not questioning whether morphemes are somehow represented in the processing system; the question is whether morphemic analyses play a role in processes tied to visual word recognition.

Much of the theoretical and empirical work regarding the role of the morpheme in visual word recognition was originally developed by Taft and Forster (1975, 1976; also see Taft, 1979a, 1979b, 1985, 1987). They argue that readers first decompose polymorphemic words into constituent morphemes. Readers then access lexical files that are listed under the root morpheme. For example, if the word *characteristic* were presented, the reader would first access the root word *character*, and once this root word was accessed, the subject would search through a list of polymorphemic words with the same root morpheme, e.g., *characteristic, uncharacteristic, characterized, characteristically, uncharacteristically*, and so on.

As noted above, there have been a number of studies reported in the literature that support the notion that there is a morphemic level of analysis in visual word recognition. For example, Taft (1979a, 1979b) found an effect of printed word frequency of the root morpheme (the sum of frequencies of all words with a given root) in lexical decision performance for items that were equated in surface frequencies (see, however, caveats by Bradley, 1980). This would appear to support the contention that root morphemes do play a special role in word recognition and it is not simply the frequency of the actual lexical string that is crucial.

Possibly, a more compelling finding in the literature deals with long-term morphemic priming effects (Stanners, Neiser, & Painton, 1979). In these studies, subjects are most often presented a sequence of lexical decision trials. At varying lags within the sequence, subjects might be presented two forms of a given word with the same root. The interesting comparison is the influence of an earlier presentation of a given root form on later lexical decisions to the actual root. For example, if either *jump* or *jumped* is presented earlier in a lexical decision task, what impact does this presentation have on later lexical decision performance on the root form *jump*? Stanners, Neiser, Hernon, and Hall (1979) found that both *jump* and *jumped* equally primed later lexical decisions to *jump*. Presumably, subjects had to access *jump* to recognize *jumped*, and hence there was as much long-term priming from *jumped* as for the actual stem itself. Interestingly, Lima (1987) has found that mere letter overlap does not produce such an effect. For example, she has reported that *arson* does not prime *son*, but *dishonest* does prime *honest*. Thus, it does not appear that mere letter overlap is producing this long-term priming effect (see review by Feldman & Andjelkovic, 1992, for a summary of evidence favoring nonorthographic accounts of morphemic priming effects.)

I have painted a relatively simple picture of morphemic analyses in word recognition. However, as Rayner and Pollatsek (1989) have pointed out, the influence of morphological processing in word recognition tasks (assuming there is an effect) is far from simple. For brevity, we simply list some of the findings here: First, suffixes that produce inflections of a given word but keep the grammatical class (e.g., *jump* vs. *jumped*) produce full long-term repetition priming in lexical decisions, whereas suffixes that produce derivations of a given form that belong to a different grammatical class produce reduced priming

(e.g., *select* vs. *selective*) (Stanners, Neiser, Hernon, & Hall, 1979). Second, Stanners et al. found that there is reduced priming from verbs that have a less transparent relationship; for instance, there is reduced priming from *spoken* to *speak* compared to *jumped* to *jump*. Third, Bradley (1980) failed to provide evidence of root morpheme frequency or surface frequency effects for words that end in *-ION*. Thus, different types of morphemes apparently behave differently. Fourth, prefixed words (e.g., *rejuvenate*) produce faster lexical decision latencies (Taft, 1981) and shorter fixation durations in an on-line reading task (e.g., Lima, 1987), compared to words with prefix-like beginnings that are not actually prefixed stems (e.g., *repertoire*). If the parser first strips prefixes, then how would the parser know not to strip the prefix-like beginnings in nonprefixed stems? Fifth, it appears that morphological structure in nonwords (e.g., *walken* vs. *wilken*) is also accessed in the lexical decision task and produces interference in making nonword decisions (e.g., Caramazza, Laudanna, & Romani, 1988; Taft & Forster, 1976).

Seidenberg (1987) has also noted that there are often low-frequency bigram troughs at morphemic boundaries. He argues that prefixes are relatively common bigrams in the language, and hence the frequency of the bigram within a prefix is typically higher than the frequency of the bigram that straddles the prefix and the next morpheme. Hence, it is possible that at least some of the prefix-stripping effects that have been reported in the literature are due to the presence of a bigram trough at prefix boundaries. Again, one might expect such morphemic-like effects to naturally fall from a Seidenberg and McClelland type model, without a functional role for the morpheme directly represented.

Rapp (1992) addressed Seidenberg's concern about the presence of bigram troughs in producing some of the morphemic effects in word recognition research. She reported a lexical decision study in which the corresponding letters within a letter string were displayed in two colors. The colors of the letters either were consistent with morphological boundaries (e.g., *UNTIE*) or inconsistent with morphological boundaries (e.g., *UNTIE*). Moreover, she compared items that had bigram troughs present with items that did not. The results indicated that words that included color differences at morphological boundaries produced reliably faster lexical decisions compared to words that did not and this effect was not modulated by the presence of a bigram trough. Thus, to the extent that Rapp's lexical decision task reflects more natural processes en route to word recognition in reading, the influence of morphological analyses do not appear to be simply due to the presence of a bigram trough that overlaps morphological boundaries. Of course, one might also ask whether there are other measures of orthographic redundancy that may account for the apparent influence of morphological analyses en route to word recognition.

D. Summary of Sublexical but Supraletter Organization

In the preceding sections, I have provided an abbreviated overview of some of the sublexical structures that researchers have argued are engaged in the journey from print to sound. There is still considerable debate regarding the nature of the impact of these variables. As I have tried to emphasize, one of the major issues is whether the influence of sublexical structures naturally falls from interactions amongst many lexical representations, as in Seidenberg and

McClelland's (1989) connectionist model, or whether these variables fall from direct parsing of words into these sublexical structures. Before turning to further discussion of this issue, it is important to complete our discussion of the empirical results that need to be accounted for by an adequate model of word recognition. It is in this light that we now turn to lexical-level variables.

IV. LEXICAL-LEVEL VARIABLES

By lexical-level variables, I simply refer to the influence of variables that have been quantified at the whole word level. For example, word frequency is a lexical variable. Specifically, a researcher can investigate the influence of the printed frequency of occurrence of a given word (e.g., *dog*) on word recognition task performance. Here, frequency is defined at the lexical level instead of the sublexical level.

A. Word Frequency

The frequency with which a word appears in print has a strong influence on word recognition tasks. Such effects have been observed in lexical decision performance (e.g., Forster & Chambers, 1973), pronunciation performance (e.g., Balota & Chumbley, 1984), perceptual identification performance (e.g., Broadbent, 1967), and on-line reading measures such as fixation duration and gaze duration measures (e.g., Inhoff & Rayner, 1986; Rayner & Duffy, 1986). This, of course, should not be surprising because printed word frequency should be related to the number of times one experiences a given word; and, experience with an operation should influence the ease of performing that operation.

The influence of word frequency in word recognition tasks has been accounted for by two rather broad classes of word recognition models (Forster, 1989). The ACTIVATION class of models is based in large part on Morton's (1969, 1970) classic logogen model and includes the more recent interactive activation and connectionist offspring, described above. For example, according to Morton's logogen model, frequency is coded via the resting level activations in word recognition devices (logogens). High-frequency words, because of the increased likelihood of experience, will have higher resting level activations than low-frequency words. Therefore, in order to surpass a word recognition threshold, the activation within such a logogen will need to be boosted by less stimulus information for high-frequency words than for low-frequency words. The second class of word recognition models is referred to as ORDERED SEARCH models (e.g., Forster, 1976, 1979; Rubenstein, Garfield, & Millikan, 1970). According to these models, the lexicon is serially searched with high-frequency words being searched before low-frequency words. For example, as shown in Figure 5, Forster (1976) has argued that the lexicon may be searched via several indexing systems: orthographic, phonological, or syntactic/semantic access bins. Each of these bins involves a frequency-ordered search, whereby high-frequency words are searched before low-frequency words, and once the target is located, the subject has immediate access to the word's master lexical representation. As one might expect, there are also hybrid models that include both activation and search characteristics, such as the Becker (1980), Paap et al. (1982), and Taft and Hambly (1986) models. For example, Becker suggests that

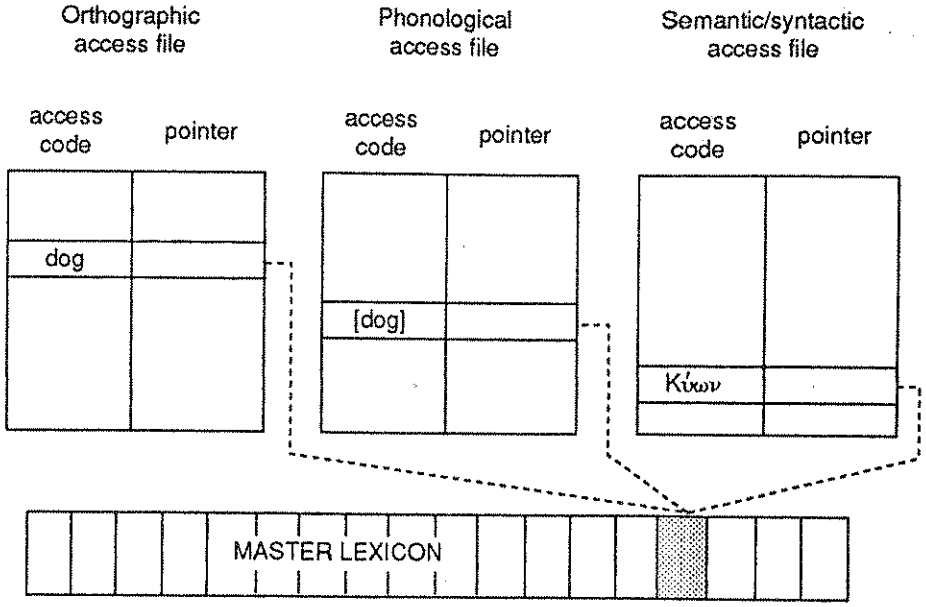

FIG. 5 Architecture of Forster's (1978) serial search model of word recognition.

activation processes define both sensorily and semantically defined search sets. These search sets are then compared to the target stimulus via a frequency-ordered search process.

An important question that has arisen recently regarding word frequency effects is the locus of the effect in the tasks used to build models of word recognition. All the above models attribute the influence of frequency to processes leading up to and including the magical moment of word recognition (see Balota, 1990). However, there is more recent evidence that suggests there are (a) decision components of the lexical decision task (Balota & Chumbley, 1984; Besner, Davelaar, Alcott, & Parry, 1984; Besner & McCann, 1987), (b) post-access components related to the generation and output of the phonological code in the pronunciation task (Andrews, 1989; Balota & Chumbley, 1985; Connine, Mullennix, Shernoff, & Yelen, 1990), and (c) sophisticated guessing aspects of the threshold identification task (Catlin, 1969, 1973) that are likely to exaggerate the influence of word frequency.

