Chapter 9

Visual Word Recognition: The Journey from Features to Meaning (A Travel Update)

David A. Balota, Melvin J. Yap, and Michael J. Cortese

1. INTRODUCTION

1.1. The Word

Well, it is been more than a decade since the literature on visual word recognition has been reviewed for the Gernsbacher (1984a) *Handbook of Psycholinguistics*, and there continues to be considerable interest in understanding the processes tied to the "word." Understanding the journey from features to meaning has clearly made progress, but there is still some distance to go. In the present chapter, we will provide an update on the major issues that were covered in the 1994 chapter, and introduce many new issues that have arisen during the interim. This chapter will focus on isolated visual word recognition research; other chapters in this volume are devoted to auditory word recognition, and recognizing words in sentential context. 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, we 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, we will briefly outline why word recognition research has been central to a number of quite distinct developments in cognitive psychology, psycholinguistics, and cognitive neuroscience. Second, we will 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, we will review the literature on context and priming effects in word recognition; again, highlighting major theoretical developments and controversies. Fourth, we will discuss some limitations regarding inferences that are possible based on the available data, and highlight some recent developments that have provided additional leverage on such issues.

Handbook of Psycholinguistics
ISBN: 0-12-369374-8

1.2. 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 upon. 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 visual word recognition has been devoted to identifying the functional roles of these different levels.

Second, word recognition research has been central 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 historically provided some of the major fuel for arguments regarding self-encapsulated linguistic processing modules (see Fodor, 1983). As we shall see, the issue of how attentional control signals might modulate processes involved in word recognition has received renewed interest recently, and hence, notions of automaticity and modularity have been re-evaluated.

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 identifying the underlying subordinate critical features of a given pattern (e.g., Neisser, 1967). Written words are relatively well-defined patterns. Historically, because words have been the central unit of analysis in much of the verbal learning and memory research that dominated experimental psychology between 1950s and 1960s, there was considerable interest in developing norms that quantify different components of words (e.g., Kučera & Francis', 1967, word frequency norms; Noble's, 1960, meaningfulness norms; Osgood, Suci, & Tannenbaum's, 1957, semantic differential norms). As we shall see, there has been a resurgent interest in developing norms that help quantify the different characteristics of words (see for example, the chapter by Burgess, this volume). Clearly, the importance of the lexical unit in developing models of pattern recognition is due in part to the efforts devoted to defining the stimulus.

AQ1

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 (for example, see Seidenberg & McClelland, 1989). As we shall see, word-level analyses appear to be an ideal battleground for pitting symbolic, rule-based models against connectionist models of cognition.

In sum, word recognition research has been central to work in cognitive psychology and psycholinguistics 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. In fact, word-level analysis was the initial focus of neuroimaging studies (see Petersen, Fox, Posner, Mintun, & Raichle, 1989), and continues to be central to the efforts in the burgeoning field of cognitive neuroscience. 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.

2. FEATURES, LETTERS, AND MODELING CONSTRAINTS

We shall now review some of the variables that have been pursued in word recognition research. First, we shall attempt to break the word down into smaller, more tractable bits. Second, we will discuss work that addresses how orthography maps onto phonology in English. Third, we will discuss the influence of variables that can be quantified at the whole word level, e.g., frequency, familiarity, age of acquisition, orthographic neighborhood size, along with a set of additional semantic variables. Fourth, we will provide an overview of the influence of single word context on isolated word recognition, via a review of the priming literature. Sprinkled within each of these sections will be discussion of the major theoretical models and issues.

2.1. Features

A common approach to understanding 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 initial work in this area was developed by Gibson and Gibson (1955), who forcefully argued that feature-level analyses were an essential aspect of pattern recognition and, more generally, perceptual learning. These primitive features were 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, 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, 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).

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 (and Nobel Prize winning) work by Hubel and Wiesel (1962, 1968). These researchers 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. 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 upon 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.

At the same time behavioral and neural evidence was accumulating in support of features being used in pattern recognition, one of the first computational models of pattern recognition was developed. This was a model of letter recognition developed by Selfridge (1959); Selfridge and Neisser 1960). The model initially coded the stimulus into a set of 28 visual features that provided support for the letters that were most consistent with those features. The Pandemonium model had the capacity to learn which features were especially discriminating among letters, and adjusted the weights for these features accordingly. As we shall see, the Pandemonium model predates by some 20 years important developments in letter and word recognition models. It is quite amazing that the Pandemonium model worked so well given the computational hardware limitations in 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 detected, how does one put the features together to identify the letter T? We typically do not perceive free-floating features (for a review of the binding issue, see Treisman, 1996, 1999). Second, what happens in the feature analytic models when distortions occur that modify the feature, e.g., does a 15° rotated vertical line still activate the vertical line detector? 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). Interestingly, there is some evidence that there may be differences in the way people process printed words and cursive handwriting. For example, case mixing disrupts reading performance with printed words (Mayall, Humphreys, & Olson, 1997) but can actually facilitate performance with handwriting (Schomaker & Segers, 1999). **AQ2** This suggests that distinctive word contours are more critical in handwriting recognition (for a description of a computational model of handwriting, see Schomaker & Van Galen, 1996). 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 binding process (see, however, Shulman, 1990). As we will see the distinction between parallel and serial processing in word recognition has been a central area of debate in recent models (for a recent discussion, see Rastle & Coltheart, 2006). Finally, are features within letters the critical level of analysis in word recognition or are there supraletter and/or even word-level features (e.g., Purcell, Stanovich, & Spector, 1978) more important? Although there has been considerable progress in understanding how features contribute to pattern recognition (for a review, see Quinlan, 2003), there are still many questions that need to be resolved in mapping features onto letters. In lieu of getting bogged down in some of the more important fundamental aspects of visual perception, let us take the leap of faith and assume we have made it to the letter. Surely, things must get a bit more tractable there.

2.2. Letters

Assuming that features play a role in letter recognition, and letters are crucial in word recognition, one might ask what variables are important in letter recognition. For example, does the frequency of a given letter in print influence its perceptability? 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, i.e., letters vs. words. Third, this is our first exposure of a general theme that runs across the word recognition literature, i.e., different tasks or analyses yield different patterns of data, and so it is incumbent upon the researcher to build a task of not only the targeted dimensions in word processing, but also the tasks that are used to tap these dimensions.

2.3. Features, Letters, and Word Interactions: Some Initial Models

Important theoretical issues regarding letter recognition date back to questions that were originally posed by Cattell (1885). The interest here is to define 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, i.e., words are made up of letters. However, Cattell (1885, 1886) reported evidence 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, the conundrum is that if words can be better perceived than letters then how can 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 word processing, and such information is unavailable when single letters are presented. For example, if you thought you saw the letters T and H at the beginning of a short briefly presented word and the letter T at the end then you are likely to guess that there was the letter A between the TH and T, producing the word THAT. This 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 phenomena 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 pre-existing lexical knowledge to bias their response one way or the other (for an alternative view, see Krueger & Shapiro, 1979; Massaro, 1979). Hence, it appeared that subjects actually *see* letters better when embedded in words than when embedded in nonwords. This finding was termed the word-superiority effect and was also reported in a study by Wheeler (1970), so it sometimes also referred to as the Reicher-Wheeler effect.

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 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 historical 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 1 provides an overview of the architecture of the McClelland and Rumelhart (1981) model. Here, one can see the 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 1, 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.

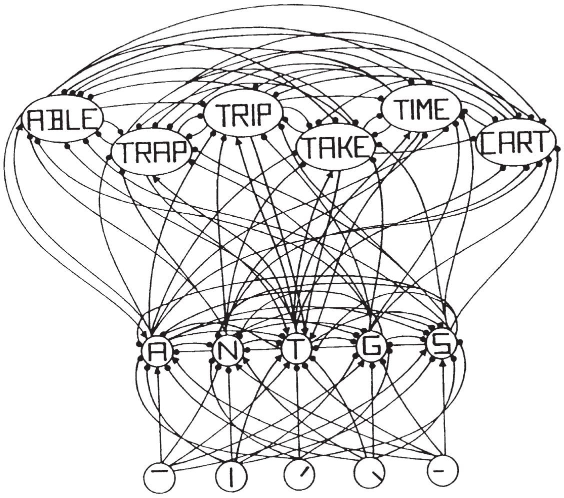

Figure 1. McClelland and Rumelhart's (1981) interactive activation model of letter recognition.

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 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, i.e., 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. Such an effect would appear to be easily accommodated within the McClelland and Rumelhart-type architecture by assuming that there are also top-down influences from the letter level to the feature level.

Interestingly, there is another phenomenon called the pseudoword superiority effect that would at first glance to be problematic for the McClelland and Rumelhart model. Specifically, 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). However, the interactive activation model can also accommodate this effect. Specifically, 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 recent developments in word recognition research.

Although some orthographic regularity effects appear to naturally fall from this model, there are some additional intriguing insights from the model regarding other orthographic regularity effects. 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. (Bigram frequency also has very little influence on other lexical-processing tasks, such as naming or lexical decision; for example, see Andrews, 1992; Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995.) One question that arises from this apparent lack of an influence of bigram frequency is why there are influences of neighbors only 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, Tresselt, & Wolin, 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-type 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 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. However, there are some limitations to such harsh coding. As discussed below in the section on orthographic neighborhood effects, the coding of position of letters within words has become a very active area of research recently.

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 (for an alternative view, see Massaro & Cohen, 1994). A particularly important aspect of this model is that “other” similar lexical-level representations appear to have an influence on the ease of recognizing 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 shall now turn to some discussion of the dimensions that define “similarity” in such networks.

3. 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 distinct issues that arise in this area, one of the major theoretical issues has been the specification of the "rules" that are used 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.

3.1. 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 PHILOSOPHY, and two phonemes in a word like UPHILL. Likewise, even single letters are quite ambiguous such as the C in the word 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 upon 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 do not appear to fit within any predictable class (e.g., CHAMOIS). The important question is to what degree are patterns predictable when one considers similarities across words within the language. For example, some correspondences appear to be relatively invariant (predictable invariant patterns), e.g., 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., /b/p/ and /p/b/), 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 such as in *ambulance* and *amber*, but only one phoneme at word ending positions, such 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 rule-type constraints that are simply based upon 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. Using sophisticated permutation analyses, Kessler and Treiman (2001) have also shown that the spelling-to-sound consistency of a syllabic segment (i.e., onset, vowel, and coda) increases substantially when the other two segments are taken into account. These results also support the contention that English spelling is not as chaotic or irregular as popularly thought (Kessler & Treiman, 2003).

For demonstrative purposes, we have only touched upon 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 system developed by Venezky, and others, Coltheart (1978) estimated that 10–15% of the words would still be unpredictable, i.e., irregular. Likewise, Wijk (1969) 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. Interestingly, Coltheart, Curtis, Atkins, and Haller (1993), using a learning algorithm to generate grapheme-to-phoneme rules, found that these rules mispronounced 22% of monosyllabic words, a figure which is consistent with Henderson's estimate. Also, because of computational limits inherent in two-layer networks (Hinton & Shallice, 1991), the two-layer network of Zorzi, Houghton, and Butterworth (1998) dual-process model was found to be incapable of learning exception words, and it failed to learn correct phonological codes for about 19% of its 2774 monosyllabic word corpus.

Of course, even if one could develop a rule-based system of spelling-to-sound translation that would accommodate all words 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 are critical in fluent word recognition. Hence, instead of providing a detailed discussion of the enormously complex rule systems that have been developed to capture the mapping of orthography onto phonology in English, the present discussion will focus on the empirical evidence regarding how readers use sublexical information in word recognition tasks. The interested reader is referred to Henderson (1982), Wijk (1966, 1969), and Venezky (1970) for excellent treatments of the search for rule-based translations of spelling-to-sound in English (for a description of the algorithms used to identify single-letter, multiletter, and context-sensitive rules, see Coltheart et al., 1993).

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

3.2.1. Dual Route Perspective

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 is possible that there is a second route (the lexical or direct route) that also plays a role in recognizing words. In the second, lexical, route the reader may map the orthographic string onto a lexical

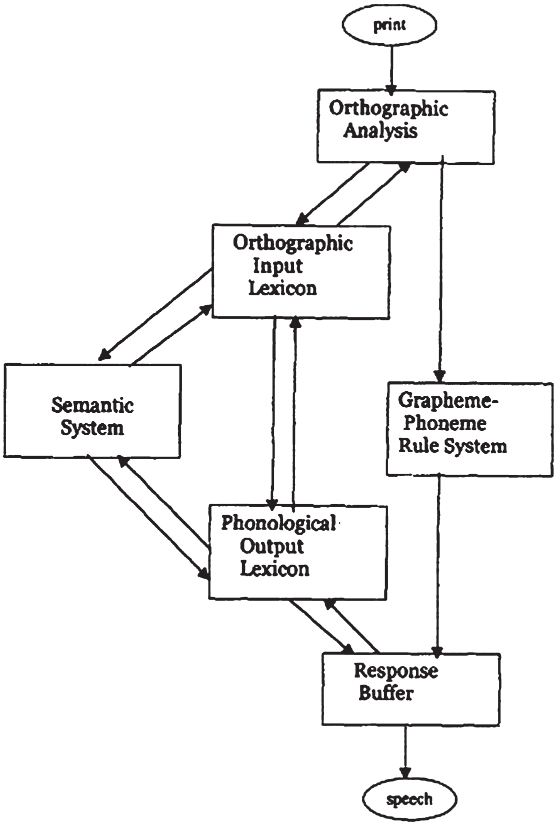

Figure 2. Coltheart et al.'s (2001) DRC model of word recognition.

AQ38

representation and then access the programs necessary for pronouncing a given word aloud either directly from that representation or via access to a semantic representations. Figure 2 displays the dual-route cascaded (DRC) model of word reading developed by Coltheart, Rastle, Perry, Langdon, and Ziegler (2001). In their model, the lexical route is a straightforward extension to the interactive-activation model discussed above. One notable difference is that in the Coltheart et al. model, separate lexicons exist for orthography and phonology.

It is important to note here that because the world's orthographies differ with respect to the regularity of spelling-to-sound correspondences, orthographies also appear to 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-Croation, then one might find little or no impact 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 in 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 the orthographic depth hypothesis). 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. Similarly, there is evidence that readers of a shallow orthography like Serbo-Croatian make lexical decisions based on a prelexically computed phonological code; in contrast, phonological effects are relatively difficult to obtain in English lexical decision (Frost, 1998). 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 is there for a role of an assembled 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 is the relative ease with which individuals can name nonwords (e.g., *blark*) 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-type 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. Activation-synthesis theorists argued that pronunciation performance is always generated via analogies to words represented in the lexicon, thus minimizing an 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 that involved the performance of acquired dyslexics, who appeared to produce a double dissociation between the two routes. Specifically, one class of dyslexics, surface dyslexics, appears 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 *PINT* such that it rhymes with *HINT* (e.g., Marshall & Newcombe, 1980; McCarthy & Warrington, 1986; Shallice, Warrington, & McCarthy, 1983). A second class of acquired dyslexics, phonological (deep) dyslexics, appears 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. Recently, Coltheart et al. (2001) simulated these two acquired dyslexias in the DRC model by selectively lesioning different components of the model. Specifically, surface dyslexia was simulated by lesioning the orthographic lexicon, while phonological dyslexia was simulated by dramatically slowing the sublexical process. These simulations nicely mimicked the neuropsychological data. For example, the degree of impairment of the orthographic lexicon produced regularization error rates that correlated highly with actual regularization error rates exhibited by surface dyslexics of varying severity. Furthermore, the model also correctly simulated the pseudohomophone advantage shown by phonolo gical dyslexics. Specifically, these individuals pronounce pseudohomophones (e.g., BRANE) more accurately than non-pseudohomophonic nonwords (e.g., BRONE), reflecting the larger impact of the lexical route as the influence of the sublexical route is decreased (for a review of pseudohomophone effects in naming performance, see Reynolds & Besner, 2005a).