Consider, for example, the Balota and Chumbley (1984) model of the lexical decision task displayed in Figure 6. Balota and Chumbley have suggested that subjects will use any information that is available to discriminate words from nonwords in this task. Two pieces of information that are obvious discriminators between words and nonwords are the familiarity and meaningfulness (FM dimension) of the stimuli. Obviously, nonwords are less familiar and also less meaningful than words. However, both words and nonwords vary on these dimensions; in fact, the distributions may overlap (e.g., the nonword *chummingly* may be more familiar and meaningful than the low-frequency word *ortolidian*). Frequency effects in the lexical decision task may be exaggerated because low-frequency words are a bit more similar to the nonwords on the FM dimension than are high-frequency words. Hence, when there is insufficient information to make a fast "word" response, the subject is required to engage in an extra checking process (possibly checking the spelling of the word). This

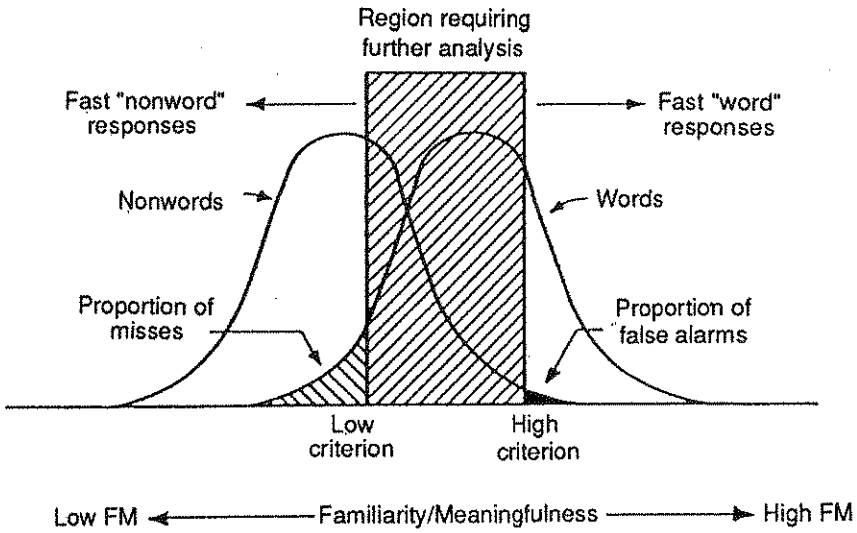

FIG. 6 Balota and Chumbley's (1984) two-stage model of the lexical decision task.

time-consuming extra checking process is more likely to occur for low-frequency words than for high-frequency words, thereby exaggerating any obtained influence of word frequency. (It is worth noting here that the two major implemented quantitative models of the lexical decision task, Ratcliff & McKoon, 1988, described below; and Seidenberg & McClelland, 1989, both make use of a familiarity dimension in modeling lexical decision performance.)

As noted, there has been considerable controversy in the literature regarding the locus of word-frequency effects in the tasks used to build word recognition models (e.g., see Balota & Chumbley, 1990; Monsell, Doyle, & Haggard, 1989; Savage, Bradley, & Forster, 1990). Of course, it is not surprising that some theorists have been reluctant to accept any concerns about the major tasks used to build models of word recognition. Moreover, some researchers have misrepresented the task analysis research as indicating that task analyzers were arguing that ALL frequency effects must be due to post-access processes. The primary intent of the task analysis work was to caution researchers that not all word-frequency effects can be unequivocally attributed to access processes in the tasks that are used to measure word recognition. Although a full discussion of this work is beyond the scope of the present review, it is sufficient to note here that there is little disagreement that word frequency influences some of the processes involved in word recognition, and most likely, this variable also influences other processes that are idiosyncratic to the tasks that are used to build models of word recognition. As exemplified throughout the word recognition literature, understanding the operations in the tasks used to build models of word recognition is a paramount first step in building adequate models.

B. Familiarity

A variable that is highly correlated with frequency is word familiarity. Some researchers have argued that the available printed word frequency norms by Kucera and Francis (1967) and Thorndyke and Lorge (1944) may not be the most sensitive estimates of the impact of frequency of occurrence on lexical

representations. For example, printed word frequency counts do not take into consideration spoken word frequency. Thus, some researchers have turned to subjective rated familiarity norms to index the frequency with which a given lexical item has been processed (e.g., Boles, 1983; Connine et al., 1990; Gernsbacher, 1984; Nusbaum, Pisoni, & Davis, 1984). In fact, Nusbaum and Dedina (1985) and Connine et al. have provided evidence that one can find familiarity effects above and beyond frequency effects in the lexical decision and pronunciation tasks, respectively. Moreover, Connine et al. provided evidence that ease of production in a delayed pronunciation task (a task used to bypass many of the operations involved in lexical access) can be predicted by familiarity ratings when word-frequency measures no longer predict performance. Hence, Connine et al. suggest that subjective rated familiarity incorporates frequency in production in addition to frequency in print.

Just as in the case of printed word frequency, there is no question whether subjective rated familiarity provides a strong predictor in word recognition tasks. The more important issue (e.g., Balota, Ferraro, & Connor, 1991) is what sorts of information subjects use to make untimed subjective familiarity ratings. For example, it is possible that when subjects rate a word for familiarity they are not only rating the printed string for familiarity, but rather they may be rating the availability of a clear and vivid meaning for that word, or possibly even the extent to which an individual can identify a specific context for that word (Schwanenflugel, Harnishfeger, & Stowe, 1988). In fact, one finds quite strong correlations between familiarity and other more semantic variables such as meaningfulness, concreteness, and contextual availability (see Balota et al., 1991, for a review). The point here is simply that although there may be clear difficulties with word-frequency norms, it is also unlikely that subjective familiarity ratings will provide a pure measure of the extent to which an individual has been exposed to a given stimulus word.

C. Neighborhood Effects

Seidenberg and McClelland (1990) have suggested that average adult readers probably have 30,000 words in their vocabulary. Because these words are based on a limited number of 26 letters, there must be considerable overlap in spelling patterns across different words. One of the major tasks of an acceptable model of word recognition is to describe how the system selects the correct lexical representation among neighborhoods of highly related orthographic representations. Of course, it is possible that the number of similar spelling patterns may not influence lexical processing and that only a single representation must pass threshold for recognition to occur. However, as already mentioned, it appears that words are not recognized in isolation of other orthographically related representations.

Coltheart, Davelaar, Jonasson, and Besner (1977) introduced the *N* metric. *N* refers to the number of words that could be generated by changing only a single letter in each of the positions within a word. Although similar to phonological neighbors, discussed above, orthographic neighbors can obviously be different. For example, for the target word *hate, late* would be an orthographic and phonological neighbor, *hale* would be an orthographic neighbor, and *eight* would be a phonological neighbor.

There are two factors that need to be accommodated when discussing neighborhood effects. First, one needs to consider the influence of neighborhood size. That is, some words are embedded in relatively large orthographic neighborhoods, whereas other words are embedded in relatively small orthographic neighborhoods. Andrews (1989, 1992) has repeatedly observed an interesting effect of neighborhood size. Specifically, in both pronunciation and lexical decision performance, low-frequency words that are from large neighborhoods produce faster latencies than low-frequency words from small neighborhoods, whereas there is little or no influence of neighborhood size for high-frequency words. (This interaction is reminiscent of the influence of phonological regularity and word frequency discussed above.)

Second, one might also expect the frequency of the neighbors to play a role in word recognition tasks. In fact, Grainger (1990) and colleagues (Grainger, O'Regan, Jacobs, & Segui, 1989; Grainger & Segui, 1990) have argued that neighborhood frequency is more important than neighborhood size. They argue that performance will be worse whenever the target word has at least one orthographically similar word that is higher in frequency. Because the likelihood of a higher frequency neighbor should increase with the size of the neighborhood, the Grainger results appear to be at odds with the Andrews results. Andrews (1992) points to two issues that may help resolve the apparent discrepancy: First, she points out that the difference between the size of the small and large neighborhoods was smaller in Grainger's studies than in Andrews' studies. Moreover, Andrews points out that Grainger has only found this inhibitory effect in lexical decision, and hence it is possible that this pattern may reflect decision processes that have been tied to this task.

It would seem that both frequency of neighbors and size of neighborhoods should play a role in word recognition tasks. In this light, it is useful to mention the P. A. Luce and Pisoni (1989) neighborhood activation model, which they applied to auditory word recognition performance. This model takes into consideration target frequency, neighbor frequency, and neighborhood size via R. D. Luce's (1959) choice rule. Specifically, the probability of identifying a stimulus word is equal to the probability of the stimulus word divided by the probability of the word plus the combined probabilities of the neighbors. Of course, it is possible that the neighborhoods of the neighbors may play a role along with the degree of overlap of the neighbors. At this level, it would seem that models that attempt to capture the interaction across many lexical representations that are partially activated, as in the McClelland and Rumelhart (1981) model and, in an indirect fashion, in the Seidenberg and McClelland (1989) model, are ideally suited to accommodate such neighborhood effects. Finally, it should also be noted here that an interesting and important aspect of Luce and Pisoni's work is that such neighborhood frequency effects appear to modulate performance differently across lexical decision, pronunciation, and threshold identification. Again, we find the importance of task analyses in defining the influence of a variable in word recognition.

Before leaving this section, it should be noted that neighborhood size effects have taken on considerable currency in discriminating between the search models of word recognition (e.g., Becker, 1980; Forster, 1976, 1989; Taft & Forster, 1976) and the activation models of word recognition (e.g., McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989). Neighborhood size effects

would appear to produce particular difficulties for serial search models. Specifically, the more items that need to be searched, the slower response latency should be. This is precisely opposite to the pattern reported by Andrews, who finds that larger neighborhoods produce faster response latencies, and only for low-frequency words. However, Seidenberg and McClelland (1989) have demonstrated that their model can nicely accommodate Andrews' effects of neighborhood size.

Although the Seidenberg and McClelland model can account for some influences of neighborhood density, there are still some limitations (see Andrews, 1992; Besner, Twilley et al., 1990a). For example, because frequency and neighborhood size are coded within the same network weights, it is unclear why neighborhood density would only have an effect for low-frequency words. Moreover, it is unclear why there are no neighborhood density effects for nonwords in the lexical decision task. Of course, there may be more to a word than the orthographic and phonological neighborhoods that apparently fall from the Seidenberg and McClelland model. It is in this light that we now turn to a discussion of the possibility that the MEANING of a word may play a role in word recognition.

V. SEMANTIC VARIABLES FOR ISOLATED WORDS

There have been a number of reports in the literature that indicate that semantic variables associated with lexical representations can modulate the ease of word recognition. This is an intriguing possibility because most models of word recognition would appear to indicate that the word must be recognized before the meaning of the word is determined. That is, the subject needs to recognize the string as the word *dog* before the subject accesses the meaning of the word *dog*. However, within an interactive activation model such as McClelland and Rumelhart's, one might envisage a higher level semantic representation above the word level that may accrue activation and provide top-down feedback in a cascadic fashion to influence word recognition performance (see Balota, 1990; Balota et al., 1991, for a further discussion of this issue).

Although there has been a considerable amount of work attempting to specify which semantic variables play a role in word recognition, much of this work has been open to alternative interpretations. Here we briefly review this work, emphasizing the major findings with respect to each of the major variables. Balota et al. (1991) have provided a more in-depth review of this literature.

A. Concreteness

If one could provide evidence that concrete words (those words which can be the object of a sense verb, e.g., *touch, see, hear,* etc.) were more easily recognized than abstract words, then this would provide some evidence of a semantic influence on word recognition performance. Although the concrete/abstract dimension has been the center of considerable research in word recognition (e.g., Bleasdale, 1987; Boles, 1983, 1989; Day, 1977; de Groot, 1989; James, 1975; Kroll & Merves, 1986; Paivio & O'Neill, 1970; Rubenstein et al., 1970; Winnick & Kressel, 1965), until recently, most of the results have been equivocal

because of potentially confounding variables. For example, variables such as familiarity of the letter string were not measured in most of the studies. Because there appears to be a relatively strong correlation between familiarity and concreteness, this factor could be responsible for at least some of the observed effects (see Schwanenflugel et al., 1988).

An interesting and consistent finding that has occurred in the literature addressing concreteness effects is that there appears to be an interaction between concreteness and word frequency. Specifically, concrete words produce faster lexical decisions compared to abstract words, primarily for low-frequency words but not for high-frequency words (de Groot, 1989; James, 1975; Kroll & Merves, 1986). Although this finding appears to be relatively consistent in the word recognition literature, Kroll and Merves have argued that it is possible that this interaction may be more a reflection of postlexical decision processes that are tied to the lexical decision task rather than task-independent word recognition processes. Again, we find concerns about the dimensions of the tasks used to measure word recognition.

It is in this light that a study by Bleasdale (1987) is quite intriguing. Bleasdale was interested in the influence of semantic context on word recognition processes tied to abstract and concrete words. Because we are interested in the influence of concreteness on isolated word recognition, we will focus on results from Bleasdale's neutral context condition. In this condition, either abstract or concrete words were preceded by the word *blank*. There is no a priori reason to suspect that such neutral primes should differentially influence word identification processes for abstract and concrete words. (Although it is at least possible that *blank* may provide a better prime for concrete words than abstract words; consider *blank letter* vs. *blank love*.) Bleasdale found consistent evidence for an advantage of concrete words over abstract words in pronunciation performance. Moreover, the stimuli that Bleasdale used appeared to be well controlled on relevant variables that might be related to concreteness, such as rated subjective familiarity. Thus, the Bleasdale study provides some of the best evidence available to support an effect of concreteness on word recognition.

B. Meaningfulness

A second semantic variable that might play a role in word recognition is the MEANINGFULNESS of the stimulus. Of course, meaningfulness can be defined in many ways. For example, one might simply look up the number of meanings that are listed under a given word in the dictionary. Jastrzembski (1981) used this metric and found that words possessing many (>10) meanings produced faster lexical decisions than words possessing few (<4) meanings. However, Gernsbacher (1984) demonstrated that number of dictionary meanings of a given word does predict lexical decision time when familiarity is confounded with meaningfulness but does not predict lexical decision time when familiarity is controlled. Moreover, it is also unclear whether subjects actually have representations for all meanings. For example, the word *fudge* has 13 entries listed with it. Although possible, it seems unlikely the subjects have more than four or five distinct meanings of the word *fudge* available (see Millis & Button, 1989, for further discussion of different metrics of meaningfulness).

A second approach to meaningfulness has been to compare words that have multiple distinct meaning representations and words that appear to only have a single meaning representation. Consider the homograph *organ*; one meaning refers to 'musical instrument' and the second meaning refers to 'bodily system'. Rubenstein et al. (1970; also see Rubenstein, Lewis, & Rubenstein, 1971) found that, compared to nonhomographs, homographs produced faster lexical decisions. Unfortunately, Clark (1973) closely scrutinized the Rubenstein results and found that most of the effect was produced by a few idiosyncratic items, and hence when one conducts item analyses instead of the traditional subject analyses, the effect disappears. More recently, however, Kellas, Ferraro, and Simpson (1988) also found that homographs produced faster lexical decisions than nonhomographs. This effect was reliable both by subjects and by items (F. R. Ferraro, personal communication) and hence is not susceptible to the criticism leveled by Clark regarding the Rubenstein et al. study. Moreover, the stimuli appeared to be well equated on potentially important confounding factors such as familiarity. Finally, we (Balota and Ferraro) have recently replicated this same effect with the Kellas et al. stimuli in a pronunciation task. Thus, the effect does not appear to be task-specific.

In sum, it appears that the number of meanings available for a word does modulate performance on word recognition tasks. This of course would be expected if there were a top-down impact from a meaning-level representation to a word-level representation. It should also be noted, however, that—although the results are encouraging—because of its rather checkered past, more work clearly needs to be done to firmly establish an impact of meaningfulness above and beyond other variables in word recognition tasks.

C. Contextual Availability

Recently, Schwanenflugel et al. (1988; also see Schwanenflugel & Shoben, 1983) have provided evidence that a variable referred to as contextual availability can produce an influence on isolated word recognition in lexical decision performance above and beyond influences of correlated variables such as concreteness, familiarity, length, and so on. Contextual availability refers to how easily a subject is able to think of a particular context or circumstance in which a given word might occur. As examples, Schwanenflugel et al. present the words *baseball* and *emotion,* both of which would be rated relatively high in contextual availability, and the words *inversion* and *sloop,* both of which would be rated relatively low in contextual availability. The results of three lexical decision experiments indicated that contextual availability reliably predicted lexical decision performance above and beyond familiarity, word frequency, and word length.

Of course, just as in the case of familiarity, one might ask what sorts of information subjects use in making untimed contextual availability ratings. Are subjects somehow relying on the "wordness" of the stimuli, and wordness is the dimension that is actually predicting lexical decision performance? In this light, it would be quite useful to determine whether contextual availability might also predict pronunciation performance. However, until such data become available, the role of contextual availability in isolated lexical deci-

sion performance would appear to provide at least some converging evidence that meaning can play a role in a task used to measure isolated word recognition.

D. Other Semantic Variables That Produce Effects in Isolated Word Recognition Paradigms

Because of space limitations, I shall only briefly mention a few other findings that would appear to indicate that meaning can have an early impact in word recognition performance. First, there is evidence that concreteness of a word can influence the time taken to generate an associate from that word (e.g., de Groot, 1989). Because subjects must recognize a word en route to generating an associate, this effect might be due to word recognition processes. Second, and along these same lines, Chumbley and Balota (1984) have found that the time taken to generate associates from one group of subjects can be used as a predictor of lexical decision performance for the same set of words when presented in isolation to a second group of subjects, above and beyond other related variables such as frequency, length, and so on. Third, Whittlesea and Cantwell (1987) found that providing meaning for a nonword can produce a word superiority effect, and also a study by Forster (1985) found that providing meaning for a nonword can produce a masked form priming effect in the lexical decision task. Both the word superiority effect and the masked form priming effect would appear to tap relatively early lexical processes. Finally, there is evidence from masked semantic priming studies (e.g., Balota, 1983; Carr & Dagenbach, 1990; Dagenbach, Carr, & Wilhelmsen, 1989; Fowler, Wolford, Slade, & Tassinary, 1981; Hirshman & Durante, 1992; Marcel, 1983) suggesting that highly masked primes (that subjects apparently cannot consciously recognize) produce semantic priming effects, that is, facilitate the processing of related targets compared to unrelated targets (see, however, concerns raised by Holender, 1986, and the accompanying commentaries for a full discussion of the degree of conscious processing of the primes in these studies). At the very least, such threshold priming effects suggest that under presentation conditions that minimize conscious processing of the prime, meaning access can still occur.

E. Summary

The possibility that meaning-level representations play a role in isolated word recognition has relatively far reaching implications for current models of word recognition. Most of the available models emphasize the stages that subjects use in accessing the mental lexicon, with virtually no allowance for a top-down meaning-level influence. However, when reminded that the role orthographic patterns play in reading is to convey meaning and not simply to convey lexicality, then one might easily envisage an architecture that incorporates a relatively early influence of meaning. At this level, it should be no surprise that meaning-level representations may contribute to relatively early perceptual analyses and aid in constraining the percept, that is, recognition of the word.

VI. CONTEXT/PRIMING EFFECTS

Heretofore, I have primarily discussed the literature that deals with variables that influence isolated visual word recognition. Of course, readers typically encounter words in the context of other words. We now turn to a summary of the influences of contexts (hereafter referred to as primes) on word recognition processes. In these studies, two strings of letters are typically presented, and the researcher manipulates the relation between the two strings. For example, the strings may be orthographically related (*couch–touch*), phonologically related (*much–touch*), or semantically related (*feel–touch*). By manipulating the types of relationships between the primes and targets, one can obtain evidence regarding the architecture of the word recognition system. For a more detailed discussion of this literature, the reader is referred to Neely (1991) for single-word semantic priming studies and Stanovich and West (1983) for sentence context studies. In addition, the reader is referred to Simpson (this volume) for a review of the role of context in the processing of ambiguous words.

A. Orthographic Priming Effects

An interesting approach to identifying the access code in word recognition is the masked orthographic priming paradigm developed by Evett and Humphreys (1981; also see Humphreys, Besner, & Quinlan, 1988; Humphreys, Evett, Quinlan, & Besner, 1987). In this paradigm, subjects are briefly shown two strings of letters that are both preceded and followed by pattern masks. The two strings vary in terms of orthographic, phonological, or semantic relatedness. Here, we focus on the orthographic priming conditions. There are a number of interesting findings in the original Evett and Humphreys study: First, on most trials, subjects were unable to consciously identify the prime items, and hence any influence of the prime items presumably reflects early access processes. Second, subjects were better at identifying the second letter string when it shared letters with the first letter string even though these shared letters were presented in different case (i.e., there was priming from *lert* to LOST). Third, this effect occurred even when the prime items were nonwords.

There has been some controversy regarding the later finding of nonword orthographic priming. The theoretical controversy regarding the nonword orthographic priming effect concerns *where* in the processing system the orthographic priming effect takes place. If one only finds orthographic priming for words and not for nonwords, then this paradigm is most likely tapping lexical-level processes instead of sublexical orthographic processes. In fact, Forster (1987) and Forster and Davies (1984) failed to find orthographic priming effects from nonwords to target words in a lexical decision task. However, it is also possible that masked nonword primes in the lexical decision task bias the nonword response and hence produce some conflict with the word response on target word trials. In this light, it is noteworthy that Manso de Zuniga, Quinlan, and Humphreys (1987) and Sereno (1991) have reported masked orthographic priming effects in the pronunciation task. Hence, again the appearance of an effect of a variable appears to depend on the task used to measure that variable, and such a task dependency appears to have quite an important theoretical consequence.

It should also be noted here that there is considerable evidence that orthographic codes can be accessed in the parafovea and can be used to facilitate later processing of visually presented words (e.g., Balota, Pollatsek, & Rayner, 1985; Balota & Rayner, 1983; Rayner, McConkie, & Ehrlich, 1978; Rayner, McConkie, & Zola, 1980; see Balota & Rayner, 1991, for a review). The work by Rayner et al. (1980) provides the clearest evidence for orthographic parafoveal priming. In this study, parafoveal letter strings were presented in a pronunciation task. During the subject's saccade to the parafoveal target word, the first parafoveal string (e.g., *chovt*) was replaced by a different string (*chart*). Rayner et al. found that pronunciation latency was (a) facilitated if the preview string and the target string shared the first two or three letters, (b) not dependent on a match in case (upper or lower case) between the prime and target, and (c) uninfluenced by the lexicality of the parafoveal previews. (For further details of the work addressing eye movements, parafoveal processing, and reading, see Rayner and Sereno, this volume.)

In sum, there appears to be converging evidence from both foveal priming studies and parafoveal priming studies that orthographic codes that are not dependent on case-level information do facilitate word recognition performance. Hence, there is relatively strong evidence from this area for an abstract (case-independent) orthographic access code.

B. Phonological Priming Studies

As described above, there has been considerable debate concerning the role of phonological codes in word recognition. The extremes range from "all words must be recognized via a phonological (assembled) code" to the notion "all words are only accessed via an orthographic (addressed) code." Although there is some controversy regarding the role of a phonological code in visual word recognition, there is considerably less debate regarding the importance of phonological codes in reading text, wherein phonological codes produce representations that appear better suited for aspects of comprehension that place considerable demands on the working memory system (e.g., Baddeley, Eldridge, & Lewis, 1981; Besner, 1987; Slowiaczek & Clifton, 1980). The more narrow issue here is whether phonological codes are used in the word recognition process. With this in mind, we now turn to the phonological priming literature.