3.2.2. Parallel distributed processing

Although it would appear that there is compelling evidence for a dual-route architecture, there are important alternative models that have been developed by Seidenberg and McClelland (1989) and Plaut, McClelland, Seidenberg, and Patterson (1996) that also do an excellent job of handling some of the major findings that were originally viewed as strong support for the dual-route model. These parallel-distributed-processing (PDP) models 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 models were 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 models is that the McClelland and Rumelhart model involves localized representations for the major processing codes (i.e., features, letters, and words), whereas, the later models involve distributed representations, e.g., 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 latter models 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 shall now turn to a brief introduction to the Seidenberg and McClelland model, which was the first in a series of parallel distributed processing models of word recognition.

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 of 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 of 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 presented 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 (including 13 homographs, resulting in 2884 unique letter stings) at a rate that is proportional to their natural frequency of occurrence 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 some nonwords that were never presented. Although there is clearly some controversy regarding the degree to which the model actually captures aspects of the data (e.g., see 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 it presented a powerful challenge to the available word-recognition models.

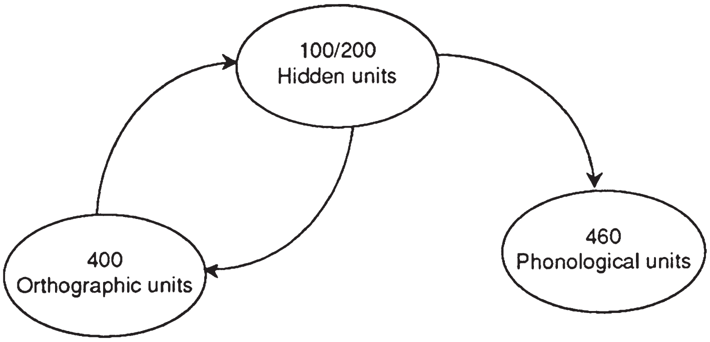

Figure 3. Seidenberg and McClelland's (1989) implemented connectionist architecture.

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, 1984a). The interaction is as follows: for high-frequency words, there is very little impact of the correspondence between orthography and phonology (but see Jared, 1997), whereas, for low-frequency words there is a relatively large impact of such a correspondence. The dual-route framework accommodated this finding by assuming that for high-frequency words the frequency modulated lexical route is faster than the frequency independent 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 (e.g., PINT), 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 provides a natural account for this interaction, this pattern also nicely falls from the Seidenberg and McClelland single route model. That is, the error scores produced by the model (a metric that is assumed to map onto response latencies) for high-frequency regular words and exception words are quite comparable, however, for low-frequency words, the error scores are larger 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 for a specific set of words, and the predictions from Seidenberg and McClelland's model for this same set of words are displayed in Figure 4.

Interestingly, the spelling-sound consistency of a word's neighborhood also influences naming performance, and this neighborhood effect appears to produce an additional influence above and beyond the grapheme-to-phoneme regularity (Glushko, 1979; Jared, McRae, & Seidenberg, 1990). Consistency refers to the degree to which similarly spelled words are pronounced similarly. In particular, studies of consistency have focused on the rime (i.e., the vowel and subsequent consonants in a monosyllabic word). A word that shares both the orthographic rime and phonological rime with most or all of its neighbors is relatively consistent, whereas a word that shares the orthographic rime with its neighbors but has a different pronunciation than most of its neighbors is relatively inconsistent. Regular words that have many "friends" (e.g., *spoon* is consistent because of *moon*, *noon*, etc.) are named faster than regular words that have many "enemies" (e.g., *spook* is inconsistent because of *book*, *took*, etc.). Jared et al. (1990) provided evidence that there are consistency effects in pronunciation primarily under conditions when the neighbors

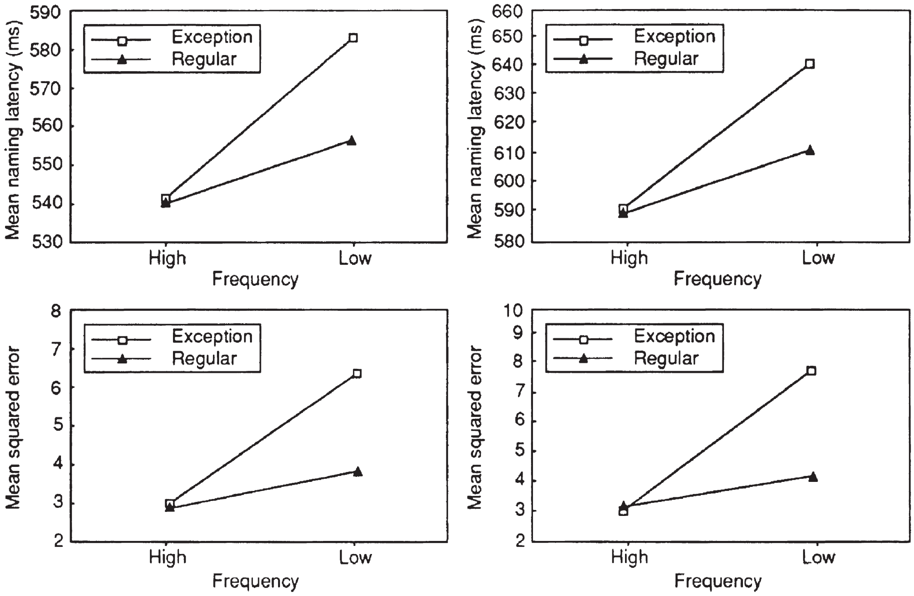

Figure 4. Results and simulations of the Seidenberg (1985, left graph) and Seidenberg et al. (1984a, Experiment 3, right graph) studies: Experimental results (upper graphs) and simulations from the Seidenberg and McClelland (1989) model (lower graphs).

that have consistent spelling patterns (i.e., friends) are higher in frequency than the neighbors that have inconsistent spelling patterns (i.e., enemies). Such neighborhood frequency effects would appear to fall quite nicely from 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 appealing that the Seidenberg and McClelland model can capture such neighborhood effects, without the appeal to rules.

3.2.3. Regularity vs. consistency revisited

AQ3

Because many irregular words (i.e., words whose pronunciation violates GPC rules) are also inconsistent at the rime level, regularity and consistency have typically been confounded. However, these two dimensions are indeed separable (e.g., Andrews, 1982; Kay & Bishop, 1987). Obviously, distinguishing regularity and consistency is important in testing contrasting predictions of models of word recognition. Specifically, the DRC model predicts large effects of regularity and small effects of consistency, and PDP models predict small effects of regularity and large effects of consistency. In general, the results of the studies that have distinguished between consistency and regularity have shown that rime consistency has a larger influence than regularity on latencies and errors than regularity (for discussion, see Cortese & Simpson, 2000; Jared, 2002). In fact, Cortese and Simpson found that the PDP model of Plaut et al. (1996) simulated the naming data on a selected set of words that crossed regularity and consistency better than the Coltheart et al. (1997) DRC model.

3.2.4. Regularity vs. consistency in words and nonwords

Of course, consistency is a continuous variable that can be measured at various levels (e.g., rimes, graphemes). In large-scale studies, Treiman and colleagues (Treiman, Kessler, & Bick, 2002; Treiman et al., 1995) have found that rime-level consistency is a better predictor of word naming performance than grapheme-to-phoneme level consistency. However, it appears that for nonword naming performance, the pattern is a bit more complicated. For example, in contrast to the results by Treiman and colleagues regarding word naming performance, Andrews and Scarratt (1998) reported that nonword reading is affected more by consistency at the grapheme-to-phoneme level than by rime-level consistency. Moreover, in their analysis of 20 nonwords (taken from Seidenberg, Plaut, Petersen, McClelland, & McRae, 1994) in which regularity and consistency pull in opposite directions, Cortese and Simpson (2000) found that grapheme-to-phoneme rules predicted the preferred pronunciation in 14 nonwords, whereas rime consistency predicted the preferred pronunciation in only 5 nonwords. Consider *jind*. The GPC rule for *i* is /I/, but consistency favors the /aInd/ pronunciation found in *bind, blind, hind, mind,* etc. Seidenberg et al. found that 23 of 24 participants pronounced *jind* in a fashion that is consistent with GPC rules. Therefore, it is quite possible that subjects may rely on different types of information when pronouncing a set of nonwords than when processing words.

Zevin and Seidenberg (2005) have recently claimed that consistency effects in nonword naming tasks are more consistent with the PDP perspective than the DRC perspective. By varying the training experience with each new run of a PDP model (also see Harm & Seidenberg, 1999), the model could be tested in terms of pronunciation variability (i.e., the degree to which pronunciations vary across subjects) that is exhibited by college readers (Andrews & Scarratt, 1998; Treiman et al., 2002). Both the PDP model and college readers exhibited considerable variability in their pronunciation of nonwords derived from inconsistent words (e.g., *chead, moup*), but not in their pronunciation of nonwords derived from consistent words (e.g., *nust*). This characteristic is difficult to assess in the DRC model because it is not clear how rules are acquired in the most recent model and how different versions of the model could be implemented. In addition to the insights provided regarding consistency effects in nonword naming, the extension of the models to individual variability, as opposed to overall mean performance, is an important next step in model development.

3.2.5. Feedback consistency

Heretofore, we have been primarily discussing the directional *feedforward* mapping of orthography onto phonology in our consideration of regularity and consistency effects.

For example, PINT is feedforward inconsistent because it does not rhyme with its orthographic neighbors (e.g., *mint, hint, tint,* etc.). However, there is another form of mapping which reflects a *feedback* influence. Specifically, feedback consistency reflects the manner in which a specific phonological pattern is spelled in different ways. Figure 5 illustrates the syllabic structure derived by linguistic distinctions between onsets and rimes (see further discussion below) and also shows how consistency can be computed along four dimensions: (a) feedforward onset, (b) feedforward rime, (c) feedback onset, and (d) feedback rime. For example, the rime in *tone* is feedback inconsistent because /on/ is spelled OWN as in GROWN, and OAN, as in MOAN. As one might guess, many words are inconsistent in both directions. Stone, Vanhoy, and Van Orden (1998) first decoupled feedforward consistency from feedback consistency and the effects of both variables were obtained in lexical decision performance. In addition, reliable and equivalent feedback consistency effects were reported by Balota, Cortese, Sergent-Marshall, Spieler, and Yap (2004) for lexical decision and naming performance, whereas in French Ziegler, Montant, and Jacobs (1997) found larger feedback consistency effects in lexical decision than in naming performance. The influence of feedback consistency in visual word recognition is theoretically important because it suggests that phonological activation provides feedback onto the orthographic representation (also see Pexman, Lupker, & Jared, 2001) during isolated visual word processing. However, it should also be noted that there is currently some debate regarding the unique effect of feedback consistency. For example, Peereman, Content, and Bonin (1998) have argued that feedback consistency effects in French are eliminated when familiarity is controlled (also see Kessler, Treiman, & Mullennix, 2005).

AQ4

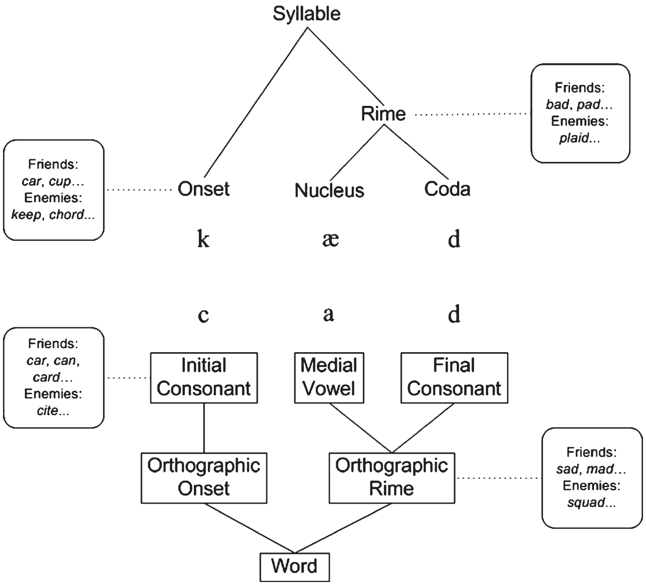

Figure 5. Feedforward and feedback onset and rime organization for single syllabic structure.

3.2.6. Potential problems with Seidenberg and McClelland model

Consistency effects would appear to arise naturally from the PDP architecture developed by Seidenberg and McClelland (1989). Although this model provided an interesting alternative to the dual-route model, it also generated a number of important problems that needed to be resolved (for a discussion of these issues, see Coltheart et al., 1993). First, it is unclear how such a model might handle the fact that some acquired dyslexics appear to only have an intact assembled route, while others appear to only have an intact lexical route (for some discussion of this issue, see Patterson, Seidenberg, & McClelland, 1989). Second, as described below, it appears that 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. Third, Besner (1990) and Besner et al. (1990a, 1990b) have documented that the phonological error scores and the orthographic error scores do a rather poor job of simulating some characteristics of nonword performance. Fourth, the Seidenberg and McClelland model mapped error scores indirectly onto response latency instead of providing a direct metric for response latency.