Meyer, Schvaneveldt, and Ruddy (1974) reported one of the first studies of phonological priming effects. In the critical conditions of this study, subjects were presented with pairs of words that had (a) similar orthographic forms but different phonological forms (e.g., *couch–touch*), (b) similar orthographic and phonological forms (e.g., *bribe–tribe*), or (c) dissimilar orthographic and phonological forms (*chair–tribe*). Meyer et al. found that subjects were faster to make lexical decisions to orthographically and phonologically related pairs, compared to unrelated pairs, and slower to make lexical decisions to orthographically related but phonologically unrelated word pairs, compared to unrelated pairs. This pattern suggests a strong role of phonological information as an access code (also see Hillinger, 1980). However, H. G. Shulman, Hornak, and Sanders (1978) found that Meyer et al.'s results were only observed when the nonwords were pronounceable. When the nonwords were unpronounceable, Shulman et

al. actually found facilitation for the orthographically related but phonologically unrelated pairs (e.g., *couch–touch*). Again, we find that the decision processes associated with the lexical decision task constrain one's interpretation of this finding. Tanenhaus, Flanigan, and Seidenberg (1980) provided evidence for orthographic and phonological activation in a Stroop paradigm wherein subjects were simply asked to name the color of the target word. Presumably because of competition between activated codes and the pronunciation response, subjects were actually slower to name the color of the target word when the primes and targets were orthographically and/or phonologically similar, compared to when they were unrelated.

Evett and Humphreys (1981) used the masked priming paradigm, described above, to investigate the viability of a phonological access code, under conditions wherein conscious processing was limited. They used conditions similar to the Meyer et al. study described above. The results of this study indicated that there was priming for pairs that were orthographically and phonologically related (e.g., *bribe–tribe*) compared to pairs that were orthographically related but phonologically unrelated (*break–freak*). Moreover, the effect occurred across case changes. In addition, in a similar masked priming paradigm, Humphreys, Evett, and Taylor (1982) found that identification accuracy was higher for targets (e.g., *chute*) that followed homophonic primes (e.g., *shoot*) compared to targets that followed graphemically related (e.g., *short*) or unrelated primes (*trail*). However, there was no facilitation from a nonword phonologically related prime (e.g., *smorl–SMALL*). Thus, the results from Humphreys et al. suggest that the phonological priming effect, as opposed to the orthographic priming effect discussed above, may be lexically mediated.

In an interesting variation of the masked phonological priming paradigm, Perfetti, Bell, and Delaney (1988) briefly presented target words (e.g., *made*) that were followed by nonword masks that were either phonologically related (e.g., *mayd*) or unrelated (*mard*) to the target word. The results indicated that under these conditions, subjects were more accurate at identifying the target word when it was masked by a phonologically related nonword compared to a phonologically unrelated nonword. Because these phonological effects occur quite strongly for nonwords, it would appear that, in contrast to the results of the Humphreys et al. study discussed above, these effects appear to be at a prelexical level.

It is also worth noting two studies by Van Orden (1987; Van Orden, Johnston, & Hale, 1988). These studies involved a semantic categorization task, in which subjects had to decide whether a given word was a member of a semantic category. The intriguing finding here is that subjects produced a considerably higher error rate for words that were homophones of an exemplar (e.g., *meet* for the category *food*), compared to an orthographically related control (e.g., *melt*). This finding would suggest a clear role of phonological information in accessing the semantics necessary for category verifications. Recently, Jared and Seidenberg (1991) have replicated and further specified the influence of phonology in this paradigm. Although there are a number of intriguing findings in the Jared and Seidenberg study, one of the major outcomes of this study is that the influence of phonology in the Van Orden paradigm primarily occurs for low-frequency words. This pattern appears to be consistent with the earlier

observation of an interaction between frequency and spelling-to-sound regularity that was observed in word pronunciation performance.

Finally, it is also worth noting that just as in the case of orthographic priming, there is also recent evidence of phonological priming in the parafoveal priming paradigm described above. Specifically, Pollatsek, Lesch, Morris, and Rayner (1992) found that previews that were homophonic with targets (e.g., *site–cite*) facilitated performance (both in pronunciation latencies and fixation durations) compared to nonhomophonic previews that were controlled for orthographic similarity (e.g., *cake–sake*). Again, this pattern would appear to support a role for phonology as an access code.

C. Semantic Priming Effects

The SEMANTIC (associative) PRIMING paradigm is by far the most well researched area of priming. This enterprise began with a seminal study by Meyer and Schvaneveldt (1971). They found that subjects were faster to make lexical decisions to each word in a pair of words when the words were related (e.g., *cat–dog*) compared to when the words were unrelated (e.g., *cat–pen*). The prevailing Zeitgeist was ready to welcome such a finding for a number of reasons: First, the dependent measure was response latency, and response latency measures were becoming the mainstay of cognitive experiments. Second, the study nicely demonstrated top-down contextual influences (e.g., semantic relations) on what would appear to be bottom-up word recognition processes. This was a major emphasis in Neisser's (1967) *Cognitive Psychology* that had been published a few years earlier. Third, the effect was quite robust and easily replicated. Fourth, the semantic priming task appeared to be ideally suited to map out the architecture of meaning-level representations and the retrieval operations that act on such representations; both of these issues would at least appear to be critical to higher level linguistic performance.

There is little controversy that across the three major tasks used to build word recognition models (threshold identification, lexical decision, and pronunciation), words are better recognized when embedded in semantically related contexts compared to unrelated contexts. However, there are many questions that one might ask about this effect. For example, one might ask whether this benefit is truly semantic or simply involves associative co-occurrence across words within a language. Here, by a semantic relationship I am referring to words that share semantic features (E. E. Smith, Shoben, & Rips, 1974) and/or entail subordinate or superordinate semantic category relations (e.g., Collins & Quillian, 1969). For example, *mouse* and *rat* are semantically related because both are members of the rodent category. However, *mouse* and *cheese* are primarily associatively related because these words do not share a simple semantic category relationship or involve much overlap in semantic features. Of course, teasing apart semantic influences from associative influences has been rather difficult because these relationships typically co-occur. In an attempt to address this issue, researchers have identified items that are of the same category (e.g., *glove–hat*) but do not entail a strong associative relation, e.g., are not produced in associative production norm studies in which subjects are asked to generate associates to a given word (e.g., Palermo & Jenkins, 1964). The

results from three such studies (e.g., Lupker, 1984; Schreuder, Flores d'Arcais, & Glazenborg, 1984; Seidenberg, Waters, Sanders, & Langer, 1984) indicate that there is still some priming with such stimuli in both lexical decision and in pronunciation, although the pure semantic effects are somewhat smaller in pronunciation.

However, one must be somewhat cautious in accepting the conclusion that there are pure nonassociative semantic priming effects. This caution is warranted for the following reasons: First, and foremost, it is unclear whether the relatively small pure semantic priming effects might be due to some lingering associative level relationship for words that researchers have argued only have a semantic relationship (e.g., *glove–hat* are probably more likely to co-occur compared to the pair *glove–pen*). Second, as noted below, there is evidence that priming can occur across mediated pairs within the memory network. Thus, it is at least possible that some of the priming from *glove* to *hat* is due to *glove* priming *clothes* and *clothes* priming *hat*. Third, when one considers low category dominance pairs, words that are categorically related but may have little associative relationship, one finds that there is relatively little priming in pronunciation performance (Keefe & Neely, 1990; Lorch, Balota, & Stamm, 1986), however, in lexical decision performance, there appears to be equivalent priming for high and low category dominance pairs (e.g., Lorch et al., 1986; Neely, Keefe, & Ross, 1989). The difference between pronunciation and lexical decision performance is particularly noteworthy here. A number of researchers have suggested that at least part of the priming effect observed in the lexical decision task may be due to a type of postlexical checking processes. Subjects can use the relatedness between the prime and target to bias their "word" response because nonwords by definition are never semantically related to the primes. In fact, Neely et al. (1989) have found that the priming effect for low-dominance exemplars (words that are acceptable but are produced relatively infrequently in category-exemplar production norms, e.g., *bird–goose*) in the lexical decision task depends on the ratio of nonwords to words. Neely et al. argue that the nonword/word ratio should modulate the utility of the checking process in the lexical decision task. Also, the fact that the pronunciation task yields little priming for low dominance category exemplars may reflect a decreased reliance on such a checking process in the pronunciation task. Hence, these data also question the evidence for a pure semantic priming effect in access processes.

1. Mediated Priming Effects

At an intuitive level, the finding that subjects are better at recognizing words that are embedded in related contexts compared to unrelated contexts is no great surprise. (Of course, it is clearly not so intuitive what mechanisms are responsible for such effects.) However, the priming literature has also provided some very counterintuitive findings. Consider the two words *lion* and *stripes*. These two words do not have any obvious direct relation but do have an indirect relation through the word *tiger*. Such items have been referred to as mediated pairs, and the research addressing mediated priming effects has provided some interesting results. First, in a lexical decision task in which subjects only respond to the target string, there is little evidence for mediated priming (cf. Balota & Lorch, 1986; DeGroot, 1983; Den Heyer, Sullivan, & McPherson, 1987). However, if one changes the lexical decision task so that

subjects either (a) make lexical decisions about the prime and target (McNamara & Altarriba, 1988) or (b) only make a response to word targets and not to nonword targets (Den Heyer et al., 1987), mediated priming does occur in the lexical decision task. Moreover, when one now turns to the pronunciation task, one does find mediated priming effects (Balota & Lorch, 1986). Researchers have again argued that checking processes tied to the lexical decision task can strongly control when mediated priming effects will be found in this task (e.g., Balota & Lorch, 1986; McNamara & Altarriba, 1988; Neely, 1991). The notion is that checking for a relationship between the prime and target will not yield a successful outcome for mediated prime–target pairs, because such pairs do not share any obvious relationship. Thus, a negative outcome from the checking process may override the mediated influence from the prime to the target.

2. Threshold Priming Effects

A second counterintuitive finding in this literature deals with the threshold semantic priming effects, noted earlier. In these experiments, researchers first determine each subject's threshold wherein he or she can no longer discriminate between the presence or absence of a stimulus. These thresholds are then used in a later semantic priming task, in which the prime is presented at this threshold and the target is presented in a lexical decision task. The intriguing finding here is that there is still evidence for semantic priming effects, under conditions in which subjects apparently can no longer make presence/absence decisions about the prime item (Balota, 1983; Carr & Dagenbach, 1990; Dagenbach et al., 1989; Fowler et al., 1981; Marcel 1983; Marcel & Patterson, 1978). There have also been similar findings reported in the pronunciation task (Carr, McCauley, Sperber, & Parmelee, 1982; Hines, Czerwinski, Sawyer, & Dwyer, 1986). Although, as noted, there is some concern regarding whether subjects are truly at an objective presence/absence threshold (see Cheesman & Merikle, 1984; Holender, 1986; Merikle, 1982), it is clear that primes presented under very degraded conditions still produce semantic priming effects. As in the mediated priming studies, these studies appear to indicate that conscious access to a prime–target relationship does not appear to be a necessary condition for semantic priming effects.

3. Backward Priming Effects

The third area that is somewhat counterintuitive in this area is backward priming. There are two types of backward priming effects. First, there is evidence (Balota, Boland, & Shields, 1989; Kiger & Glass, 1983) that indicates one can still find semantic priming (*dog–cat* versus *pen–cat*) even when the prime (*dog* or *pen*) is presented temporally after the target (*cat*). These results suggest that early on in target processing, subsequent related prime information/activation can actually catch up to influence response latencies to the target. Such an effect would appear to most naturally fall from a cascadic framework in which partial activation is released from representations before such representations have reached threshold (see earlier discussion of the McClelland & Rumelhart, 1981, model).

A second type of backward priming effect is backward semantic priming. In backward semantic priming, prime–target pairs are presented that entail directional relations, e.g., *bell* is related to *boy* in the *bell–boy* direction, but

not in the *boy–bell* direction. Koriat (1981) and Seidenberg, Waters, Sanders, and Langer (1984) have reported evidence of backward priming in the lexical decision task. However, when one turns to the pronunciation task, there is relatively little evidence of backward priming (Seidenberg, Waters, Sanders, & Langer, 1984), except under short stimulus onset asynchronies (SOAs, the temporal interval between the onset of the prime and the onset of the target) and auditorily presented primes (see Peterson & Simpson, 1989). The prevailing account for the difference between the backward priming effects in pronunciation and lexical decision is that the priming effects in the pronunciation task are more directional from the prime to the target, whereas in the lexical decision task subjects may check for a relationship between the target and the prime (see Neely, 1991). Thus, if the subject checks in a backward fashion from the target (*bell*) to the prime (*boy*), a relationship will be found to bias the word response in the lexical decision task.

D. Syntactic Priming

If associative/semantic context does indeed influence lexical processing, then it is quite possible that syntactically appropriate versus inappropriate contexts might also influence lexical processing. In fact, effects of syntactic context on word recognition might be quite informative. At one level, one might argue that associative pathways between syntactically appropriate words might be represented within the lexicon, simply due to associative co-occurrence of such pairs (cf. Ratcliff & McKoon, 1988). Likewise, one might argue that syntactic tags within lexical representations might produce priming to consistent syntactic representations. On the other hand, one might argue that syntactic representations are only engaged after word recognition, and hence one might not expect syntactic priming effects in word recognition tasks.

One of the first syntactic priming studies was reported by Goodman, McClelland, and Gibbs (1981). Goodman et al. found that subjects were faster to make lexical decisions to targets (e.g., *oven*) that followed syntactically appropriate primes (e.g., *my*) compared to syntactically inappropriate primes (e.g., *he*). Seidenberg, Waters, Sanders, and Langer (1984) replicated this pattern in a lexical decision task but only obtained marginal effects in the pronunciation task. As in the priming studies mentioned above, Seidenberg et al. argued that the syntactic priming effect in the lexical decision task was probably due to some postlexical processing of the relation between the prime and target. However, it would appear that the Seidenberg et al. arguments are not totally correct, because West and Stanovich (1986) obtained relatively large syntactic priming effects in both the pronunciation task and the lexical decision task.

More recently, Sereno (1991) noted that the past syntactic priming effects may have been due to attentional expectancies that subjects may have built up because of relatively long prime–target stimulus onset asynchronies used in these studies (see discussion below regarding long SOAs and attentional expectancies). If this were the case, then the previously observed syntactic priming effects may not have been due to influences at the lexical level. In order to test this possibility, Sereno relied on the three-word masking paradigm developed by Forster and Davies (1984). On each trial, subjects were first presented an unrelated forward masking word for 500 ms and then were briefly presented

the prime word for 60 ms, which was followed by the target word presented until a response was made. Because the prime is both forward- and backward-masked, conscious processing, and hence backward checking, is presumably limited. Sereno used both a lexical decision task and a pronunciation task to insure that any observed priming effects could not be simply due to residual conscious processing of the prime and its influence on the backward checking process in the lexical decision task. The results indicated that there were no syntactic priming effects in the pronunciation task but there were reliable syntactic priming effects in the lexical decision task. Based on this pattern, Sereno argued that the observed syntactic priming effects in the lexical decision task are most likely due to post-lexical processes that are tied to that task.

Although the Sereno results are quite compelling, it is also worth noting a study by Samar and Berent (1986). In this study, Samar and Berent recorded evoked responses in a syntactic priming lexical decision task. The interesting finding in this study is that there was a reliable evoked response component peaking 140 ms after target presentation that discriminated between conditions in which words were presented in syntactically appropriate contexts (e.g., *the-job, we-bring*) compared to syntactically inappropriate contexts (e.g., *the-bring, we-job*). Although there are some aspects of this study that may diminish the strength of the arguments (e.g., relatively long 500 ms prime–target SOA, and the use of the lexical decision task), the relatively early peak in the evoked response would appear to support an early role for syntactic analyses (see Bentin, 1989, and Kutas and Van Petten, this volume, for further reviews of the evoked response literature.)

E. Prime Type by Factor Interactions

Of course, the importance of the semantic priming literature is not simply the demonstration that certain factors produce facilitation in the lexical decision task and pronunciation tasks; its importance is also due to the intriguing interactions that have been uncovered. Again, because of space limitations, we can only list some of the more important interactions: (a) semantic priming effects are larger for low-frequency words than for high-frequency words (Becker, 1979); (b) semantic priming effects are larger for degraded words compared to nondegraded words (Becker & Killion, 1977; Borowsky & Besner, 1991) (the interactive effects of semantic priming and degradation become even more intriguing when one considers that word frequency effects are typically additive with stimulus degradation; Becker & Killion, 1977; Borowsky & Besner, 1991; see Besner & Smith, 1992, for a recent attempt to accommodate the combined influences of frequency, degradation, and semantic context); (c) semantic priming effects are larger for poor readers than good readers (Stanovich, 1980); (d) semantic priming effects are larger for the lexical decision task than for the pronunciation task (e.g., Balota & Lorch, 1986); (e) semantic priming effects in pronunciation increase across prime–target SOAs equally for high-strength and low-strength prime–target pairs (Balota & Duchek, 1988; Lorch, 1982); (f) semantic priming effects are larger for lists that contain a high proportion of related to unrelated prime–target trials compared to lists that contain a relatively low proportion of related to unrelated prime–target trials, hereafter referred to as the PROPORTIONAL RELATEDNESS EFFECT (e.g., den Heyer,

Briand, & Dannenbring, 1983; Keefe & Neely, 1990; Neely et al., 1989); (g) the proportional relatedness effect is larger at long prime–target SOAs than at short prime–target SOAs (e.g., den Heyer et al., 1983); (h) the proportional relatedness effect does not occur for low-dominance prime–target exemplars in the pronunciation task but does occur for such items in the lexical decision task (Neely et al., 1989); (i) prime-induced expectancy effects are larger when the prime–target SOA is relatively long compared to when the prime–target SOA is short (Balota, Black, & Cheney, 1992; Burke, White, & Diaz, 1987; Favreau & Segalowitz, 1983; Neely, 1977); (j) facilitation of response latencies to targets in the lexical decision task following related primes compared to neutral primes (e.g., *xxxxx*) decreases across SOAs, whereas inhibition of unrelated contexts compared to neutral primes increases across SOAs (e.g., Favreau & Segalowitz, 1983; Neely, 1977).

F. Theoretical Accounts of Semantic Priming Effects

The importance of the semantic priming paradigm has not simply been restricted to models of word recognition, but has also extended to more general issues concerning representation and retrieval processes. I shall now briefly discuss some of the theoretical issues that have been nourished by this literature. I refer the reader to Neely (1991) for further discussion of these theoretical mechanisms.

1. Automatic Spreading Activation

The notion that semantic/lexical memory may be represented by nodes that reflect concepts and that such conceptual nodes are interconnected via associative/semantic pathways has been central to a number of developments in cognitive psychology (e.g., Anderson, 1976, 1983; Collins & Loftus, 1975; Posner & Snyder, 1975). As Anderson (1983) points out, the spreading activation metaphor has probably been most strongly supported by the semantic priming paradigm. According to the spreading activation framework, when a node in memory becomes activated via stimulus presentation or via internal direction of attention, activation spreads from that node along associative pathways to nearby nodes. Thus, the reason that subjects are faster to recognize *dog* when it follows *cat*, compared to when it follows *pen*, is that the underlying representations for these two words are connected via an associative/semantic pathway, and when *cat* is presented activation spreads from its underlying node to the node underlying *dog*. Thus, the representation for *dog* needs less stimulus information to surpass threshold.

Although there is a limited-capacity version of spreading activation theory (e.g., Anderson & Bower, 1973), by far most of the work in the priming literature has addressed the AUTOMATIC nature of the spreading activation mechanism. In one of the clearest expositions of this mechanism, Posner and Synder (1975) argued that the automatic spreading activation mechanism (a) is fast-acting, (b) is independent of subjects' conscious control, and (c) primarily produces facilitation for related targets and little inhibition for unrelated targets, compared to an appropriate neutral baseline condition (see Neely, 1977). Because of current controversies regarding the adequacy of a given neutral prime condition (see, e.g., Balota & Duchek, 1989; DeGroot, Thomassen, & Hudson, 1982;

Jonides & Mack, 1984; Neely, 1991), I focus primarily on Posner and Synder's first two characteristics.

There are a number of important semantic priming results that would appear to support Posner and Synder's automatic spreading activation mechanism. First, with respect to the notion that priming effects are independent of consciously controlled processing, the finding that one still obtains priming effects under conditions in which the primes are briefly presented and highly masked (e.g., Balota, 1983; Fowler et al., 1981; Marcel, 1983) would appear to provide strong support for this assumption. In addition, the finding that there are mediated priming effects at relatively short prime–target SOAs (e.g., from *lion* to *stripes*), when it is unlikely that subjects could generate an attentional expectancy for the mediated target, also supports the notion of an automatic spread of activation within a memory network. Finally, the findings that prime-expectancy instructions (Neely, 1977) and relatedness proportion manipulations have relatively little impact at short SOAs (den Heyer et al., 1983) support the notion that the automatic spreading activation mechanism is relatively fast acting (i.e., occurs at short SOAs) and is independent of subjects' conscious expectations.

Although there appears to be strong support for something akin to an automatic spreading activation mechanism, there are some caveats that need to be noted. One issue that has been relatively controversial is whether priming effects occur across unrelated words (e.g., facilitation from *lion* to *tiger* in *lion–chalk–tiger* compared to *frog–chalk–tiger*). It is unclear why an unrelated word would disrupt an automatic spreading activation mechanism. However, Gough, Alford, and Holley-Wilcox (1981), Masson (1991), and Ratcliff and McKoon (1988) have all reported failures to obtain priming across unrelated words. Interestingly, there are two recent studies by Joordens and Besner (1992) and McNamara (1992) that have obtained such priming effects. Of course, one might expect such priming effects to be relatively small because the unrelated word may have the effect of shifting attention away from the related prime, and this shift may override any pure spreading activation effect. In fact, Joordens and Besner note that there have been small but nonsignificant effects in the predicted direction in the earlier studies. A second potential problem with the automatic nature of spreading activaton is that semantic priming effects can be eliminated when subjects process the primes in a very shallow fashion, as when responding to whether a given letter is in the prime or an asterisk is beside the prime (e.g., Henik, Friedrich, & Kellog, 1983; M. C. Smith, 1979; M. C. Smith, Theodor, & Franklin, 1983). Unless the shallow processing task eliminates processing of the prime at the lexical level, one should expect automatic spreading activation and semantic priming effects under shallow processing conditions (see, Besner, Smith, & MacLeod, 1990, for further discussion of this issue). Finally, Balota et al. (1992) have recently provided evidence that prime-expectancy instructions (e.g., when subjects are instructed to expect exemplars from the *tree* category when presented the prime *metals*) can influence pronunciation performance even at very short prime–target SOAs. Thus, although there is some evidence in support of an automatic spreading activation mechanism involved in semantic priming tasks, it appears that we still do not fully understand the constraints under which this mechanism operates.

2. Attentional/Expectancy Effects

A second mechanism that presumably underlies semantic priming effects is a more attention-based expectancy factor (Balota, 1983; Becker, 1980; Favreau & Segalowitz, 1983; Neely, 1976, 1977). Here, when the prime is presented, subjects generate expectancies about potential candidate targets. When the expectancy is correct, facilitation occurs; however, when the expectancy is incorrect, inhibition occurs. This expectancy-based model of priming falls naturally from the work of Posner and Snyder (1975) and Neely (1977), wherein instructional manipulations and list context effects have larger influences at long SOAs (when expectancies have had time to be generated) than at short SOAs. Of course, at one level the impact of an attentional-based expectancy mechanism should not be surprising because it simply reflects the probability of correctly predicting the target word. The more intriguing work here is the specification of the parameters that modulate the expectancy effects, that is, the rate at which expectancies are generated across time, the duration at which the expectancy is maintained, and the characteristics of such an expectancy set (see Becker, 1980, 1985, for a detailed discussion of a semantic expectancy model).