3.2.7. Further developments in the PDP architecture

In response to the challenges to the Seidenberg and McClelland (1989) model, Plaut et al. (1996) substantively updated the representations and architecture of the PDP model (henceforth PMSP96) to address these problems. First, the model incorporated improved orthographic and phonological representations that allow it to not only correctly pronounce all the monosyllabic words in the training corpus, but also to name nonwords with a much greater facility. Second, in Seidenberg and McClelland's (1989) "triangle" model framework (see Figure 6), skilled reading is supported by the joint contributions of a phonological and a semantic pathway. Although only the phonological pathway was implemented in the Seidenberg and McClelland mode, a prototype semantic pathway was implemented in PSMP96 (see Simulation 4), which may be useful in accommodating meaning-level influences in word recognition (see Strain, Patterson, & Seidenberg, 1995). Third, the authors extensively discussed how the network could handle the acquired dyslexia data that, as described above, was central to the development of the DRC model. For example, it is possible to simulate phonological dyslexia, i.e., better word reading than nonword reading, by a selective impairment of the phonological pathway. Similarly, surface dyslexia, i.e., normal nonword but impaired exception word reading, was satisfactorily simulated after training a new network that incorporated an isolated, semantically supported phonological pathway. In normal readers, the semantic and phonological pathway work together to support the pronunciation of exception words. Should the semantic pathway be damaged, the semi-competent isolated phonological pathway manifests symptoms similar to that of surface dyslexia (Plaut, 1997). Of course, at this point one might ask whether the inclusion of a semantic "route" makes the PDP model functionally equivalent to a dual-route model. For example, does the network, over the course of training, partition itself into two sub-networks, one that handles regular words, and one that handles exception words? Plaut et al. (1996) tested this intriguing hypothesis and found little support for this contention. Generally, the system did not fractionate

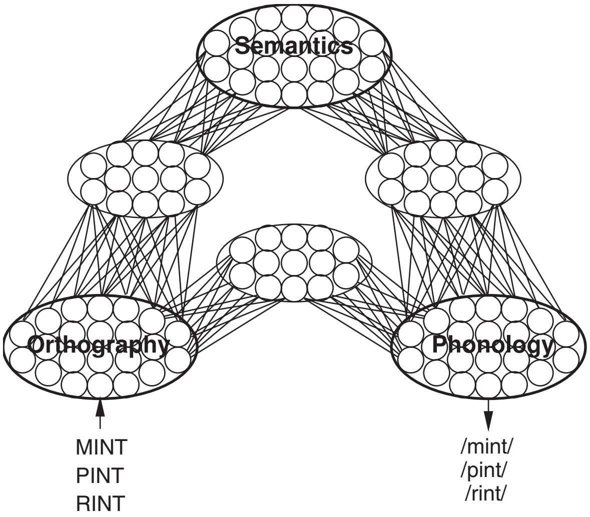

Figure 6. Seidenberg and McClelland's (1989) triangle connectionist framework for lexical processing.

itself such that one part learned spelling-sound rules and another part encoded the exceptions to these rules, but both components contributed to performance. Finally, it is important to note that the Plaut et al. model was a recurrent network that eventually settled into a steady state and hence response latencies could be evaluated in the model. That is, activation of phonological units changes over time as information in the system accumulates and is shared among network units. This property contrasts with the error measure that was used to evaluate the original Seidenberg and McClelland model. In the Plaut et al. model, when a word is recognized, its corresponding grapheme units become activated, and, in turn, this activation is propagated throughout the network.

Considerable debate continues between advocates of PDP and DRC approaches to word recognition. Although the PDP models seem ideally suited for handling consistency effects, the DRC model is particularly adept at handling data consistent with serial processing. Consider, for example, the *position of irregularity effects*. An irregular/inconsistent word can be irregular/inconsistent at the first phoneme position (e.g., *chef*), the second phoneme position (e.g., *pint*), the third phoneme position (e.g., *plaid*), or beyond (e.g., *debris*). Because the sublexical process in the DRC operates in a serial fashion, it is more prone to earlier than later irregular/inconsistent sublexical interference. In contrast, the PDP model processes words in parallel, and so does not predict a position of irregularity effect. Although there have been some methodological concerns noted (see Cortese, 1998; Zorzi, 1999), the evidence indicates that latencies are longer for words that have early irregular/inconsistent patterns than late irregular/inconsistent patterns (e.g., Coltheart & Rastle, 1994; Cortese, 1998; Rastle & Coltheart, 1999). These results, along with others (see Rastle & Coltheart, 2005), appear to support the serial component of the DRC-type framework over the parallel nature of the PDP framework. Of course, it is possible that such effects may ultimately reflect input or output processes beyond the scope of the currently implemented PDP framework. Indeed, Seidenberg (2005) acknowledges that important challenges exist, but the PDP approach accounts for many behavioral phenomena as well as providing a more natural interface between reading and underlying principles of the nervous system. In this light, Seidenberg argues that the PDP architecture is important because it generalizes well to other cognitive domains. However, a proponent of the DRC approach would argue that such generality should not outweigh the fact that the devil is in the details of the fit of a particular model of word recognition with the available evidence (see, for example, Rastle & Coltheart, 2005). Clearly, the debate continues.

More recent connectionist models of reading have shifted their emphasis from understanding how people *pronounce* letter strings aloud to understanding how *meaning* is computed (Seidenberg, 2005). For example, Harm and Seidenberg (2004) proposed a model that considers how the meaning of a word is computed by orthographic and phonological processes working cooperatively. It is also apparent that one glaring limitation of both dual-route and connectionist models is their inability to process multisyllabic words. One model that has made some progress in this respect is the connectionist multi trace memory model of Ans, Carbonnel, and Valdois (1998). While a full description of this interesting model is beyond the scope of this chapter, the Ans et al. model proposes two sequential procedures for reading: first, a holistic procedure that draws on knowledge about entire words, and if that fails, an analytic procedure that is dependent on the activation of subsyllabic segments. Including two reading procedures allows the model to name monosyllabic words, multisyllabic words, and nonwords, and also allows it to account for dissociations between skilled and pathological reading. Although the Ans et al.'s model may *prima facie* resemble the dual-route model, it does not compute phonology from orthography using different computational principles. Instead, pronunciation is always supported by the memory traces laid down by previously encountered exemplars.

3.3. Superletter Sublexical Codes : What's the Evidence for their Functional Role?

At this point, it should be noted that we have yet to discuss specific types of sublexical but supraletter influences on word recognition. We have generally grouped together a set of effects under the regularity/consistency umbrella, focusing on the theoretical implications of such effects for current models. We shall 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 rimes, morphology, and syllables. The goal here is to simply acquaint the reader with the attempts that have been used to decompose the sublexical units.

3.3.1. Onsets and rimes

As noted earlier, researchers have made a distinction between the onset and rime unit within syllables. For example, Treiman and her colleagues (e.g., Treiman, 1989; 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 both in speech production and recognition of visual strings. This subsyllabic distinction is between the onset and rime 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 rime of a word involves the following vowel and any subsequent consonants. For example, in SIP, SLIP, and STRIP, /IP/ would be the rime. Thus, syllables have a subsyllabic organization in that each syllable is composed of an onset and a rime.

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 rimes in English. For example, there is evidence from the types of speech errors that speakers produce (Dell, 1986; MacKay, 1972), the distributional characteristics of phonemes within syllables (Selkirk, 1982), along with the types of errors that subjects produce in short-term memory tasks (Treiman & Danis, 1988). Thus, the support for the onset and rime distinction clearly extends beyond the work in visual word recognition, and is driven more by phonological principles that have been developed in linguistics.

In one of the first studies addressing onset and rime organization in 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 rime. Now, consider performance on conditions where the strings again correspond to words but they are not broken at onsets and rimes. 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 rimes. The results of the Treiman and Chafetz experiments indicated that anagram solutions were better when the breaks corresponded to onset-rime 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-rime division (e.g., CR//ISP, TH//ING) or a break that did not match the onset-rime division (e.g., CRI//SP and THI//NG). The results indicated that lexical decisions were reliably faster when the break matched the onset-rime division. Thus, Treiman and Chafetz argued that onset and rime units play a role in visual word recognition.

3.3.2. Syllables

If the distinction between onsets and rimes 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 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 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 into (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; i.e., most subjects can accurately decompose most words into syllables. The more important issue is whether this information is used in accessing the lexicon for visually presented words.

Prinzmetal, Treiman, and Rho (1986) reported an intriguing 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 each of the 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 in the same color of the O, compared to the K, whereas, the V in ANVIL might be more likely to be reported in the color of the I, compared to the N. This is precisely the pattern obtained in the Prinzmetal et al. study.

It is interesting to note here that Adams (1981) provided evidence that the letters that border adjacent syllables often 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. Although Seidenberg (1987, Experiment 3) provided some initial evidence that the effects observed in the original Prinzmetal et al. paradigm were due to such bigram troughs, as opposed to actual syllabic boundaries, more recent work by Rapp (1992) found that one can obtain syllabic effects even when one controls for such bigram troughs.

The role of the syllable has not been implemented in most models of word recognition that have been primarily built to process monosyllabic words. One exception to this is the connectionist model proposed by Ans et al. (1998, discussed earlier). Based on the evidence discussed above and the findings from the literature on spoken word processing (e.g., Stevens & Blumstein, 1978), this model parses words into syllabic units in the phonological output. Presumably, this phonological output could serve as an access to a semantic system; however, this was not implemented in the current model. AQ5

More recently, Rastle and Coltheart (2000) have proposed a complex set of rules for syllable segmentation, stress assignment, and vowel reduction for disyllable words in their DRC model. In their study, the assignment of stress to a set of nonwords by the model was similar to that provided by human subjects. Also, words that violated the rules resulted in longer naming latencies, an effect that is consistent with predictions of the DRC model. However, it is important to note that, like previous studies on regular and irregular monosyllabic words, regularity in the Rastle and Coltheart study was confounded with spelling-sound consistency (Chateau & Jared, 2003). In their naming study of disyllabic words, Chateau and Jared found that the feedforward consistency of the segment containing the first vowel grapheme and subsequent consonants and the second vowel grapheme predicted naming latencies and errors. Moreover, the consistency measures derived by Chateau and Jared nicely predicted the outcome reported by Rastle and Coltheart. Of course, if readers use sublexical rules when processing multisyllabic words, the DRC model would be better equipped to explain such a result, but consistency effects are better handled by PDP models. Clearly, more research on multisyllabic words is necessary to determine both the behavioral influence of syllables and stress patterns en route to word recognition and also the best way to model such effects.

3.3.3. 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 nonword 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. Linguistic models of lexical representation assume that there is some base form of representation and a set of rules that are 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) en route to word recognition. As in the case of syllables, we are not questioning whether morphemes are represented in the processing system, the question is whether morphemic analyses play a role in processes tied to visual word recognition.

Much of the early 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 was 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, etc. 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, 1979). This would appear to support the contention that root morphemes do play a special role in word recognition and it is not simply the raw frequency of the actual lexical string that is crucial.

Another approach to morphological analyses in word recognition involves long-term morphemic priming (e.g., Stanners, Neiser, & Painton, 1979a, 1979b). In these studies, subjects are most often presented a sequence of lexical decision (word/nonword) 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 (1979a, 1979b) 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 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 (for a summary of evidence favoring non-orthographic accounts of morphemic priming effects, see review by Feldman & Andjelkovic, 1992).

Because the PDP perspective has achieved prominence as a general theory of language processing, research on morphological decomposition has taken on new theoretical significance. One main reason that this topic has received such attention is that distinct morphemic representations do not exist in PDP models (e.g., Plaut & Gonnerman, 2000; Rueckl, Mikolinski, Raveh, Miner, & Mars, 1997). Rather, morphemic effects are thought to emerge from interactions among orthography, phonology, and semantics (Gonnerman, Seidenberg, & Andersen, 2005). A recent cross-modal lexical decision study by Gonnerman et al. (2005) found support for this view. They reported that facilitation for visually presented targets was related to the semantic and phonological overlap found in prime–target pairs. In contrast, morphemic overlap did not produce additional facilitation above and beyond semantically and phonologically related items. For example, *sneer* facilitated *snarl* to the same degree as *teacher* facilitated *teach.* Also, weakly related pairs (e.g., *lately-late*) produced less facilitation than strongly related pairs.

AQ6

Interestingly, Rastle, Davis, and New (2004) reported a morphological effect that was independent of semantics. In their lexical decision study, masked primes (presented for 42 ms) that maintained a morphological relationship only (e.g., *corner-corn*) facilitated targets as much as primes that maintained both a semantic and morphological relationship with the target (e.g., *cleaner-clean*), whereas a control condition (e.g., *brothel, broth*) did not produce priming. Thus, it appears from the Rastle et al. study that decomposition is somewhat independent of the semantic information available from the stem. This outcome seems more consistent with localist models (e.g., the DRC model) than distributed models (e.g., PDP models). However, given the Gonnerman et al. results discussed above, it is clear that further work is needed on this important topic.

We have only touched upon some of the very interesting issues that have arisen in morphological analyses in visual word recognition. We suspect that this will be an area of very active research in the future, and refer the reader to Baayen and Schrender (2003), Feldman and Basnight-Brown (2005), and Sandra and Taft (1994) for more comprehensive treatments of this important area. **AQ7**

4. LEXICAL-LEVEL VARIABLES

By lexical-level variables, we refer to the impact 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 a given word (e.g., DOG) on word recognition task performance.

4.1. Length

One might ask whether there is a word length effect in visual 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. Interestingly, there has been some disagreement on this simple topic. There is clear evidence that longer words take more time in perceptual identification (McGinnies, Comer, & Lacey, 1952), and produce longer fixation durations in reading (see Just & Carpenter, 1980), but the effect of length in lexical decision naming have been a bit more inconsistent (for a review, see New, Ferrand, Pallier, & Brysbaert, 2005).

The role of letter length in naming performance has been the focus of a number of recent studies. For example, Gold et al. (2005) found that individuals with a loss of semantic/lexical input, produced exaggerated length effects, compared to individuals with dementia of the Alzheimer's type. Gold et al. suggested that these results may be supportive of greater reliance on the serial sublexical route in individuals with semantic dementia. Consistent with this possibility, Weekes (1997) found length effects for nonwords and no length effects for words. Coltheart et al. (2001) interpreted the Weekes results as being critical to the DRC, i.e., the small or non-existent length effects for words is due to the parallel pathway used in the lexical route, whereas, the large length effects for nonwords reflects the serial analysis demanded by the sublexical route. In a study of speeded naming performance of over 2400 single syllable words, Balota et al. (2004) obtained clear effects of length that were modulated by word frequency. Low-frequency words produced larger length effects than high-frequency words.

There is some controversy regarding length effects in the lexical decision task. Because the lexical decision task has been taken as a premier task to develop word recognition models, this is a troublesome finding (for a review, see Henderson, 1982). Chumbley and Balota (1984) reported relatively large length effects in the lexical decision task when the word and nonwords were 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 using a relatively small range of lengths of words. In this light, the recent study by New et al. (2005) is noteworthy. Specifically, they analyzed length effects in a dataset of 33,006 of lexical decision latencies taken from Balota et al. (2002). They found an interesting quadratic relationship between length and lexical decision performance, such that there was a facilitatory effect from 3 to 5 letter in length null effect for 5–8 letters in length and a clear inhibitory effect for 8–13 letter words. The long words appear to demand some serial processing. Interestingly, the short words indicate that there may be an ideal length, based on the average length of words, and that very short words actually may produce a decrement in performance. Finally, it should also be noted that frequency does appear to modulate the length effect, since Balota et al. (2004) reported that length effects were larger in lexical decisions for low-than high-frequency words, similar to the pattern obtained in speeded naming performance mentioned above. Thus, the effects of word length in lexical decision performance appear to depend on both the frequency and the particular lengths of the words.

4.2. Word Frequency

The frequency with which a word appears in print has an influence on virtually all word recognition tasks. For example, word frequency effects have been found in lexical decision performance (e.g., Forster & Chambers, 1973), naming performance (e.g., Balota & Chumbley, 1984), perceptual identification performance (e.g., Broadbent, 1967), and online reading measures such as fixation duration and gaze duration measures (e.g., Rayner & Duffy, 1986; Schilling, Rayner, & Chumbley, 1998). This, of course, should not be surprising because printed word-frequency should be related to the number of times one experiences a given word; experience with an operation should influence the ease of performing that operation.

Although it would appear to be obvious why word-frequency would modulate performance in word recognition tasks, the theoretical interpretation of such effects have been quite varied. For example, the activation class of models based in large part on Morton's (1969, 1970) classic Logogen model, assume that 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. Coltheart et al. (2001) DRC model nicely captures frequency effects via the activation patterns in the lexical route. the PDP models of Seidenberg and McClelland (1989) and Plaut et al. (1996) assume that frequency is coded in the weights associated with the connections between the units. Interestingly, there are hybrid models (e.g., Zorzi et al., 1998), which implement lexical and sublexical processing using connectionist principles.

A third class of word recognition models that we have yet to describe are 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 7, Forster (1976) has argued that the lexicon may be searched via several indexing systems: orthographic, phonological, syntactic/semantic access bins. Each of these bins involves a frequency ordered search, i.e., 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 lexicon representation. Although such a model may seem cumbersome, Murray and Forster (2004) have recently provided intriguing evidence supporting this position, since rank frequency (as in rank in the search bin) appears to be a better predictor of word-frequency effects than actual log frequency values. It is noteworthy that there are additional models that are hybrids of the activation and search models such as in the Becker (1980), Paap et al. (1982), and the Taft and Hambly (1986) models. For example, Becker suggests that 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.

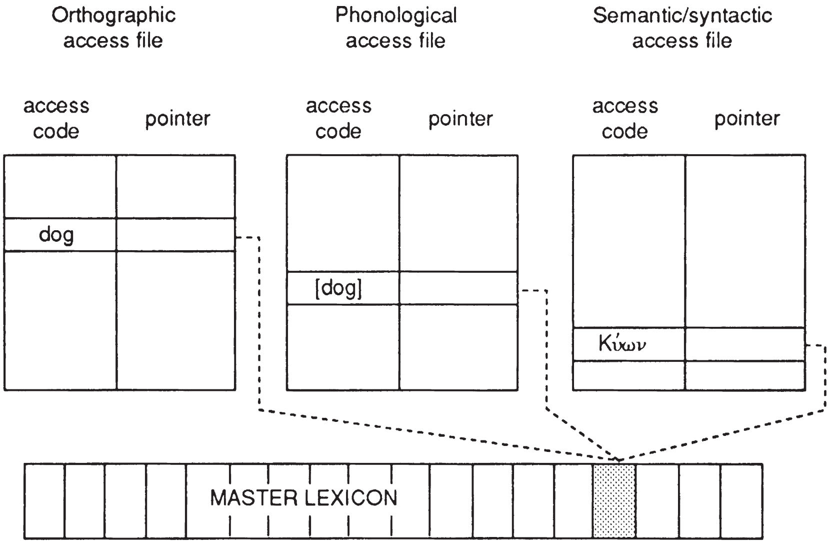

Figure 7. Architecture of Forster's (1976) serial search model of word recognition.

An important question that has arisen regarding word frequency effects is the locus of the effect in the tasks used to build models of word recognition. The models mentioned above all suggest that frequency is central to the interworkings of the models, as it should be. However, there is also 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 (Balota & Chumbley, 1985; Connine, Mullennix, Shernoff, & Yelens, 1990), and sophisticated guessing aspects of the threshold identification task (Catlin, 1969, 1973) that are likely to exaggerate the influence of word frequency. Because of the importance of task analyses we will use this as an opportunity to review some of these issues regarding the lexical decision task.

AQ8

Consider, for example, the Balota and Chumbley (1984) model of the lexical decision task displayed in Figure 8. Balota and Chumbley have suggested that because of the demands of the task, subjects place particular emphasis on two pieces of information that are obvious discriminators between words and nonwords, i.e., the familiarity and meaningfulness (FM dimension) of the stimuli. 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 is probably 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 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 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. Hence, one should expect a larger influence of word-frequency in the lexical decision task than in the naming task, and in general this is what is found

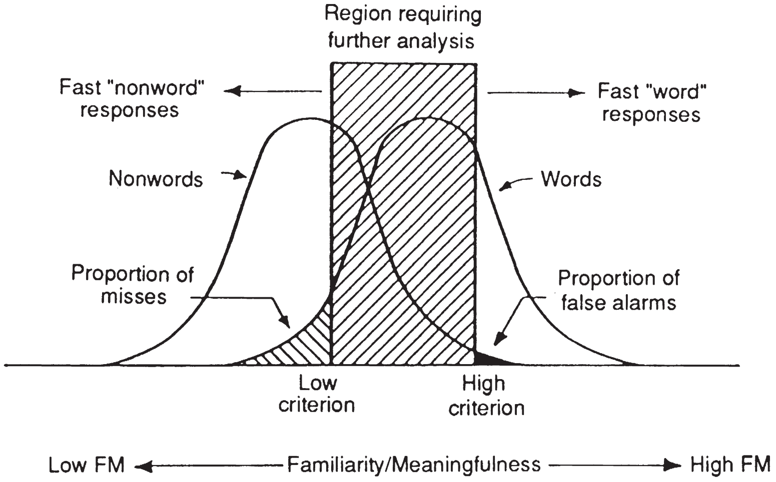

Figure 8. Balota and Chumbley's (1984) two-stage model of the lexical decision task.

(see Balota et al., 2004). Balota and Spieler (1999) have implemented a hybrid model of the lexical decision tasks that not only accommodates word-frequency effects and other effects, but also accounts for the reaction time distributional aspects of performance, i.e., the shape of the reaction time distribution. It is also important to note that Ratcliff, Gomez, and McKoon (2004) have also argued that decision processes tied to the lexical decision task are critical in understanding word frequency effects, along with other variables, and have nicely modeled such effects with a single-process diffusion model.

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 Andrews & Heathcote, 2002; Balota & Chumbley, 1990; Monsell, Doyle, & Haggard, 1989). Of course, the primary intent of the task analysis work is 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 processes involved in word recognition, and hence will need to be incorporated into all models of word recognition. However, as exemplified throughout this review, understanding the operations in the tasks used to build models of word recognition is a paramount first step in building adequate models.

4.3. Familiarity

A variable that is highly correlated with frequency is word familiarity. Familiarity is typically based on untimed ratings. For example, subjects may be asked to rate each word on a 7 point scale ranging from extremely unfamiliar to extremely familiar with this word. The importance of familiarity norms was motivated by Gernsbacher (1984a, 1984b) who persuasively argued that the available printed word frequency norms by Kučera and Francis (1967) and Thorndike and Lorge (1944) may not be the most sensitive estimates of the impact of frequency of occurrence on lexical representations. For example, frequency norms typically do not take into account spoken word frequency, and are based on dated and relatively limited samples of word use. Gernsbacher (1984a, 1984b) pointed out that *boxer*, *icing*, and *joker* have the same objective frequency value (according to Kučera & Francis, 1967) as *loire*, *gnome*, and *assay.* Recently, there are a number of more extensive norms that have been developed based on a multifold increase in the sample size compared to the original norms (e.g., Baayen, Piepenbrock, & Van Rijn, 1993; Burgess & Livesay, 1998; Zeno, Ivens, Millard, & Duvvuri, 1995). As one might expect, when comparing different frequency norms, the more recent norms are better predictors of both naming and lexical decision performance than the still commonly used Kučera and Francis (1967) norms (see Balota et al., 2004; Zevin & Seidenberg, 2002). Hopefully, cognitive science researchers who are investigating or controlling word frequency will begin to use these more recent norms. AQ9

Although the norms are becoming better, it is still the case that they are only a proxy for frequency of exposure. Hence, researchers still argue that subjective familiarity ratings are a better measure of sheer exposure to a word. However, one might ask what sorts of information do subjects use when making an untimed familiarity rating? Standard instructions for familiarity ratings tend to be vague and may encourage the use of other types of information, such that more meaningful stimuli tend to be rated more familiar. In fact, Balota, Pilotti, and Cortese (2001) found that the familiarity ratings of Toglia and Battig (1978) were related to meaningfulness, a semantic variable. As an alternative to standard familiarity ratings, Balota et al. (1999) had participants rate monosyllabic words in terms of subjective frequency. Participants estimated how often they read, heard, wrote, said, or encountered each word based on the following scale: 1 = never, 2 = once a year, 3 = once a month, 4 = once a week, 5 = every two days, 6 = once a day, 7 = several times a day. Balota et al. found that these ratings were less influenced by meaningfulness than the Toglia and Battig (1978) familiarity ratings. Hence, subjective frequency ratings may be more appropriate than traditional familiarity ratings because they are less influenced by semantic factors. Indeed, Balota et al. (2004) found that the subjective ratings produced a powerful predictive unique influence on both lexical decision and naming performance above and beyond a host of other correlated variables, such as objective word frequency, length, neighborhood size, spelling-to-sound consistency, etc.

4.4. Age of Acquisition

Within the past decade there has been considerable interest in the influence of the age at which words are acquired on various measures of lexical processing (for a recent review, see Juhasz, 2005). There have been a number of reports suggesting that age of acquisition (AoA) produces a unique influence on word recognition performance (e.g., Brown & Watson, 1987; Morrison & Ellis, 1995) above and beyond correlated variables such as word frequency. The intriguing argument here is that early acquired words could play a special role in laying down the initial phonological and/or semantic representations that the rest of the lexicon is built upon. Moreover, early acquired words will also have a much larger cumulative frequency of exposure across the lifetime.

There are at least two important methodological issues regarding AoA effects. The first concerns the extent to which AoA affects performance in word recognition tasks like naming and lexical decision. One of the problems with assessing this issue is that AoA is correlated with many other variables, including length, frequency, and imageability. Moreover, one might expect an AoA effect not because early acquired words have a special influence on the lexicon, but rather because early acquired words have a greater cumulative frequency effect, even when objective frequency is held constant (for example, see Lewis, Gerhand, & Ellis, 2001). Although most studies have not teased these apart the influence of cumulative frequency, Juhasz and Rayner (2003) found a unique effect of AoA in eye fixation data in reading and Bonin, Barry, Méot, and Chalard (2004) have demonstrated significant effects of objective AoA in word naming and lexical decision performance.

The second issue is concerned with whether or not AoA should be considered an outcome variable (Zevin & Seidenberg, 2002, 2004) or a standard independent (or predictor) variable. Zevin and Seidenberg have argued that AoA predicts word recognition performance because the age at which a word is learned is affected by many factors, and hence, this is related to the correlated variables issue noted above. They focus on frequency trajectory, which reflects the distribution of exposures that one has with words over time. Some words such as *potty* occur fairly frequently during early childhood but not adulthood, whereas other words such as *fax* occur frequently during adulthood, but not childhood. Therefore, frequency trajectory should influence AoA, and indeed the two variables are correlated. In addition, Zevin and Seidenberg (2004) examined the influence of frequency trajectory and cumulative frequency in naming. They found little evidence for frequency trajectory, whereas cumulative frequency produced a unique effect on naming performance (for an alternative interpretation of the Zevin and Seidenberg findings, however, see Juhasz, 2005). Given the potential theoretical importance of AoA, it appears that this variable will continue to be at the center of considerable empirical and theoretical work in the next several years.

4.5. Orthographic Neighborhood Effects

Although estimates vary, the average adult reader is likely to have at least 50,000 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 from other orthographically related representations.

Coltheart, Davelaar, Jonasson, and Besner (1977) introduced the orthographic neighborhood or 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. For example, the orthographic neighbors of the word FALL include MALL, FELL, FAIL, BALL, FULL, CALL, among others. There are two major ways that researchers have investigated the influence of N. First, consider the influence of the sheer number of orthographic neighbors. In naming performance, the results are rather straightforward: as the number of orthographic neighbors increases, response latency decreases, and this effect is larger for low-frequency words than high-frequency words (see Andrews, 1989, 1992; Balota et al., 2004). In contrast, in lexical decision performance, increases in N increase response latencies to nonwords, and for word targets the results range from facilitatory Andrews (1989, 1992; Forster & Chen, 1996) to no effect (Coltheart et al., 1977) to some conditions producing inhibitory effects (see, for example, Johnson & Pugh, 1994). In an excellent review of this literature, Andrews (1997) has argued that the variance across the studies of orthographic neighborhood size in lexical decision appears to be in part due to variability in list contexts (e.g., nonword type). It should also be noted that there is evidence of facilitatory effects of large Ns in semantic classification studies (Forster & Shen, 1996; Sears, Hino, & Lupker, 1999). Finally, it should be noted that the evidence from eye-fixation patterns while people are reading indicate that there is an inhibitory effect of words with large Ns. Importantly, Pollatsek, Perea, and Binder (1999) have shown that with the same set of words that produces facilitatory effects in lexical decision performance, these words produce inhibitory effects in eye-fixation durations. Clearly, the effects of orthographic N are highly dependent upon the task constraints, and most likely a host of other variables such as individual processing speed (see, e.g., Balota et al., 2004).

A second way to investigate the influence of orthographic neighborhoods is to consider the frequency of the neighbors, i.e., does the stimulus have higher-frequency neighbors or lower-frequency neighbors? In lexical decision performance, there is evidence that targets with higher-frequency neighbors indeed produce inhibition in lexical decision performance, compared to words with lower-frequency neighbors (e.g., Grainger, 1990, 1992; Grainger & Jacobs, 1996; Carreiras, Perea, & Grainger, 1997, but see Pollatsek et al., 1999). However, there is even some conflict here, because in a series of experiments involving both naming and lexical decision performance, Sears et al. (1995) found facilitation for low-frequency targets with large neighborhoods and higher-frequency neighbors. **AQ10**

Given that word recognition unfolds across time, it is not surprising that both frequency of the neighbors and the size of the neighborhoods should play a role in word recognition tasks. In this light, it is useful to mention the 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 is noteworthy that recent simulations by Sears et al. (1999) have shown that both the Plaut et al. (1996) and the Seidenberg and McClelland (1989) models appear to predict facilitatory effects of neighborhood size that are greater for low-frequency words than for high-frequency words, which is overall most consistent with the data in this area.

Facilitatory neighborhood effects for low-frequency words would appear to be difficult to accommodate within models that have a competitive interactive activation component (e.g., the DRC model of Coltheart et al., 2001, or the multiple read-out (MROM) model of Grainger & Jacobs, 1996). Specifically, the larger the neighborhood, the more competition one should find. Moreover, the facilitatory effects of N produce particular difficulties for serial search models, such as Forster's classic bin model. Specifically, the more items that need to be searched, the slower response latency should be. This is opposite to the most common pattern reported in this literature.

An interesting variation on the influence of orthographic N is the transposed letter effect. Specifically, Chambers (1979) and Andrews (1996) found that words like SLAT produce slower response latencies in lexical decision performance, because these items have a highly similar competitor SALT. Andrews (1996) also found this pattern in naming performance. Note that SLAT is not an orthographic neighbor of SALT, but is very similar because two letters in adjacent positions are switched. As Perea and Lupker (2003) have recently argued, the influence of transposed letter stimuli is inconsistent with most available models of word recognition, because these models typically code letters by positions within the words. These results are more consistent with recent input coding models such as SOLAR (Davis, 1999), and SERIOL (Whitney, 2001) that use spatial coding schemes for input of letters, that are not simply position specific (also see Davis & Bowers, 2004). Clearly, this is an important new area of research that extends the original work on orthographic N effects and has important ramifications of how the visual system codes the spatial position of the letters within words.