3. Backward-Checking Accounts

As noted above, a number of researchers have argued that priming effects in the lexical decision task may reflect influences at a postlexical decision level (e.g., Balota & Lorch, 1986; DeGroot, 1984; Forster, 1979, 1981; Neely, 1976, 1977; Neely & Keefe, 1989; Seidenberg, Waters, Sanders, & Langer, 1984; Stanovich & West, 1983). Subjects can rely on finding a relationship between the prime and target to bias the "word" response in the lexical decision task, because nonwords are never related to the primes. This would have the effect of facilitating "word" decisions to related prime–target trials and possibly inhibiting "word" decisions to unrelated prime–target trials. As described above, there is considerable support for such a mechanism in the lexical decision task. For example, the finding that there is backward priming in the lexical decision task (e.g., priming from *boy* to *bell*) suggests that subjects can use the target to check in a backward direction (*bell* to *boy*) about any potential relationship to the prime item. Although the backward-checking mechanism would appear to be primarily a nuisance variable tied to the lexical decision task, one might argue that this checking process may reflect a tendency in natural language processing to integrate meanings across words (see Neely & Keefe, 1989, for a full discussion of the backward checking mechanism).

4. Ratcliff and McKoon's (1988) Compound Cue Model

Ratcliff and McKoon have developed a model that takes a quite different approach to priming effects in the lexical decision task. The model is based on a formal model of episodic recognition memory developed by Gillund and Shiffrin (1984). In Ratcliff and McKoon's model, items in short-term memory serve as a compound cue, with the more recently presented items having a larger influence on the output of the retrieval process. If the prime and target are associated, then this will provide a higher familiarity value than if the prime and target are not associated. Familiarity is then used to predict response latency via a random-walk decision process (Ratcliff, 1978) wherein

high-familiar compound cues produce relatively fast "word" decisions and low-familiar compound cues (e.g., nonwords) produce relatively fast "nonword" decisions. Intermediate values of familiarity produce relatively slower and less accurate decisions. Hence, if familiarity is modulated by the degree to which primes and targets are either directly associated or share associates in memory, then one should find that related prime–target pairs will produce higher familiarity values and faster response latencies in the lexical decision task than unrelated prime–target pairs.

The compound cue model has a number of positive characteristics. First, the model is based on formal memory models. The quantitative aspect of this model is a clear strength over other theories of semantic priming. Second, as Ratcliff and McKoon point out, their model handles a number of important findings in the priming lexical decision literature that spreading activation or attentional expectancy mechanisms do not appear to handle. For example, the model nicely accounts for backward priming effects in the lexical decision task because the target *boy* and prime *bell* can serve as a compound cue that is directly related in memory and hence produces a high familiarity value. The model can also handle single-step mediated priming effects (from *lion* to *stripes* via *tiger*) but apparently cannot handle two-step mediated priming effects (e.g., *mane* to *stripes* via both *lion* and *tiger*, see McKoon & Ratcliff, 1992, and McNamara, 1992, for further discussion of mediated priming effects).

Although the compound cue model does provide an interesting alternative to prime-induced mechanisms, there are some limitations to this approach. For example, the model is primarily a model of the lexical decision task and hence does not account for the wealth of interesting priming data from the pronunciation task. The tripartite (spreading activation, attentional expectancies, and backward checking) framework accounts for both lexical decision and pronunciation results by assuming logogen-type word recognition devices that are also connected to a phonological output system used for pronunciation. Second, and more important, the distinction between the compound cue model and the spreading activation framework may be more apparent than real. In both frameworks, it is necessary to map the influence of relationships between words onto priming effects. Within the spreading activation framework, this mapping involves the preactivation of related concepts in memory, whereas within the compound cue model, this mapping is based on a rule that computes familiarity based on associations within long-term memory. At this level, the major distinction between the spreading activation framework and the compound cue model involves the mapping process.

G. Summary of Context/Priming Effects

The semantic priming literature has provided an extremely rich database for developing models of context effects, memory retrieval, and word recognition. Because of space limitations, I was unable to provide a review of other important models of semantic priming effects such as Becker's (1980) verification model, Norris' (1986) plausibility-checking model, Forster's (1981) bin model, and Masson's (1991) recent connectionist model. Each of these models provide intriguing alternative perspectives on semantic priming effects. However, at this point in theory development, I would agree with Neely (1991) that no single model of priming is available that readily accounts for the richness and diversity

of this literature, and it would appear that multiple mechanisms will need to be postulated to account for the breadth of priming effects. At least at the present stage of development, it would appear that some variants of spreading activation, attentional prediction, and backward checking will need to be incorporated into a model to account for most of the observed semantic priming effects.

VII. CONCLUDING REMARKS

In the present chapter I have attempted to provide the reader with an overview of the major issues addressed in the word recognition literature. To conclude, I would like to address some of the major themes that have spanned a number of the sections.

First, in each of the sections, there has been evidence initially supporting a rather straightforward theoretical analysis, and then there have been reports by trouble-makers that constrain the strength of the theoretical inferences available from a given task. For example, even in the word superiority paradigm, there have been arguments that partial information from the target letter could, in conjunction with the word envelope, allow subjects to use a sophisticated guessing strategy to bias the correct choice (e.g., Krueger & Shapiro, 1979; Massaro, 1979). If this is the case, then the word superiority effect may not reflect top-down impacts in perception, but rather biases that occur at postperceptual levels, based on partial information. Similar concerns were raised about the threshold identification, lexical decision, and pronunciation tasks. Of course, task analyses can be frustrating for theoreticians. However, before inferences can be made regarding the underlying locus or loci of a given variable, one needs to be especially careful in developing tasks that faithfully reflect such processes. Clearly, the adequacy of any theory rests on the adequacy of the tasks used to build that theory.

A second consistent theme that has surfaced in this review is whether separable sublexical analyses are performed en route to word recognition, or the apparent influences of sublexical analyses are in large part merely a consequence of the activation and inhibition patterns across many lexical representations. Although some effects appear to be modeled quite well by interactive activation and parallel distributed processing systems, there have also been results that appear inconsistent with such systems. There are at least two likely outcomes to this area of work: First, more of the apparent sublexical effects may fall from these models when networks that are closer to the size of an adult's vocabulary are implemented (see Seidenberg & McClelland, 1990). Second, it may be necessary to implement sublexical processing modules within such connectionist models to incorporate these types of analyses. Ultimately, however, the potential for this level of explanatory power makes the connectionist modeling in word recognition very appealing.

A third theme in the present review is actually a type of statistical interaction that has been repeatedly observed. The vast majority of interactions in this literature are of the nature that Factor A has more of an effect at the level of Factor B that produces the slowest or least accurate performance. Consider, for example, word frequency. I have reviewed evidence indicating that compared to

high-frequency words, low-frequency words produce larger effects of bigram frequency, phonological regularity, word-body strength, concreteness, semantic priming, task (lexical decision task vs. category verification vs. pronunciation), repetition priming, and neighborhood size, among others. There are at least two noteworthy aspects of these interactions. First, one may wish to argue that because of the development of automaticity, high-frequency words are recognized via routes that effectively bypass many sublexical stages of analyses. Hence, if one is interested in identifying many of the intriguing sublexical aspects of word recognition, one should primarily investigate the processing of low-frequency words. Alternatively, as Loftus (1978) has noted, on a simply statistical level, this particular type of interaction is one of the most difficult to interpret. In fact, it is possible that if one considered percentage of overall response latency change as a function of the levels for Factor A and B, many of these interactions would disappear. In this light, it may be the interactions that still remain after considering percentage change as the dependent measure that may be most informative regarding underlying mechanisms.

Finally, I should note that there are many issues that have not been reviewed simply because of space limitations. For example, I have not discussed word recognition performance in beginning readers and some of the lexical-level breakdowns that appear to occur in developmental dyslexia (see Seymour, 1986, for a review). Also, I decided to attempt to provide an overview of the well-established empirical findings at many different levels of the processing stream, at some cost to treatments of the diverse and rather elegant theoretical accounts of such findings. The emphasis on the empirical literature was in part due to the fact that alternative theoretical accounts at least in some cases appear to be equally plausible. As Anderson (1978) has pointed out, this can often be a striking limitation in cognitive theory development. We are still in search of the critical experiment in many cases to discriminate models. It is quite possible that we will be forced to await physiological evidence to help discriminate between some of these models. Although clearly still relatively early in its development, cognitive neuroscience has considerable potential to contribute to our understanding of the word recognition system.

In light of this chapter, I am hopeful that the reader agrees that at some level the word is to cognitive psychologists and linguists as the cell is to biologists. Both entail many substructures and interact with many higher level systems. The present overview of the word recognition literature may seem rather imposing, and sometimes it would appear that little progress is being made. However, I clearly do not feel that this is the case and believe that considerable progress has been made. At this level, the seductive simplicity of understanding lexical-level analyses surely is more apparent than real. As is often the case in a discipline, the more we know about a system, the more we realize what we need to know to understand that system.

ACKNOWLEDGMENTS

This work was, in part, supported by NIA grant 54297A. Thanks are extended to Sheila Black, Lisa Connor, Janet Duchek, Lisa Foley, Richard Ferraro, and especially Morton Gernsbacher for helpful comments on earlier drafts of this manuscript.

REFERENCES

Abrams, R. A., & Balota, D. A. (1991). Mental chronometry: Beyond reaction time. *Psychological Science, 2,* 153–157.

Adams, M. (1981). What good is orthographic redundancy? In H. Singer & O. J. L. Tzeng (Eds.), *Perception of print* (pp. 197–221). Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review, 85,* 249–277.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22,* 261–295.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory.* Washington, DC: Hemisphere Press.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition, 10,* 565–575.

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

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

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

Ashby, F. G. (1982). Deriving exact predictions from the cascade model. *Psychological Review, 89,* 599–607.

Baddeley, A. D., Eldridge, M., & Lewis, V. J. (1981). The role of subvocalization in reading. *Quarterly Journal of Experimental Psychology, 33A,* 439–454.

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

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

Balota, D. A., Black, S. R., & Cheney, M. (1992). Automatic and attentional priming in young and older adults: Reevaluation of the two-process model. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 485–502.

Balota, D. A., Boland, J. E., & Shields, L. (1989). Priming in pronunciation: Beyond pattern recognition and onset latency. *Journal of Memory and Language, 28,* 14–36.

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

Balota, D. A., & 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.

Balota, D. A., & Chumbley, J. I. (1990). Where are the effects of frequency in visual word recognition tasks? Right where we said they were! Comment on Monsell, Doyle, and Haggard (1989). *Journal of Experimental Psychology: General, 119,* 231–237.

Balota, D. A., & Duchek, J. M. (1988). Age-related differences in lexical access, spreading activation, and simple pronunciation. *Psychology and Aging, 3,* 84–93.

Balota, D. A., & Duchek, J. M. (1989). Spreading activation in episodic memory: Further evidence for age independence. *Quarterly Journal of Experimental Psychology, 41A,* 84–93.

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

Balota, D. A. and Ferraro, F. R. (1993). A dissociation of frequency and regularity effects in pronunciation performance across young adults, older adults, and individuals with senile dementia of the Alzheimer type. *Journal of Memory and Language, 32,* 573–592.

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

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

Balota, D. A., & Rayner, K. (1983). Parafoveal visual information and semantic contextual constraints. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 726–738.

Balota, D. A., & Rayner, K. (1991). Word recognition processes in foveal and parafoveal vision: The range of influence of lexical variables, In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading* (pp. 198–232). Hillsdale, NJ: Erlbaum.

Baron, J., & Thurston, I. (1973). An analysis of the word-superiority effect. *Cognitive Psychology, 4,* 207–228.

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

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

Becker, C. A. (1985). What do we really know about semantic context effects during reading? In D. Besner, T. G. Waller, & E. M. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5) (pp. 125–166). Toronto: Academic Press.