4.6. Phonological Neighborhood Effects

Although the influence of orthographic neighbors has dominated work in visual word recognition, it is quite possible that phonological neighbors may also play a role. Indeed, work by Yates, Locker, and Simpson (2004) has recently shown that lexical decision performance is facilitated by words with large phonological neighborhoods (also see Yates, in press). Here, a phonological neighbor reflects a change in one phoneme, e.g., GATE has the neighbors HATE and GET, and BAIT. Yates et al. have also noted that previous studies of orthographic neighborhood size have typically confounded phonological neighborhood size. Although this is a relatively new area of exploration, it indeed is quite intriguing regarding the role of phonology in early access processes (see earlier discussion of feedback consistency effects), and has potentially important implications for how phonology is coded in the extant models (also see Ziegler & Perry, 1998).

5. 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 (see review by Balota, Ferraro, & Connor, 1991, of the early work in this area). This is an intriguing possibility because many models of word recognition would appear to indicate that the word must be recognized before the meaning of the word is determined. For example, within a logogen model, the lexical representation will need to reach threshold before the meaning of the word becomes available. How could it be otherwise? How could the system have access to the meaning without knowing what the stimulus is? Of course, this has some similarity to the word superiority effect described earlier wherein it was argued that the word level information is activated before the letters that make up the word have been recognized, via cascaded top-down activation. In fact, recent computational models by Coltheart et al. (2001) and Plaut et al. (1996) would appear to be able to handle such cascaded influences of meaning en route to making a speeded naming response.

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 shall briefly review this work emphasizing the primary findings with respect to each of the major variables.

5.1. Concreteness/Imageability Effects

Because concreteness is highly correlated with imageability, we will lump these variables together here. Concreteness refers to whether a word can be the object of a sense verb (e.g., touch, see, hear, etc.), whereas imagability typically involves subjects rating words on a low to high imageability scale. One might expect that high-imageable words (e.g., CARROT) may be better recognized than less imageable words (e.g., FAITHFUL), because of the additional influence of the more salient referent being activated. Although the early evidence suggested that there were indeed effects of the concreteness/imageability variables (e.g., Boles, 1983; Day, 1977; Paivio & O'Neill, 1970; Rubenstein et al., 1970; Winnick & Kressel, 1965), some of this work was questioned because of the potential for confounding variables (see, for example, Schwanenflugel, Harnishfeger, & Stowe, 1988). However, there are indeed studies that are less susceptible to such criticism and have confirmed that there are concreteness/imageability effects in lexical decision, which are larger for low-frequency words than high-frequency words (e.g., De Groot, 1989; James, 1975; Kroll & Merves, 1986). Of course, this finding in and of itself is not terribly compelling for an influence of meaning en route to word recognition performance, because one could argue that subjects place a premium on semantics in discriminating words from nonwords in the lexical decision task. Hence, the results from the naming task are indeed more noteworthy. Although the effects are clearly smaller, there is also evidence of an effect of concreteness/imageability in naming (e.g., Bleasdale, 1987). In the Balota et al. (2004) study of over 2000 monosyllabic words, they found that there was a reliable unique effect of imageability in naming (based on norms developed by Cortese & Fugett, 2004) after other variables were controlled, but this effect was quite small compared to lexical decision.

AQ11

Imageability has played a special role in recent empirical work in naming. Specifically, Strain et al. (1995) found an intriguing interaction between word frequency, spelling-to-sound consistency, and imageability. They found that low-frequency words with inconsistent spelling to sound consistencies produced the largest imageability effects. This was viewed as reflecting greater input from preexisting semantic representations for items with relatively low spelling to sound mapping, i.e., low-frequency inconsistent words, which they viewed as consistent with the tripartite connectionist framework, as exemplified by Plaut et al. (1996) model. It should also be noted, however, that there is some controversy regarding potential correlated variables that may have contributed to this effect (see Monaghan & Ellis, 2002; Strain, Patterson, & Seidenberg, 2002).

5.2. Meaningfulness

A second semantic variable that could play a role in word recognition is the meaningfulness of the stimulus. One way of measuring meaningfulness is simply to count the number of dictionary meanings for each word. Again, the early work in this area was controversial (for further discussion of different metrics of meaningfulness, see Millis & Button, 1989). For example, Jastrzembski (1981) found initial evidence for a facilitatory effect of number of dictionary meanings, while, Gernsbacher (1984a, 1984b) argued that this was likely due to familiarity being confounded with meaningfulness. Azuma and Van Orden (1997) found an effect of number of meanings in lexical decision performance, but this seemed to depend on the relatedness of the meanings for a word. In fact, Azuma and Van Orden argued that the relatedness of the meanings is more important than the sheer number of meanings. As described below, this may be related to more recent notions of semantic connectivity. Balota et al. (2004) found a small effect unique effect of subject-rated meaningfulness that was larger in lexical decision than in naming performance. Finally, it is noteworthy that Rodd (2004) has recently provided evidence that the effect of number of meanings in speeded naming is larger for inconsistent spelling to sound mappings. This, of course, is consistent with the theoretically important observation of an increased influence of a semantic variable (imageability) for low-frequency inconsistent items, reported by Strain et al. (1995) described above.

As noted, meaningfulness is typically defined by the number of dictionary meanings, which can vary in subtle but related ways. For example, the word DOG can mean the four-legged animal, but it can serve as an adjective such as in “My car is a real dog,” wherein the meaning of the word DOG is extended to another form. These might be considered different shades of the same meaning as opposed to distinct meanings of the word. In this light, there has also been some intriguing work investigating word recognition performance on homographs (e.g., the word ORGAN has two very different meanings referring to musical meaning and bodily system). It appears that such items can produce a facilitatory effect in both naming and lexical decision performance (see Hino & Lupker, 1996; Hino, Lupker, & Pexman, 2002). Interestingly, although one finds facilitation in naming and lexical decision, Hino et al. (2002) found inhibition in a semantic categorization. These authors argued that only when attention is directed to retrieve semantic information, as in the semantic categorization task, will one find interference effects. A similar pattern was observed by Balota and Paul (1996) in a semantic relatedness judgment task. Finally, in neutral contexts, on-line measures of reading performance, as reflected by eye-fixation durations, suggest that there is interference when ambiguous words have relatively equally dominant interpretations (e.g., CLUB means to hit and organization, with similar frequencies), as if the meanings are competing for interpretation (for a review, see Duffy, Morris, & Rayner, 1988; Morris, this volume). Again, we find that task constraints strongly modulate the influence of a variable.

5.3. Grounding Semantics in Large-Scale Databases

There have been a number of recent attempts to ground semantics via analyses of large databases of natural language. This approach avoids some of the pitfalls in trying to quantify meaning as feature lists (e.g., the word DOG may include the features furry, barks, four-legged, pet) or some abstracted prototype (e.g., the model DOG that is based on your experience with all DOGs). These more recent approaches include Burgess and Livesay’s (1998) hyperspace analogue of language (HAL) and Landauer and Dumais’ (1997) latent semantic analysis (LSA). HAL and LSA capture the meaning of words from the context in which a given word occurs with. Hence, the meaning of DOG is an evolving concept dependent upon an individual’s experience with DOG in various linguistic contexts.

Buchanan, Westbury, and Burgess (2001) have shown that estimates from HAL indeed predict lexical decision performance (for a detailed discussion of this work, see Burgess, this volume). It is noteworthy that an early study by Schwanenflugel, Harnishfeger, & Stowe (1988) provided evidence that a variable referred to as contextual availability can have an influence on isolated word recognition in lexical decision performance above-and-beyond influences of correlated variables such as concreteness, familiarity, length, etc. Contextual availability refers to how easily a subject is able to think of contexts in which a given word might occur.

An intriguing alternative approach has recently been developed by Steyvers and Tenenbaum (2005). They have utilized recently developed graph theoretic techniques to look at metrics of connectivity (along with other metrics) of meanings of words in a set of large-scale databases including, Roget's (1911) Thesaurus, Miller's (1990) WordNet, and Nelson, McEvoy, and Schreiber's (1998) word association norms. Based on these analyses of these databases, Steyvers and Tenenbaum have shown that semantic memory has a small-scale network structure in which a relatively small number of concepts serve as communication hubs for the rest of the semantic network. If semantic networks are represented in terms of the structure hypothesized by Steyvers and Tenenbaum, then words characterized by a high degree of connectivity with other words may be processed more quickly than words characterized by sparse connections. Indeed, Steyvers and Tenenbaum found evidence for such an effect in naming and lexical decision performance, above and beyond more standard lexical variables (also see Balota et al., 2004).

5.4. Additional Semantic Variables that Produce Effects in Isolated Word Recognition Paradigms

Because of space limitations, we shall only briefly mention a few other findings that would appear to indicate that meaning can have an early influence 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, etc. 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) indicated 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 (reviewed below) suggesting that highly masked primes (that subjects apparently cannot consciously recognize) produce semantic priming effects, i.e., facilitate the processing of related targets compared to unrelated targets (see Holender, 1986, and the accompanying commentary for a 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.

5.5. 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 relatively little direct influence of meaning-level variables. 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, i.e., recognition of the word. Although recent connectionist and DRC models of word processing acknowledge such effects, the devil is in the details of implementing such meaning-level influences.

6. CONTEXT/PRIMING EFFECTS

Heretofore, we 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-letter strings 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), semantically related (FEEL-TOUCH), or unrelated (NAIL-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 rich literature, see Neely (1991), Hutchison (2004), McNamara (2005) for important reviews of the semantic priming literature and Kinoshita and Lupker (2003) for a volume dedicated to masked priming effects.

6.1. 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 presented two letter strings that are both preceded and followed by pattern masks. The two letter 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 these masked priming studies: first, on most trials, subjects are unable to consciously identify the prime items and hence any influence of the prime items presumably reflects early access processes. Second, subjects are better at identifying the second-letter string when it shares letters with the first-letter string even though these shared letters are presented in different case. For example, relative to a baseline (e.g., harmless-ATTITUDE), there are priming effects for both *identity priming* (e.g., attitude-ATTITUDE) and *form priming* (e.g., aptitude-ATTITUDE). Third, in lexical decision, evidence for nonword repetition priming (e.g., flirp-FLIRP) is clearly less powerful than word repetition priming (Forster, 1998). Although earlier studies actually failed to find nonword repetition priming effects in lexical decision (see, for example, Forster and Davis, 1984), more recent studies have observed reliable effects (Bodner & Masson, 1997; Sereno, 1991). Fourth, in masked repetition priming studies, the effects of target word frequency and prime-target repetition are additive (Forster & Davis, 1984), a finding which is more consistent with search-class than with activation-class models of lexical access. Fifth, eye-tracking studies by Rayner, McConkie, and Zola (1980) using orthographic priming techniques have provided compelling evidence for a case independent orthographic code being used to access words in the parafovea while reading (for reviews, see Balota & Rayner, 1991; Rayner, 1998).

A particularly intriguing aspect of the masked priming literature is that within a range of short-duration primes, there is a relatively linear relationship between the *duration* of the masked prime and the *magnitude* of the priming effect (Forster & Davis, 1984). Specifically, a prime with a duration of 30ms produces a priming effect of about 30ms, whereas a prime with a duration of 20ms produces a priming effect of about 20ms. Forster (1998) has argued that this is most consistent with an *Entry Opening* process where the prime has the influence of opening the target's lexical representation, allowing the target to be processed more rapidly. This Entry Opening account of masked priming nicely accommodates the equivalent masked repetition effects for high-frequency and low-frequency words, i.e., the masked prime has the effect of opening the lexical representation (Forster & Davis, 1984). However, it is unclear how the Entry Opening model accounts for nonword repetition priming effects, since nonwords, by definition, have no pre-existing lexical representations. To address such nonword effects, Bodner and Masson (1997) have proposed that masked priming effects are driven by a nonlexical locus, specifically, the retrieval of episodic memory traces established during previous encounters with the stimulus (for an episodic trace view of lexical processing, see Goldinger, 1998). This account implies that masked nonword primes operate nonlexically to facilitate orthographic processing (for an alternative explanation, but see Forster, 1998).

Finally, task-specific effects have also been observed in masked priming. For example, there is evidence for a phenomenon called the *masked onset priming effect.* This effect was first reported by Forster and Davis (1991), who found that naming latencies to a target were facilitated when the prime and target shared the initial letter (e.g., save-SINK) compared to when they did not (e.g., farm-SINK). Further work by Kinoshita (2000) has revealed that this effect is position-dependent and is observed *only* when the initial onset (not the letter) is shared. For the pair bingo-BLISS, which has a common initial letter but different onsets (i.e., /B/ vs. /BL/), the effect was eliminated. Kinoshita argued that this supported a serial left-to-right procedure in naming performance, and may reflect articulatory planning rather than orthography-to-phonology computations (see also Schiller, 2004). The onset effect is only observed with tasks that require articulation, such as speeded naming, and not with lexical decision (Forster & Davis, 1991). Positing an articulatory nonlexical priming component for speeded naming may also explain why nonword repetition priming effects, which are equivocal in lexical decision, are more consistent in speeded naming (Masson & Isaak, 1999).

6.2. Phonological Priming Studies

There has been considerable debate concerning the role of phonological codes in word recognition (for an excellent review of this literature, see Frost, 1998). The extremes range from all words must be recognized via a phonological (assembled) code to the notion that many words (e.g., high-frequency words for skilled readers) are only accessed via an orthographic (addressed) code. Although there is 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). It is possible that such phonological codes become active after lexical access has taken place in such reading studies. The more narrow issue here is whether phonological codes are necessary in the word recognition process. With this in mind, we now turn to the phonological priming literature.

Evett and Humphreys (1981) used the masked priming paradigm, described above, also to investigate the viability of a phonological access code, under conditions wherein conscious processing was limited. 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), suggesting a lexical locus for the priming effect.

Evidence for phonological mediation has also been obtained with an associative priming paradigm, which permits conscious, albeit brief, processing of primes. For example, Lukatela and Turvey (1994) compared priming effects across four conditions at different stimulus onset asynchronies (SOAs): *standard semantic* priming (e.g., TOAD-FROG), *word homophonic* priming (e.g., TOWED-FROG), *nonword homophonic* priming (e.g., TODE-FROG), and an orthographic control condition (e.g., TOLD-FROG). At short (i.e., 50ms) SOAs, the three related conditions produced comparable facilitation priming effects, compared to the control condition. However, at longer SOAs (i.e., 250ms), TODE became a stronger prime than TOWED. These findings reinforce the role of phonology in early visual lexical access, and also suggest that although word homophone primes (i.e., TOWED) are initially effective, they are quickly suppressed when the system detects the mismatch between their orthography and the addressed spelling of TOAD.

It is important to point out that the validity of the findings described above rests on the assumption that the orthographic control (e.g., TOLD) is as orthographically similar to the target (e.g., FROG) as the homophone (e.g., TOWED) (Pollatsek, Perea, & Carreiras, 2005). Some have failed to replicate the homophone/pseudo-homophone advantage described above (see, for example, Davis, Castles, & Iakovidis, 1998) and Pollatsek et al. argued that this inconsistency may be due to imperfect matching of controls to homophones. After controlling for this potential confound, Pollatsek et al. still observed early phonological effects in a Spanish lexical decision task, strengthening the assertion that phonological coding of the primes takes place relatively early in the word recognition process.

Interestingly, the importance of phonological codes in word identification has been demonstrated in both orthographically shallow languages, where there is a direct mapping between orthography and pronunciation (e.g., Serbo-Croatian, for a review, see Carello, Turvey, & Lukatela, 1992) and orthographically deep languages, where the mapping appears to be more arbitrary (e.g., Chinese, for a review, see Tan & Perfetti, 1998). Clearly, phonological information can constrain visual word recognition even in logographic scripts where one would expect meaning to be derived directly from ideograms (see Hoosain, 1991). For example, Tan and Perfetti (1999) sequentially presented pairs of Chinese words in a meaning-judgment task, in which they were asked to judge whether the two words had the same meaning or not. On trials where participants were supposed to make a “no” judgment (i.e., the two words had different meanings), the “no” response had longer latencies when the foil was homophonous with the base word compared to when it was not.

There have been additional tasks used to investigate the early influence of phonological processes. For example, Van Orden (1987; Van Orden, Johnston, & Hale, 1988) used 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 rates 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 suggests a clear role of phonological information in accessing the semantics necessary for category verifications, and nicely converges with the results from the Tan and Perfetti (1999) study with Chinese characters. Jared and Seidenberg (1991) replicated this pattern showing that this effect is more likely to occur for low-frequency words. This pattern also 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 (also, see Rodd, 2004). Ziegler, Ferrand, Jacobs, Rey, and Grainger (2000) used an *incremental priming technique*, by manipulating the duration of the prime, which provides a window into the time-course of masked priming effects. They found clear orthographic and phonological priming effects in both naming and lexical decision performance, with the naming task being more dependent upon phonological priming. This study is particularly noteworthy because it provides a method to help understand the temporal locus of such priming effects. Finally, it is also worth noting that just as in the case of orthographic priming, there is also evidence of phonological priming in the parafoveal priming paradigm in more natural reading contexts. Specifically, Pollatsek, Lesch, Morris, and Rayner (1992) found that previews that were homophones of targets (e.g., site-cite) facilitated performance (both in pronunciation latencies and fixation durations during reading), compared to nonhomophonic previews that were controlled for orthographic similarity (e.g., cake-sake). Lee, Binder, Kim, Pollatsek, and Rayner (1999) have extended this work with the *fast-priming* paradigm (for a description, see Sereno & Rayner, 1992), a task which taps early stages of word processing. They observed an interesting prime by word frequency interaction; specifically, homophonic priming was primarily obtained with high-frequency word primes. Taken together, these findings not only support the role of phonology as an access code, but also suggest that lexical information may be guiding phonological coding early in a fixation during reading (Lee et al., 1999).

6.3. "Semantic" Priming Effects

The semantic (associative) priming paradigm is clearly the most studied area of priming. (Because of space limitations, the present section will be limited to single word priming studies, see Morris, this volume, for a review of sentential semantic priming effects.) 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 appeared to be a bottom up, stimulus driven word recognition processes. This was a major emphasis in Neisser's (1967) *Cognitive Psychology* that was 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 upon such representations; both of these issues would at least appear to be critical to higher-level linguistic performance.

6.3.1. Semantic or associative effects?

There is little controversy that across the major tasks used to build word recognition models (threshold identification, lexical decision, pronunciation, and on-line measures of eye-movements during reading), words are better recognized when embedded in semantically related contexts compared to unrelated contexts. However, there are many questions that have arisen regarding this effect. For example, one might ask if the effect is truly "semantic" (i.e., reflects similarity in semantic features, Smith, Shoben, & Rips, 1974 or category membership, Collins & Quillian, 1969), or if it primarily reflects associative relationships among items. For example, DOG and CAT share a semantic and associative relationship, whereas RAT and CHEESE appear to primarily share an associative relationship. Two recent reviews of this topic appear to come to somewhat different conclusions. Lucas (2000) argued that there was indeed evidence that semantic priming effects truly reflected "semantic" information, whereas, Hutchison (2003) concluded that, with a few exceptions, a simple associative account could handle most of this literature. Of course, teasing apart semantic influences from associative influences has been rather difficult because these relationships typically co-occur. In an attempt to

AQ12

address this issue, researchers have attempted to identify 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 (see, for example, Palermo & Jenkins, 1964). The results from three such studies (e.g., Lupker, 1984; Schreuder, Flores d'Arcais, & Glazenborg, 1984; Seidenberg, Waters, Sanders, & Langer, 1984b) 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.

One must be 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, but "pure," semantic priming effects might be due to some lingering associative-level relationship for words that researchers believe 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. As noted below, 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 post-lexical 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 upon the ratio of nonwords to words. Neely et al. argue that the nonword/word ratio should modulate the likelihood of the checking process being engaged in the lexical decision task. Hence, because of the task-specific list context effect in this study (i.e., the effect of the nonword/word ratio), one may question the argument for a pure semantic priming effect in access processes (also see Balota & Paul, 1996). In the following discussion, we will use the term "semantic" priming effects, however, the reader by now should understand that many of these effects could be primarily "associative" in nature.

6.3.2. 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. (Although, as described below, it is 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 lexical decision performance in which subjects only respond to the target string, there is little evidence for mediated priming (cf. Balota & Lorch, 1986; De Groot, 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 respond to nonword targets (Den Heyer, Sullivan, & McPherson, 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.

AQ13

6.3.3. Threshold priming effects

A second important finding in this literature deals with threshold semantic priming effects, mentioned earlier. In the initial studies in this area, 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 subject's threshold and the target is presented in a lexical decision task. The intriguing finding here is that there still is 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, Carr, & Wilhelmsen, 1989; Fowler, Wolford, Slade, & Tassinary, 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, 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. It is noteworthy that the threshold priming literature has also been extended to functional neuroimaging techniques. For example, in an event-related potential/functional magnetic resonance neuroimaging study, Dehaene et al. (1998) used number primes that were so briefly presented that participants were unable to discriminate them from foils. Nevertheless, these primes influenced performance on a semantic comparison task (press one key if the target is less than 5 and another key if the target is greater than 5), and modulated hemodynamic measures of brain activity. As in the mediated priming studies, these studies indicate that conscious access to a prime–target relationship does not appear to be a necessary condition for obtaining semantic priming effects.

In some studies of threshold priming, stimuli and/or targets are repeated across trials, with thresholds being carefully monitored. There has been a recent debate about whether such effects in these paradigms reflect unconscious access to meaning at the whole-word level (Abrams & Greenwald, 2000; Damian, 2001; Naccache & Dehaene, 2001). For example, Abrams and Greenwald (2000) argued that threshold priming effects in these studies may reflect automatized stimulus–response mappings that develop as participants make responses to visible targets across trials (for an alternative view, Damian, 2001, but see Kunde, Kiesel, & Hoffmann, 2003). Specifically, after participants repeatedly (and consciously) classify *smut* and *bile* as negative words, *smile* (*smut-bile* hybrid) subsequently functions as a negative valence masked prime (Abrams & Greenwald, 2000). No significant priming is found for masked primes that had *not* earlier appeared as a target to be classified. These findings question the traditional premise that threshold primes are analyzed at the whole-word level, and suggest that subconscious processing may instead involve sublexical analyses. There is, however, some recent evidence that these findings are specific to words. With numbers, masked primes are able to provide access to long-term semantic memory (Greenwald, Abrams, Naccache, & Dehaene, 2003). The important point here is that one needs to be cautious in interpreting "threshold" priming effects when stimuli are repeated across trials.

AQ14

It is also noteworthy that the masked priming paradigms have been extended to the domain of social psychology. The overarching question of interest is whether affective states can be automatically triggered by threshold-level primes. For example, Fazio, Sanbonmatsu, Powell, and Kardes (1986) found evidence for automatic attitude activation using an adjective connotation task (i.e., rate a target as "good" or "bad"). They observed that participants rated a negative valenced target (e.g., DISGUSTING) more quickly when it was preceded by a negative prime (e.g., COCKROACH). Wittenbrink, Judd, and Park (1977) found similar priming effects under highly masked prime conditions. In a highly cited paper by Devine (1989) using the Neely (1977) automatic and controlled distinction in semantic priming, there was clear evidence of automatic activation of racial prejudice at short SOAs that was ultimately controlled at longer SOAs (see also Payne, 2001; Lambert et al., 2003). In reviewing this literature, Fazio (2001) has argued that such attitude priming is automatic and unconscious (also see De Houwer, Hermans, & Eelen, 1998).

Automatic influences from masked primes have also been detected using more ecologically valid paradigms. Bargh and Chartrand (1999) provide a comprehensive review of this literature. For example, Bargh, Chen, and Burrows (1996) found that participants presented with highly masked primes that presumably activated "rudeness" traits (e.g., *rude*, *impolite*, *obnoxious*) were more likely to interrupt a subsequent conversation than if they were primed with "politeness" traits (e.g., *respect*, *considerate*, *polite*). Collectively, the evidence from attitude and affect priming in social psychology is in-line with the evidence from semantic masked priming in visual word recognition. Given the cascadic nature of the models that we discussed earlier, such a pattern might be expected. However, this literature also clearly demonstrates that one needs to be cautious and use converging evidence to evaluate whether such effects are in the purest sense unconscious.

6.3.4. Backward priming effects

The third area that is counterintuitive 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 vs. 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 to again fall from a cascadic framework in which partial activation is released from representations before such representations have reached threshold.

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 et al. (1984b) 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 et al., 1984b), except under short stimulus onset asynchronies (see Kahan, Neely, & Forsythe, 1999; Peterson & Simpson, 1989). It is possible that at short SOAs, there is sufficient temporal overlap between the target and the context to produce the first type of backward priming, noted above, even in naming.

6.4. Syntactic Priming

If associative/semantic context does indeed influence lexical processing, then it is quite possible that syntactically appropriate vs. 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 (c.f., Ratcliff & McKoon, 1988). Likewise, one might argue that syntactic tags within lexical representations might produce priming to consistent syntactic representations. Alternatively, 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 et al. (1984b) 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 post-lexical processing of the relation between the prime and target. At first, it appeared that Seidenberg et al.’s 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. However, Sereno (1991) argued that the past syntactic priming studies have used relatively long prime–target SOAs, and hence may be due to attentional expectancies. In a series of careful studies, with highly masked primes, Sereno found clear syntactic priming effects in lexical decision that were eliminated in naming, consistent with the Seidenberg et al.'s original arguments about task-specific post-lexical checking processes.

6.5. 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 and naming tasks, but its importance extends to the intriguing interactions that have been uncovered. As an example, consider the following intriguing pattern of interactive effects: (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 non-degraded words (Becker & Killion, 1977; Borowsky & Besner, 1991); (c) there are additive effects of stimulus degradation and word frequency (see Balota & Abrams, 1995; Becker & Killion, 1977; Borowsky & Besner, 1991). Traditionally, this constellation of findings has been used to support independent, sequentially organized stages in lexical processing (Borowsky & Besner, 1993; Plourde & Besner, 1997; Sternberg, 1969). In contrast, Plaut and Booth (2000) have argued that a single-mechanism PDP model, implemented with a sigmoid activation function, can more parsimoniously simulate these effects, along with additional findings in the literature. This debate has recently resurfaced, with Borowsky and Besner (2005) contending that there is insufficient evidence that the PDP model implemented by Plaut and Booth (2000) can simultaneously achieve high lexical decision accuracy and correctly simulate the joint effects of stimulus quality, word frequency, and priming in speeded lexical decision. Instead, they argue that the available evidence is more consistent with serially organized processing stages that are differentially sensitive to degradation, semantic relatedness, and word frequency. Evidence for independent stages of processing is especially intriguing when considering the human word recognition architecture.

6.6. Theoretical Accounts of Semantic Priming Effects

The importance of the semantic priming paradigm has not simply been restricted to models of word recognition, but also has extended to more general issues concerning representation and retrieval processes. We shall now briefly discuss some of the theoretical issues that have been nurtured by this literature, but the interested reader should see Neely (1991), Hutchison (2004), and McNamara (2005) for a full discussion of these theoretical mechanisms.

6.6.1. 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. When a node in memory becomes activated via stimulus presentation or via internal direction of attention, the notion is that 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 because the underlying representation 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 Snyder (1975) argued that the automatic spreading activation mechanism was (a) fast-acting, (b) 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, for example, Balota & Duchek, 1989; De Groot, Thomassen, & Hudson, 1982; Jonides & Mack, 1984; Neely, 1991), we will primarily focus on Posner and Snyder's first two characteristics.

There are a number of important semantic priming results that would appear to support Posner and Snyder's automatic spreading activation mechanism. First, the evidence for semantic priming under highly masked priming conditions, reviewed above, is consistent with the notion that priming effects are independent of consciously controlled processing (e.g., Balota, 1983; Dehaene et al., 1998; Fowler et al., 1981; Marcel, 1983). Second, the evidence that there are mediated priming effects at relatively short prime–target SOAs (e.g., from LION to STRIPES), when it is unlikely that subjects have sufficient time to 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, Briand, & Dannenbring, 1983), but strong influences at long SOAs, 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. AQ15

Although there appears to be support for something akin to an automatic spreading activation mechanism, there are some caveats. For example, initially, there was little evidence of priming effects occurring across unrelated word (e.g., facilitation from LION to TIGER in LION-CHALK-TIGER compared to FROG-CHALK-TIGER, e.g., Gough, Alford, & Holley-Wilcox, 1981; Masson, 1991; Ratcliff & McKoon, 1988). Clearly, if the effect is automatic, one would expect such effects. In this light, it is noteworthy that more recent studies by Joordens and Besner (1992), McNamara (1992), and Balota and

Paul (1996) 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. A second potential problem with the automatic nature of spreading activation is that semantic priming effects can be eliminated when subjects process the primes in a very shallow fashion, e.g., responding to whether a given letter is in the prime or an asterisk is beside the prime (e.g., Henik, Friedrich, & Kellogg, 1983; Smith, 1979; 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 (for further discussion of this issue, see Besner, Smith, & MacLeod, 1990). Finally, Balota, Black, and Cheney (1992) have shown that prime-expectancy instructions (e.g., subjects are instructed to expect exemplars from the TREE category when presented the prime METALS) can influence naming performance even at very short prime–target SOAs. Thus, although there is 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 (for a recent discussion of the automatic nature of spreading activation, see Neely & Kahan, 2001).

6.6.2. 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 when given the prime. The more intriguing work here is the specification of the parameters that modulate the expectancy effects, i.e., 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 (for a detailed discussion of a semantic expectancy model, see Becker, 1980, 1985).