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

Bentin, S. (1989). Electrophysiological studies of visual word perception, lexical organization, and semantic processing: A tutorial review. *Language and Speech, 32,* 205–220.

Besner, D. (1987). Phonology, lexical access in reading, and articulatory supression: A critical review. *Quarterly Journal of Experimental Psychology, 39A,* 467–478.

Besner, D. (1990). Does the reading system need a lexcon? In D. A. Balota, G. B. Flores d'Arcaise, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 73–99). Hillsdale, NJ: Erlbaum.

Besner, D., Davelaar, E., Alcott, D., & Parry, P. (1984). Basic processes in reading: Evidence from the FDM and the FBI. In L. Henderson (Ed.), *Orthographies and reading.* Hillsdale, NJ: Erlbaum.

Besner, D., & McCann, R. S. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 201–219). Hillsdale, NJ: Erlbaum.

Besner, D., & Smith, M. C. (1992). Models of visual word recognition: When obscuring the stimulus yields a clearer view. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 468–482.

Besner, D., Smith, M. C., & MacLeod, C. M. (1990). Visual word recognition: A dissociation of lexical and semantic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 862–869.

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.

Bjork, E. L., & Estes, W. K. (1973). Letter identification in relation to linguistic context and masking conditions. *Memory & Cognition, 1,* 217–223.

Bleasdale, F. A. (1987). Concreteness-dependent associative priming: Separate lexical organization for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13,* 582–594.

Boles, D. B. (1983). Dissociated imageability, concreteness, and familiarity in lateralized word recognition. *Memory & Cognition, 11,* 511–519.

Boles, D. B. (1989). Word attributes and lateralization revisited: Implications for dual coding and discrete versus continuous processing. *Memory & Cognition, 17,* 106–114.

Borowsky, R., & Besner, D. (1991). Visual word recognition across orthographies: On the interaction between context and degradation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 272–276.

Bradley, D. (1980). Lexical representations of derivational relations. In M. Aronoff & M. Kean (Eds.), *Juncture* (pp. 37–55). Cambridge, MA: MIT Press.

Broadbent, D. E. (1967). Word-frequency effects and response bias. *Psychological Review, 74,* 1–15.

Broadbent, D. E., & Gregory, M. (1968). Visual perception of words differing in letter digram frequency. *Journal of Verbal Learning and Verbal Behavior, 7,* 569–571.

Burke, D. M., White, H., & Diaz, D. L. (1987). Semantic priming in young and older adults: Evidence for age constancy in automatic and attentional processes. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 79–88.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28,* 297–332.

Carr, T. H., & Dagenbach, D. C. (1990). Semantic priming and repetition priming from masked words: Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16,* 341–350.

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 and Performance, 4,* 674–690.

Carr, T. H., McCauley, C., Sperber, R. D., & Parmelee, C. M. (1982). Words, pictures, and priming: On semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 757–777.

Catlin, J. (1969). On the word frequency effect. *Psychological Review, 76,* 504–506.

Catlin, J. (1973). In defense of sophisticated guessing theory. *Psychological Review, 80,* 412–416.

Cattel, J. M. (1885). The intertia of the eye and brain. *Brain, 8,* 295–312.

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

Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Perception & Psychophysics, 36,* 387–395.

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

Clark, H. H. (1973). The language-as-fixed fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335–359.

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

Collins, A., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 240–247.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151–216). London & New York: Academic Press.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed processing approaches. *Psychological Review, 100,* 589–608.

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

Connine, C., Mullennix, J., Shernoff, E., & Yelens, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 1084–1096.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

Dagenbach, D., Carr, T. H., & Wilhelmsen, A. (1989). Task-induced strategies and near-threshold priming: Conscious influences on unconscious perception. *Journal of Memory and Language, 28,* 412–443.

Day, J. (1977). Right-hemisphere language processing in normal right-handers. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 518–528.

de Groot, A. M. B. (1983). The range of automatic spreading activation in word priming. *Journal of Verbal Learning and Verbal Behavior, 22,* 417–436.

de Groot, A. M. B. (1984). Primed lexical decision: Combined effects of the proportion of related prime-target pairs and the stimulus onset asynchrony of prime and target. *Quarterly Journal of Experimental Psychology, 36A,* 253–280.

de Groot, A. M. B. (1989). Representational aspects of word imageability and word frequency as assessed through word associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 824–845.

de Groot, A. M. B., Thomassen, A., & Hudson, P. (1982). Associative facilitation of word recognition as measured from a neutral prime. *Memory & Cognition, 10,* 358–370.

den Heyer, K., Briand, K., & Dannenbring, G. (1983). Strategic factors in a lexical decision task: Evidence for automatic and attention-drive processes. *Memory & Cognition, 11,* 374–381.

den Heyer, K., Sullivan, A., & McPherson, C. (1987). *Mediated priming in a single-response lexical decision task.* Unpublished manuscript.

Erdmann, B., & Dodge, R. (1898). *Psychologische Untersuchungen über das Lesen*. Halle, Germany: M. Niemery.

Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology, 33A*, 325–350.

Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in first- and second-language reading of fluent bilinguals. *Memory & Cognition, 11*, 565–574.

Feldman, L. B., & Andjelkovic, D. (1992). Morphological analysis in word recognition. In R. Frost & L. Katz (Eds.), *Phonology, orthography, morphology, and meaning*. (pp. 343–360). New York: North Holland Elsevier Science Publishers.

Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Forster, K. I. (1978). Accessing the mental lexicon. In R. J. Wales & E. C. T. Walker (eds.), *New approaches to language mechanisms* (pp. 257–287). Amsterdam: North-Holland.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. Walker (Eds.), *Sentence processing: Psychological studies presented to Merrill Garrett* (pp. 27–85). Hillsdale, NJ: Erlbaum.

Forster, K. I. (1981). Priming and the effects of sentence and lexical context on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology, 33A*, 465–495.

Forster, K. I. (1985). Lexical acquisition and the modular lexicon. *Language and Cognitive Processes, 1*, 87–108.

Forster, K. I. (1987). Form-priming with masked primes: The best match hypothesis. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 127–146). Hove, England and Hillsdale, NJ: Erlbaum.

Forster, K. I. (1989). Basic issues in lexical processing. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 75–107). Cambridge MA: MIT University Press.

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., & Davies, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 680–689.

Fowler, C. A., Wolford, G., Slade, R., & Tassinary, L. (1981). Lexical access with and without awareness. *Journal of Experimental Psychology: General, 110*, 341–362.

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.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113*, 256–280.

Gibson, E., Osser, H., Schiff, W., & Smith, J. (1963). An analysis of critical features of letters, tested by a confusion matrix. In *A basic research program on reading* (Cooperative Research Project No. 639). Washington, DC: U.S. Office of Education.

Gibson, J. J., & Gibson, E. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review, 62*, 32–41.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*, 1–67.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 674–691.

Goodman, G., McClelland, K., & Gibbs, R. (1981). The role of syntactic context in word recognition. *Memory & Cognition, 9*, 580–586.

Gough, P. B., Alford, J. A., & Holley-Wilcox, P. (1981). Words and contexts. In O. L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology* (pp. 85–102). Hillsdale, NJ: Erlbaum.

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. *Perception & Psychophysics, 45*, 189–195.

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

Haas, W. (1970). *Phonographic translation*. Manchester: Manchester University Press.

Hanna, P. R., Hanna, J. S., Hodges, R. E., & Rudorf, E. H. (1966). *Phoneme–grapheme correspondences as cues to spelling improvement*. Washington, DC: U.S. Department of Health, Education, and Welfare.

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.

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

Henik, A., Friedrich, F. J., & Kellogg, W. A. (1983). The dependence of semantic relatedness effects upon prime processing. *Memory & Cognition, 11*, 366–373.

Hillinger, M. L. (1980). Priming effects with phonemically similar words: The encoding-bias hypothesis reconsidered. *Memory & Cognition, 8*, 115–123.

Hines, D., Czerwinski, M., Sawyer, P. K., & Dwyer, M. (1986). Automatic semantic priming: Effect of category exemplar level and word association level. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 370–379.

Hirshman, E., & Durante, R. (1992). Prime identification and semantic priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18*, 255–265.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences, 9*, 1–66.

Howard, I. (1972). *A directional theory of rule application in phonology*. Dissertation at Massachusetts Institute Technology.

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

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology (London), 195*, 215–243.

Humphreys, G. W., Besner, D., & Quinlan, P. T. (1988). Event perception and the word repetition effect. *Journal of Experimental Psychology: General, 117*, 51–67.

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. *Brain and Behavioral Sciences, 8*, 689–740.

Humphreys, G. W., Evett, L. J., Quinlan, P. T., & Besner, D. (1987). Priming from identified and unidentified primes. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 105–126). London: Erlbaum.

Humphreys, G. W., Evett, L. J., & Taylor, D. E. (1982). Automatic phonological priming in visual word recognition. *Memory & Cognition, 10*, 576–590.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading; Effects of word frequency. *Perception & Psychophysics, 34*, 49–57.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 130–136.

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. (1990). Naming multisyllabic words. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 92–105.

Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General, 120*, 358–394.

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

Johnston, J. C., & McClelland, J. L. (1973). Visual factors in word perception. *Perception & Psychophysics, 14*, 365–370.

Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin, 96*, 29–44.

Joordens, S., & Besner, D. (1992). Priming effects that span an intervening unrelated word: Implications for models of memory representation and retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18*, 483–491.

Juola, J. F., Leavitt, D. D., & Choe, C. S. (1974). Letter identification in word, nonword, and single letter displays. *Bulletin of the Psychonomic Society, 4*, 278–280.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.

Kay, J., & Bishop, D. (1987). Anatomical differences between nose, palm, and foot, or the body in question: Further dissection of the processes of sub-lexical spelling–sound translation. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 449–470). London: Erlbaum.

Kay, J., & Marcel, A. J. (1981). One process, not two, in reading aloud: Lexical analogies do the work of non-lexical rules. *Quarterly Journal of Experimental Psychology, 37A*, 39–81.

Keefe, D. E., & Neely, J. H. (1990). Semantic priming in the pronunciation task: The role of prospective prime-generated expectancies. *Memory & Cognition, 18*, 289–298.

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

Kiger, J. I., & Glass, A. (1983). The facilitation of lexical decisions by a prime occurring after the target. *Memory & Cognition, 11*, 356–365.

Kinney, G. C., Marsetta, M., & Showman, D. J. (1966). *Studies in display symbol legibility. Part XII. The legibility of alphanumeric symbols for digitized television* (ESD-TR066-117). Bedford MA: The Mitre Corporation.

Koriat, A. (1981). Semantic facilitation in lexical decision as a function of prime–target association. *Memory & Cognition, 9*, 587–598.

Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 92–107.

Krueger, L. E., & Shapiro, R. G. (1979). Letter detection with rapid serial visual presentation. Evidence against word superiority of feature extraction. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 657–673.

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

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293–323.

Lima, S. D. (1987). Morphological analysis in sentence reading. *Journal of Memory and Language, 26*, 84–99.

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition, 6*, 312–319.

Lorch, R. F. (1982). Priming and search processes in semantic memory: A test of three models of spreading activation. *Journal of Verbal Learning and Verbal Behavior, 21*, 468–492.

Lorch, R. F., Balota, D. A., & Stamm, E. G. (1986). Locus of inhibition effects in the priming of lexical decision: Pre- or post-lexical access? *Memory & Cognition, 14*, 95–103.

Luce, P. A., & Pisoni, D. B. (1989). *Neighborhoods of words in the mental lexicon*. Unpublished manuscript.

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

Lupker, S. J. (1984). Semantic priming without association: A second look. *Journal of Verbal Learning and Verbal Behavior, 23*, 709–733.