6.6.3. 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 post-lexical decision level (e.g., Balota & Lorch, 1986; De Groot, 1984; Forster, 1979, 1981; Neely, 1976, 1977; Neely & Keefe, 1989; Seidenberg et al., 1984b; 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 backwards 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 to integrate the meaning of the current word with the ongoing comprehension of the previous words (for a full discussion of the backward checking mechanism, see Neely & Keefe, 1989). In support of this possibility, Kahan et al. (1999) found some evidence of backward checking at short SOAs even in naming performance.

6.6.4. Compound-cue model

Ratcliff and McKoon (1988) 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 “yes” decisions and low-familiar compound cues produce relatively slow “no” 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 than unrelated prime–target pairs.

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, along with other tasks. Neely’s (1991) 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 importantly, 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 this mapping process.

6.6.5. Plaut and Booth's (2000) Single-mechanism connectionist model

In contrast to Neely's tripartite framework described above, Plaut and Booth have claimed that a distributed network model can account for semantic priming lexical decision phenomena using a *single* mechanism. Implementing a distributed attractor network simulation with distributed orthographic and semantic representations (Plaut, 1995), Plaut and Booth were able to account for a number of theoretically interesting findings, including the surprising observation that only participants with high perceptual ability exhibited the priming by frequency interaction (i.e., greater priming for low-frequency words); participants with low perceptual ability showed equal priming for both high- and low-frequency targets. Like the Seidenberg and McClelland (1989) model, however, the connectionist view of priming faces challenges. For example, as mentioned earlier, there is an ongoing debate about whether semantic priming is better accommodated by a single-mechanism account or by separate mechanisms that invoke distinct sets of computational principles (see Borowsky & Besner, 2005). Nevertheless, this work represents an interesting advance in that it includes a computationally implemented architecture that has been applied across a number of cognitive domains and accommodates some intriguing data in the priming literature and takes a step toward tackling the important topic of individual differences.

6.6.6. Masson's (1995) distributed memory model of semantic priming

Masson's model, based also on distributed connectionist principles, provides a framework for accommodating semantic priming in speeded naming that neither appeals to spreading activation nor compound cues. In this model, conceptual knowledge is represented via distributed orthographic, phonological, and semantic units that are connected by weighted pathways. Importantly, Masson's network, a Hopfield (Hopfield, 1982) net variant, does not distinguish between input, hidden, and output units. The basic principle in the model is that semantically related words have very similar patterns of activation in the semantic units. When a semantically related prime is presented, activation in the semantic units starts moving toward a pattern that is similar to the pattern of activation of the to-be-presented target. When the target appears, the overlap between its pattern and the pattern of activation in the semantic units helps the phonological units converge more rapidly on the target's pattern, and hence, speeds naming responses. This model is able to account for the *intervening stimulus effect*, which, as mentioned above, is the observation that interpolating an unrelated word between the prime and the target reduces the priming effect in naming performance, a finding that the spreading activation framework does not readily predict. However, it is also the case that this model has not been extended to the wealth of data that Neely's tripartite framework appears to be able to handle.

6.7. Summary of Context/Priming Effects

The priming literature has provided an extremely rich data base to develop models of context effects, memory retrieval, and word recognition. Because of space limitations, we were 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, and Forster's (1976) bin model. Each of these models provides intriguing alternative perspectives on semantic priming effects. At this point in theory development, it appears that no single model of priming 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 semantic priming effects.

7. ATTENTIONAL CONTROL, MODULARITY, AND TIME CRITERION MODELS

The models reviewed earlier appear to have a relatively passive architecture wherein different systems accumulate information across time. However, in some instances, it may be advantageous for the reader to modulate the contribution of a given pathway depending upon the task demands or particular reading context. For example, one might expect different emphases on distinct systems when proofreading, comprehending, or checking for grammaticality. Virtually, every theory of word recognition posits multiple ways of accessing or computing the phonological code from print. In the DRC model, one can compute a phonological code via the lexical route, which maps the whole word onto a lexical representation to access phonology, or via the sublexical route, which computes the phonology via the spelling-to-sound correspondences in the language; in PDP models, the phonology can be computed by differential emphasis on the direct orthographic to phonological connections or the indirect connections via semantics. The question naturally arises whether there is any control of which processing pathway influences performance in a given task. This is important because it brings into question the modularity of the lexical processing system (see Fodor, 1983).

One way to examine the control issue is to present words that place different demands on the lexical and sublexical information. For example, within a DRC model, nonwords should bias the sublexical pathway. However, low-frequency exception words should bias the lexical pathway, since the sublexical pathway would lead to regularization errors for low-frequency exception words, i.e., pronouncing *pint* such that it rhymes with *hint*. Monsell et al. (1992) found that naming latencies to high-frequency irregular words were faster and more accurate when embedded with other irregular words, than when mixed with nonwords. Monsell et al. suggested that exception word context directed attention to the lexical pathway, which is more appropriate for naming exception words, than the sublexical pathway. Additional studies have found similar influences of pathway priming (e.g., Rastle & Coltheart, 1999; Reynolds & Besner, 2005b; Simpson & Kang, 1994; Zevin & Balota, 2000).

Although intuitively appealing, the evidence for route priming has been quite controversial. Specifically, work by Kinoshita and Lupker (2002, 2003a) suggests that much of the earlier findings can be accounted for by a time criterion model. The time criterion perspective is important in a number of domains so we will briefly review it here.

Specifically, there is evidence that participants adopt a time criterion whereby they are likely to produce a response at a latency that is biased toward the average of the latencies in a block of trials. Consider the word-frequency effect (presumably a reflection of the lexical route). In two pure independent blocks, assume that a set of low-frequency words produces response latencies on the average of 700ms and a set of high-frequency words produces response latencies on the average of 600ms. If one now embeds these same words in the context of nonwords that produce an average response latency of 700ms, the word-frequency effect will likely diminish. That is, latencies to the low-frequency words will remain the same (because the latencies for both low-frequency words and nonwords are 700ms), whereas latencies to the high-frequency words will increase considerably, i.e., migrate toward the time criterion invoked by mean latency of the nonwords. Hence, the word-frequency effect will decrease in the context of nonwords not because of a decreased reliance on the lexical pathway, but rather because of a change in the temporal criterion to produce a response.

The evidence clearly suggests that participants do adopt a time criterion based on the difficulty of items within a block. However, there is also evidence that appears to be consistent with a pathway control perspective above and beyond the time criterion effects. For example, all of the effects reported by Zevin and Balota (2000) hold even after the response latencies to the context items are partialed out via analyses of co-variance. Of course, if the time criterion model were the only responsible variable in this study, one should not find this pattern. Moreover, Kinoshita, Lupker, and Rastle (2004) have recently provided evidence that one can indeed modulate the lexicality effect (words faster than nonwords) via list context manipulations. However, they were unable to modulate the regularity effect (regular words faster and/or more accurate than exception words) by list context manipulations. In addition, Reynolds and Besner (2005b) have recently demonstrated that one can find lexical and sublexical pathway switching above and beyond any response latency criterion effects. Although there is accumulating evidence for some level of pathway control, further work is clearly necessary in this area. Indeed, the extent to which attentional systems modulate the information in distinct pathways has important implications for future modeling endeavors, and quite naturally would accommodate task-specific influences that have been emphasized in the present chapter. Moreover, time criterion perspectives are important in understanding how the word recognition system adjusts to the local constraints of an experiment and may have important implications for other cognitive paradigms that rely on response latency measures. At this level, time criterion effects may be viewed as an example of attentional control.

8. DEVELOPMENTS OF NEW APPROACHES AND ANALYTIC TOOLS TO GUIDE THE JOURNEY FROM FEATURES TO MEANING

In the following sections, we will describe some recent developments in approaches to studying word recognition. Again, this is not a comprehensive review, but simply a brief summary to expose the reader to some of the interesting techniques that are helping researchers constrain how humans process visual words.

8.1. Neuroimaging Techniques

In the past decade, tremendous advances in neuroimaging methodology have provided another window into the dynamics of lexical processing (also see Just and Mason, this volume). Specifically, neuroimaging techniques like positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and event-related potentials (ERPs) now allow researchers to localize and measure the time course of activity of brain regions that are recruited by a particular cognitive task (Fiez & Petersen, 1998). One particularly exciting development has been the advent of *event-related* fMRI designs (Dale & Buckner, 1997). In PET studies and early fMRI studies, *blocked designs* (i.e., experimental conditions are blocked) were mandatory, making paradigms like the lexical decision task impractical. Event-related fMRI allows researchers to extract the fMRI BOLD (blood oxygen level dependent) response for specific trials and to conduct standard word recognition experiments in the scanner.

As Fiez and Petersen point out, neuroimaging allows one to make both coarse as well as refined fractionations of brain regions that are involved in reading. For example, Petersen, Fox, Posner, Mintun, and Raichle's (1988) seminal study elegantly demonstrated that remarkably distinct brain regions were activated by different levels of single-word processing. Specifically, at a relatively coarse level, PET scans revealed that compared to appropriate baseline conditions, occipital areas were active for passive viewing of words (orthography), temporal areas were active for reading words aloud (phonology), and frontal regions were active when participants generated verbs to nouns (semantics). By varying tasks demands and contrasting neural activation in reading aloud vs. a control condition, these researchers were able to identify broadly the functions of different regions.

More recently, research designs have been employed to make finer-grained differentiations of regions that support different reading operations. For example, in a PET study of speeded naming, Fiez, Balota, Raichle, and Petersen (1999) manipulated the following three variables: lexicality (word vs. pronounceable nonword), frequency (high vs. low), and spelling-to-sound consistency (consistent vs. inconsistent). As discussed earlier, these variables have been central in the developments of models of word recognition, and so it is useful to explore the underlying circuitry. Fiez et al. (1999) found a number of noteworthy effects. First, a left frontal region showed effects of consistency and lexicality, indicating that this area may be involved in orthographic-to-phonological transformation. Second, there was greater activation for low-frequency words in a left temporal region and the supplementary motor area, which implicate these regions in the access and storage of lexical-level information. Third, effects of consistency were found bilaterally in the primary motor cortex, suggesting that consistency may influence both recognition *and* motor production systems; this surprising constraint has yet to be considered by extant theories of word recognition (but see, Kawamoto, Kello, Jones, & Bame, 1998). Fourth, the left inferior frontal gyrus showed a pattern analogous to the behavioral frequency by regularity interaction discussed earlier. Just as naming latencies are particularly slow for low-frequency inconsistent words compared to low-frequency consistent words, high-frequency inconsistent words, and high-frequency consistent words (Seidenberg et al., 1984a), the left inferior frontal gyrus showed strong activation only for low-frequency inconsistent words. This study demonstrates how manipulating stimulus properties in a neuroimaging paradigm can be used to complement *and* extend theoretical accounts that have hitherto been based on behavioral data.

In the remainder of this section, we will review some recent neuroimaging studies and discuss how these studies contribute to our understand of word recognition. Obviously, due to space constraints, this review is selective. Also, rather than enumerating in minute detail which brain regions are activated by which task, we will be using more sweeping brushstrokes to describe the functional neuroanatomy of reading.

8.1.1. Is there convergence across studies?

A reasonable concern one may have regarding neuroimaging research is the extent to which findings generalize across laboratories and studies. Variability across studies may arise as a result of intersubject variability and slight differences in methodology, making it difficult to establish consistent regions of activation (Turkeltaub, Eden, Jones, & Zeffiro, 2002). A few articles have attempted to review results from multiple studies in order to answer this question. For example, Fiez and Petersen (1998) reviewed nine studies where participants read aloud single words, and found encouraging convergence between studies. Basically, they combined the data across the studies by merging foci from different experiments into a single figure (Figure 9).

Fiez and Petersen (1998) established that a set of areas are consistently active during word reading, including the supplementary motor area, the cerebellum, the anterior cingulate gyrus (BA 32), left-lateralized fusiform and lingual gyri (BA 18 and BA 37), the left inferior frontal gyrus (BA 44/45), bilateral activation in the anterior and posterior regions of the superior temporal gyrus (near BA 22), and dorsal and ventral portions of the post-central gyrus (near BA 4). Interestingly, a more sophisticated meta-analysis of 11 PET studies generated similar findings (also, see Price, 2004). A statistical map of

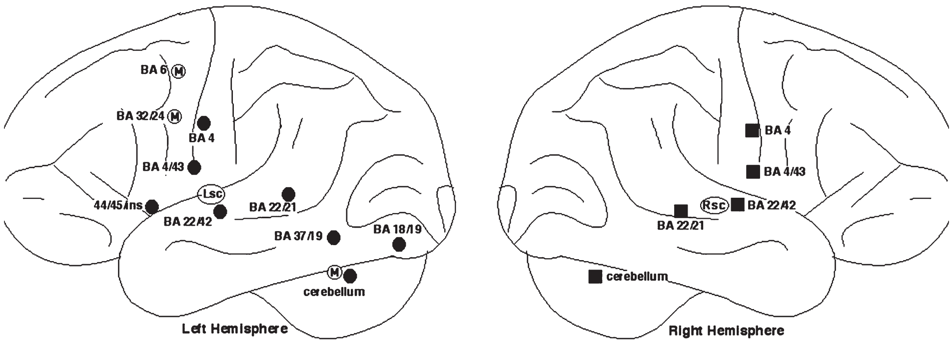

Figure 9. Brain regions that are consistently activated during word reading (Fiez & Petersen, 1998).

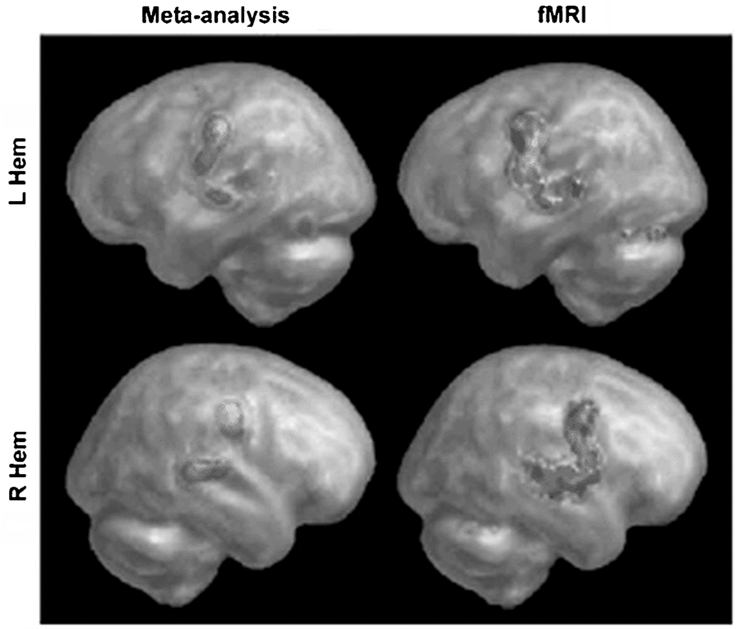

Figure 10. Statistical map generated by meta-analyses of 11 PET studies of word reading (Turkeltaub et al., 2002).

convergent foci by Turkeltaub et al. (2002) included the bilateral motor and superior temporal cortices, presupplementary motor area, left fusiform gyrus, and the cerebellum (see Figure 10). This map was successfully validated against new fMRI data of word reading, supporting the reliability of these findings.