MacKay, D. (1972). The structure of words and syllables: Evidence from words and speech. *Cognitive Psychology, 3*, 210–227.

Manelis, L. (1974). The effect of meaningfulness in tachistoscopic word perception. *Perception & Psychophysics, 16*, 182–192.

Manso de Zuniga, C. M., Humphreys, G. W., & Evett, L. J. (1991). Additive and interactive effects of repetition, degradation, and word frequency in the reading of handwriting. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading visual word recognition* (pp. 10–33). Hillsdale, NJ: Erlbaum.

Manso de Zuniga, C. M., Quinlan, P. T., & Humphreys, G. W. (1987). *Task effects on priming under masked conditions*. Unpublished manuscript.

Marcel, A. J. (1980). Surface dyslexia and beginning reading: A revised hypothesis of pronunciation of print and its impairments. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 227–258). London: Routledge & Kegan Paul.

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15*, 197–237.

Marcel, A. J., & Patterson, K. E. (1978). Word recognition and production: Reciprocity in clinical and normal studies. In J. Requin (Ed.), *Attention and performance VII* (pp. 209–226). Hillsdale, NJ: Erlbaum.

Marshall, J. C., & Newcombe, F. (1980). The conceptual status of deep dyslexia: An historical perspective. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 1–21). London: Routledge & Kegan Paul.

Massaro, D. M. (1979). Letter information and orthographic context in word perception. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 595–609.

Massaro, D. M., Venezky, R. L., & Taylor, G. A. (1979). Orthographic regularity, positional frequency and visual processing of letter strings. *Journal of Experimental Psychology: General, 108,* 107–124.

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: Erlbaum.

Mayzner, M. S. & Tresselt, M. E. (1965). Tables of single-letter and digram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements, 1,* 13–32.

Mayzner, M. S., Tresselt, M. E., & Wolin, B. R. (1965a). Tables of pentagram frequency counts for various word-length and letter position combinations. *Psychonomic Monograph Supplements, 1,* 144–186.

Mayzner, M. S., Tresselt, M. E., & Wolin, B. R. (1965b). Tables of tetragram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements, 1,* 79–143.

Mayzner, M. S., Tresselt, M. E., & Wolin, B. R. (1965c). Tables of trigram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements, 1,* 33–78.

McCarthy, R., & Warrington, E. K. (1986). Phonological reading: Phenomena and paradoxes. *Cortex, 22,* 359–380.

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

McClelland, J. L., & Johnston, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception & Psychophysics, 22,* 249–261.

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. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2). Cambridge, MA: MIT Press.

McGinnies, E., Comer, P. B., & Lacey, O. L. (1952). Visual-recognition thresholds as a function of word length and word frequency. *Journal of Experimental Psychology, 44,* 65–69.

McKoon, G., & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 18, 1151–1172.

McNamara, T. P. (1992). Theories of priming: I. Associative distance and list context. *Journal of Experimental Psychology: Learning, Memory, and Cognition 18,* 1173–1190.

McNamara, T. P., & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language, 27,* 545–559.

Merikle, P. M. (1982). Unconscious perception revisited. *Perception & Psychophysics, 31,* 298–301.

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

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.

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

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

Monsell, S., Patterson, K., Graham, A., Hughes, C. H., & Milroy, R. (1992). Lexical and sublexical translations of spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 452–467.

Morton, J. (1969). The interaction of information in word recognition. *Psychological Review, 76,* 165–178.

Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory & Cognition, 4,* 648–654.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106,* 226–254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 236–264). Hillsdale, NJ: Erlbaum.

Neely, J. H., & Keefe, D. E. (1989). Semantic context effects on visual word processing: A hybrid prospective/retrospective processing theory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 24, pp. 207–248). San Diego, CA: Academic Press.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Role of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1003–1019.

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

Noble (1952). An analysis of meaning. *Psychological Review, 59,* 421–430.

Norris, D. (1986). Word recognition: Context effects without priming. *Cognition, 22,* 93–136.

Nusbaum, H. C., & Dedina, M. (1985). *The effects of word frequency and subjective familiarity on visual lexical decisions.* Research on speech perception. (Progress Report No. 11). Bloomington: Indiana University.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). *Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words.* Research on speech perception. (Progress Report No. 10). Bloomington: Indiana University.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* Urbana: University of Illinois Press.

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.

Paivio, A., & O'Neill, B. J. (1970). Visual recognition thresholds and dimensions of word meaning. *Perception & Psychophysics, 8,* 273–275.

Palermo, D. S., & Jenkins, J. J. (1964). *Word association norms: Grade school through college.* Minneapolis: University of Minnesota Press.

Patterson, K. E. (1982). The relation between reading and phonological coding: Further neuropsychological observations. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 77–111). San Diego, CA: Academic Press.

Patterson, K. E., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 335–359). Hillsdale, NJ: Erlbaum.

Patterson, K. E., Seidenberg, M., & McClelland, J. L. (1989). Connections and disconnections: Acquired dyslexia in a computational model of reading processes. In R. G. M. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neurobiology* (pp. 131–181). Oxford: Oxford University Press.

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

Petersen, S. E., Fox, P. T., Snyder, A. Z., & Raichle, M. E. (1990). Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science, 249,* 1041–1044.

Peterson, R. R., & Simpson, G. B. (1989). The effect of backward priming on word recognition in single-word and sentence contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1020–1032.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 148–162.

Posner, M. I. (1986). *Chronometric explorations of mind.* Oxford: Oxford University Press.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Prinzmetal, W., Treiman, R., & Rho, S. H. (1986). How to see a reading unit. *Journal of Memory and Language, 25,* 461–475.

Purcell, D. G., Stanovich, K. E., & Spector, A. (1978). Visual angle and the word superiority effect. *Memory & Cognition, 7,* 3–8.

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. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review, 95,* 385–408.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14,* 191–201.

Rayner, K., McConkie, G. W., & Ehrlich, S. F. (1978). Eye movements and integration across fixations. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 529–544.

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

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice-Hall.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81,* 274–280.

Rice, G. A., & Robinson, D. O. (1975). The role of bigram frequency in perception of words and nonwords. *Memory & Cognition, 3,* 513–518.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9,* 487–494.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. (1971). Homographic entries in the internal lexicon: Effects of systematicity and relative frequency of meanings. *Journal of Verbal Learning and Verbal Behavior, 10,* 57–62.

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.

Rumelhart, D. E., & McClelland, J. L. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1). Cambridge, MA: MIT Press.

Samar, V. J., & Berent, G. P. (1986). The syntactic priming effect: Evoked response evidence for a prelexical locus. *Brain and Language, 28,* 250–272.

Savage, G. R., Bradley, D. C., & Forster, K. I. (1990). Word frequency and the pronunciation task: The contribution of articulatory fluency. *Language and Cognitive Processes, 5,* 203–226.

Schendel, J. D., & Shaw, P. (1976). A test of the generality of the word-context effect. *Perception & Psychophysics, 19,* 383–393.

Schreuder, R., Flores d'Arcais, G. B., & Glazenborg, G. (1984). Effects of perceptual and conceptual similarity in semantic priming. *Psychological Research, 45,* 339–354.

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., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 82–102.

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

Seidenberg, M. S. (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.

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,* 447–452.

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

Seidenberg, M. S., Waters, G. S. Sanders, M., & Langer, P. (1984). Pre- and post-lexical loci of contextual effects on word recognition. *Memory & Cognition, 12,* 315–328.

Selfridge, O. G. (1959). Pandemonium: A paradigm for learning. In *Proceedings of a symposium on the mechanisation of thought processes* (pp. 511–526). London: H. M. Stationery Office.

Selfridge, O. G., & Neisser, U. (1960). Pattern recognition by machine. *Scientific American, 203,* 60–68.

Selkirk, E. O. (1980). *On prosodic structure and its relation to syntactic structure.* Bloomington: Indiana University Linguistics Club.

Sereno, J. A. (1991). Graphemic, associative, and syntactic priming effects at a brief stimulus onset asynchrony in lexical decision and naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 459–477.

Seymour, M. K. J. (1986). *Cognitive analysis of dyslexia.* London: Routledge & Kegan Paul.

Shallice, T., & Warrington, E. K. (1980). Single and multiple component central dyslexic syndromes. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 119–145). London: Routledge & Kegan Paul.

Shallice, T., Warrington, E. K., & McCarthy, R. (1983). Reading without semantics. *Quarterly Journal of Experimental Psychology, 35A,* 111–138.

Shulman, G. L. (1990). Relation attention to visual mechanisms. *Perception & Psychophysics, 47,* 199–203.

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.

Slowiaczek, M. L., & Clifton, C. (1980). Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior, 19,* 573–582.

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.

Smith, M. C. (1979). Contextual facilitation in a letter search task depends on how the prime is processed. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 239–251.

Smith, M. C., Theodor, L., & Franklin, P. E. (1983). On the relationship between contextual facilitation and depth of processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 697–712.

Spoehr, K. T., & Smith, E. E. (1973). The role of syllables in perceptual processing. *Cognitive Psychology, 5,* 71–89.

Spoehr, K. T., & Smith, E. E. (1975). The role of orthographic and phonotactic rules in perceiving letter patterns. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 21–34.

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.

Stanners, R. F., Neiser, J. J., & Painton, S. (1979). Memory representation for prefixed words. *Journal of Verbal Learning and Verbal Behavior, 18,* 733–743.

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16,* 32–71.

Stanovich, K. E., & West, R. F. (1983). On priming by sentence context. *Journal of Experimental Psychology: General, 112,* 1–36.

Taft, M. (1979a). Lexical access via an orthographic code: The Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behavior, 18,* 21–39.

Taft, M. (1979b). Recognition of affixed words and the word frequency effect. *Memory & Cognition, 7,* 263–272.

Taft, M. (1981). Prefix stripping revisited. *Journal of Verbal Learning and Verbal Behavior, 20,* 284–297.

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) (pp. 83–123). Orlando, FL: Academic Press.

Taft, M. (1987). Morphographic processing: The BOSS Re-emerges. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 265–279). Hillsdale, NJ: Erlbaum.

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

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior, 15,* 607–620.

Taft, M., & Hambly, G. (1986). Exploring the cohort model of spoken word recognition. *Cognition, 22,* 259–282.

Tanenhaus, M. K., Flanigan, H. P., & Seidenberg, M. S. (1980). Orthographic and phonological activation in auditory visual word recognition. *Memory & Cognition, 8,* 513–520.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words.* New York: Columbia University, Teachers College Press.

Treiman, R., & Chafetz, J. (1987). Are there onset and rime-like units in printed words? In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 281–298) Hillsdale, NJ: Erlbaum.

Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language, 27,* 87–104.

Treiman, R., & Zukowski, A. (1988). Units in reading and spelling. *Journal of Memory and Language, 27,* 466–477.

Treisman, A. M. (1986). Features and objects and visual processing. *Scientific American, 255,* 114–125.

Treisman, A. M., & Schmidt, H. (1982). Illusory conjunctions in the preception of objects. *Cognitive Psychology, 14,* 107–141.

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.

Venezky, R. L. (1970). *The structure of English orthography.* The Hague: Mouton.

West, R. F., & Stanovich, K. E. (1986). Robust effects of syntactic structure on visual word processing. *Memory & Cognition, 14,* 1040–112.

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

Whittlesea, B. W. A. & Cantwell, A. L. (1987). Enduring influence of the purpose of experiences: Encoding-retrieval interactions in word and pseudoword perception. *Memory & Cognition, 15,* 465–472.

Wijk, A. (1966). *Rules of pronunciation for the English language.* Oxford: Oxford University Press.

Winnick, W. A., & Kressel, K. (1965). Tachistoscopic recognition thresholds, paired-associate learning, and free recall as a function of abstractness–concreteness and word frequency. *Journal of Experimental Psychology, 70,* 163–168.