Taken collectively, the regions identified by neuroimaging are broadly compatible with the neurological literature (Fiez & Petersen, 1998; but see Price, 2004, for exceptions). More importantly, these analyses generate candidate regions of interest that can be used by researchers to test new hypotheses.

8.1.2. Controversies regarding targeted areas

We have already discussed how the left inferior frontal gyrus is sensitive to spelling-to-sound consistency manipulations (Fiez et al., 1999), implicating this region in processes that transform orthographic to phonological representations. This pattern has been nicely replicated in a number of other studies (see Herbster, Mintun, Nebes, & Becker, 1997; Rumsey et al., 1997). Note that some researchers have proposed another locus for the sublexical procedure, the left posterior superior temporal region (Simos et al., 2002), although this region may be associated more with phonological decomposition (Palmer, Brown, Petersen, & Schlaggar, 2004) than with orthographic-to-phonological transformation *per se*.

Interestingly, there is far less agreement about the neural markers of lexical processing. Typically, frequency and lexicality effects have been used as indicants of lexical processing. Word-frequency effects may mark brain regions involved in the access and representation of either localist or distributed lexical-level information (Fiez et al., 1999). Although word frequency is easily the most studied variable in word recognition, it has received surprisingly little attention in the functional neuroimaging literature. The literature suggests greater activation for low-frequency words in a left temporal region (near BA 22) and the supplementary motor area (BA 6) in speeded naming (Fiez et al., 1999), and greater activation in the left inferior frontal gyrus (BA 44/45) and a number of subcortical structures in lexical decision[1] (Fiebach, Friederici, Muller, & von Cramon, 2002).

A number of studies have also used lexicality effects (greater activation for words than nonwords) as a marker for the lexical processing (Binder et al., 2003; Fiebach et al., 2002; Ischebaeck et al., 2004; Rissman, Eliassen, & Blumstein, 2003; Simos et al., 2002). It seems plausible that greater activations for words reflect access to linguistic information, which may either be orthographic, phonological, or semantic (Fiez et al., 1999). Unfortunately, there seems to be little consensus on regions that show greater activation for words. Lexical decision studies have identified diverse regions, including the left angular gyrus (BA 39), left dorsal prefrontal cortex (BA 6/8), superior frontal gyri (BA 6/8/9), left rostral–ventral cingulate gyrus (BA 32/24), left posterior cingulate gyrus and precuneus (BA 23/29-31/7), and the junction of the left posterior middle temporal and inferior temporal gyri (BA 21/37). Curiously enough, in speeded naming, it is relatively difficult to find regions that show greater activations for words than nonwords. For example, in the Fiez et al. (1999) PET study, no region showed greater activation in the word condition compared to the nonword condition (for a replication in an fMRI study, see Palmer, 2003) . This may be attributable to the lower spatial resolution of PET (compared to fMRI), or to strategic effects induced by blocking. Using a more sensitive event-related paradigm with Japanese Kana words, Ischebeck et al. (2004) found greater activation for words in the left and right temporo–parietal areas (BA 39/40), the middle part of the left middle/inferior temporal gyrus (BA 21/20) and the posterior cingulate (BA 31).

The marked discrepancy between the lexical decision and naming findings may be partly attributable to differential demands of the two tasks; a theme that has consistently arisen in the current chapter. If indeed there are large task differences in the behavior, there clearly should be consequences for the neural underpinnings. For example, the meta-analyses we discussed earlier (e.g., Fiez & Petersen, 1998; Turkeltaub et al., 2002) were based on neuroimaging studies of naming. It will be interesting to see the degree to which statistical maps based on meta-analyses of *lexical decision* data show a similar pattern. This is a theoretically important question that has yet to be answered.

---

[1] The latter finding needs to be qualified; the study was done with German stimuli, which have regular spelling-to-sound correspondences. The regularity of the orthography may have facilitated sublexical processing (reflected in left inferior frontal gyrus activation) and attenuated lexical activity.

Nevertheless, in spite of the diversity of findings, two regions seem to be consistently associated with lexicality effects (word > nonword), the left angular gyrus (BA 39) and the left middle temporal gyrus (BA 21), a pattern which is nicely consistent with the Kana-naming study (Ischebeck et al., 2004). The left middle temporal gyrus has long been associated with language processing (Fiebach et al., 2002); other studies also implicate the region in the representation and processing of lexico–semantic (Pugh et al., 1996) and phonological word forms (Cohen et al., 2000). The left angular gyrus also seems to play a role in non modality-specific semantic processing (Binder et al., 2003).

PET and fMRI are not the only windows into the functional neuroanatomy of reading. While these measures have excellent spatial resolution, the intrinsic characteristics of these signals limit their temporal resolution. Event-related potentials (ERPs) have far more exquisite temporal resolutions, and so these measures are better suited to study the time course of word recognition processes. Scalp-measured ERPs reflect the brain electrical activity that is triggered by experimental stimuli, and capture in real-time cognitive processes on a millisecond basis (Kutas & Federmeier, 2000; Kutas and van Patten, this volume). Frequency effects, for example, are apparent in ERPs between 200 and 400 ms, and are most obviously correlated with a left anterior negative component called the lexical processing negativity (LPN) (King & Kutas, 1998). In a recent representative ERP study, Hauk and Pulvermüller (2004) investigated how word length and word frequency influenced the amplitude and peak latencies of event-related potentials in lexical decision. Early effects of length and frequency were observed, and the researchers interpreted the results as consistent with lexical access occurring as early as 150ms after the onset of visually presented words. More intriguingly, large length effects were observed in ERPs but not in the behavioral data; this dissociation demonstrates that psychophysiological measures may in some cases be more sensitive than behavioral data.

To recapitulate, in the foregoing discussion, we have briefly considered the neuroanatomical correlates of selected psycholinguistic effects. Clearly, this nascent work is exciting and informative, and many issues remain unexplored. In the final portion of this section, we will consider how neuroimaging has advanced our understanding of word recognition processes above and beyond traditional behavioral work.

8.1.3. What constraints are afforded by neuroimaging techniques?

It is incontrovertible that we know a great deal more about the neuroanatomy of language today than a mere 10 years ago. Nevertheless, it is also clear that neuroimaging of cognitive processes is still a relatively new area of investigation. Even though the conclusions we have presented are preliminary and may be revised not too far in the future, we would contend that neuroimaging data is an essential adjunct to response latency and accuracy data. Most obviously, neural correlates of behavior provide another level of explanation (Marr, 1982) that reveals how reading processes are physically instantiated in the brain. Moreover, Palmer et al. (2004) have cogently argued that collecting functional neuroimaging data affords two other important advantages. One, brain activation data can powerfully complement behavioral measures. For example, young and older adults may perform identically on a task (in terms of response latencies and error rates), but show marked differences in brain activity (also see Hauk & Pulvermüller, 2004; Schlaggar et al., 2002). In addition, neuroimaging data may also be useful in informing theories and adjudicating between competing models. For example, the DRC model (Coltheart et al., 2001) and the connectionist model (Plaut et al., 1996) adopt very different architectures for naming words that have been difficult to discriminate based on behavioral data, but it may be possible that converging evidence of the role of different neural substrates dedicated to specific operations may provide important information on how the brain implements the processes involved in word recognition.

8.2. Large-Scale Studies vs. Factorial Studies of Word Recognition

Word recognition researchers have traditionally employed factorial designs where item variables of interest (e.g., length, frequency, etc.) have been "manipulated," and other factors known to affect performance have been controlled. This approach has been useful, but there are some limitations (see Balota et al., 2004). Recently, researchers have examined word recognition performance for large sets of words that are not constrained by selection factors, e.g., virtually all monosyllabic words (see Balota & Spieler, 1998; Besner & Bourassa, 1995; Kessler, Treiman, & Mullennix, 2002; Spieler & Balota, 1997; Treiman et al., 1995). Such datasets are useful in a number of ways. For example, using standard predictor variables, Balota et al. (2004) accounted for 49 and 50% of the variance in the lexical decision and speeded naming performance, respectively for a dataset of 2428 words. This is a multifold increase over current computational models (for a discussion of pros and cons for using accounted for variance as a critical variable in evaluating a model's performance, see Balota & Spieler, 1998; Seidenberg & Plaut, 1998). This outcome was obtained despite the success these computational models have had in accounting for performance at the factor level. The large-scale item-level analyses provide another potentially important constraint in the evaluation of theoretical approaches to word processing. More recently, Balota and colleagues have collected naming and lexical decision latencies for over 40,000 words (Balota et al., 2002). The English Lexicon Project website (http://elexicon.wustl.edu) provides a comprehensive data set of behavioral measures that researchers can easily access, via a search engine, along with a rich set of item characteristics. Hopefully, this dataset will be helpful in extending current models to multisyllabic words, which as noted above is a potentially serious limitation in current models. Finally, as mentioned earlier, recent attempts to ground semantics in large scale natural databases of language use (e.g., Burgess & Livesay, 1998; Landauer & Dumais, 1997; Steyvers & Tenenbaum, 2005) have also been quite informative. Clearly, the computational power available today that affords analyses of these large-scale databases appears to be providing an important additional constraint on theory development.

8.3. RT Distributional Analyses

In standard word recognition experiments, one compares the mean response latency across several conditions to determine if the predictions generated by an experimental hypothesis are correct. Of course mean performance is not the only estimate of performance across a set of trials. Researchers have long noted that means of conditions are only one estimate available from performance. For example, in the standard Stroop task (i.e., naming the color that a word appears in), Heathcote, Popiel, and Mewhort (1991) provided a useful demonstration of how the shape of a response time distribution can provide useful information beyond estimates of central tendency. They found that the incongruent condition (e.g., the word *blue* appearing in the color red), compared to the neutral condition (e.g., the word *block* appearing in the color red), increased both the skewing and the central tendency of the reaction time distribution, but amazingly, the congruent condition (e.g., the word *red* appearing in the color red) *increased* skewing and *decreased* the central tendency, which basically masked any effect in means (for a replication of this pattern, see Spieler, Balota, & Faust, 1996). These researchers have fit reaction time distributions to ex-Gaussian functions, but other functions such as the Weibull or ex-Wald could also capture useful characteristics of the reaction time distributions. As theories become more precise regarding item level performance, there should be an increased level of sophistication regarding the predictions concerning the underlying reaction time distributions. For example, Balota and Spieler (1999) found that frequency and repetition influenced these parameters differently depending on the dependent measures, i.e., naming vs. lexical decision (however, see Andrews & Heathcote, 2001). Ratcliff et al. (2004) have recently used reaction time distributions to more powerfully test a diffusion model of lexical decision performance. As models become more sophisticated, the precision of reaction time distribution analyses will be critical in their evaluation.

8.4. Individual Differences

Just as one may be losing information when averaging across items to estimate means, one is also losing information when averaging across individuals. Of course, there are standard comparisons of individual differences as a function of age, acquired or developmental dyslexia, or other neuropsychological impairment (see Perfetti, this volume), however, another possibility is that individuals may produce particular profiles of lexical processing. For example, if indeed the dual-route model is correct, one might find that some subjects rely more on lexical pathways, while other subjects rely more on sublexical pathways, and this could indeed be tied to the manner in which they were originally taught to read or inherent individual differences in capacities. The recent explosion of interest in differences in working memory capacity has been quite successful in identifying distinct cognitive processing profiles (see, for example, Engle, Kane, & Tuholski, 1999). With the advent of large datasets on individual subjects (see megastudies mentioned earlier) it is quite possible that such differences could be observed (for processing speed modulating the effects of orthographic neighborhood size, see Balota et al., 2004). Of course, this may also push researchers to more closely consider the reliability of effects, which at least within one domain, semantic priming, appear to be surprisingly low (see Stolz, Besner, & Carr, 2005).

9. CONCLUDING REMARKS

In the present chapter, we have attempted to provide the reader with an overview of the major issues addressed in the word recognition literature. To conclude, we would like to summarize 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 post-perceptual 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 should be especially careful in developing (or understanding) 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 there are separable analyses performed en route to word recognition or the apparent influences of multiple pathways 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 the strong evidence for multiple distinct access pathways. Clearly, this is still a central issue in current state of model development (see Andrews, 2006).

A third theme in the present review is the 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. We have reviewed evidence indicating that compared to high-frequency words, low-frequency words produce larger effects of bigram frequency, spelling-to-sound consistency, word-body strength, concreteness, semantic priming, task (lexical decision task vs. category verification vs. pronunciation), repetition priming, 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, or a z-score transform of the data (taking into account variability), many of these interactions would disappear (for a detailed discussion of these issues, see Faust, Balota, Spieler, & Ferraro, 1999). Clearly, the assumption of a linear relations between response latency and underlying cognitive operations is a simplifying assumption, which will ultimately need to be faced by those studying the time-course of processes involved in visual word recognition, along with other cognitive operations.

In sum, we are hopeful that the reader agrees that at some level the word is to cognitive psychologists and psycholinguist 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, this clearly is *not* the case; considerable progress has been made, especially within the last decade. Of course, 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 develop procedures for generating and constraining our questions in the future. Given the new analytic methods that have come on line recently this will indeed be a very exciting next decade of research!

ACKNOWLEDGMENTS

This work has been supported by Grant BCS 0001801 from the National Science Foundation and Grants AGO3991 and RO1 AG1702 from the National Institute on Aging. Thanks are extended to Matthew Traxler for helpful comments on an earlier version of this chapter.

REFERENCES

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

Abrams, R. L., & Greenwald, A. G. (2000). Parts outweigh the whole (word) in unconscious analysis of meaning. *Psychological Science*, *11*, 118–124.

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. (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 and 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 orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 234–254.

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

Andrews, S. (2006). *All about words: Current issues in lexical processing*. Psychology Press, Inc. **AQ16**

Andrews, S., & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: A matter of some moment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 514–544.

Andrews, S., & Scarratt, D. R. (1998). Rule and analogy mechanisms in reading nonwords: Hough dou peapel rede gnew wirds? *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1052–1086.

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

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

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

Azuma, T., & Van Orden, G. C. (1997). Why safe is better than fast: The relatedness of a word's meaning affects lexical decision times. *Journal of Memory and Language*, *36*, 484–504.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Baayen, R. H., & Schreuder, R. (Eds). (2003). *Morphological structure in language processing*, Berlin: Mouton de Gruyter.

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., 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., Cortese, M. J., Hutchison, K. A., Neely, J. H., Nelson, D., Simpson, G. B., & Treiman, R. (2002). *The English lexicon project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords.* Accessed January 30, 2004, on the Washington University Web site: http://elexicon.wustl.edu.

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

Balota, D. A., & 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., & 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., & Paul, S. T. (1996). Summation of activation: Evidence from multiple primes that converge and diverge within semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 827–845.

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory and Cognition*, *29*, 639–647.

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.

Balota, D. A., & Spieler, D. H. (1998). The utility of item-level analyses in model evaluation: A reply to Seidenberg and Plaut. *Psychological Science*, *9*, 238–240.

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

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, *54*, 462–479.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*, 230–244.

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.

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

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

Besner, D., & Bourassa, D. C. (1995). Localist and parallel distributed processing models of visual word recognition: A few more words. Paper presented at the Annual Meeting of the Canadian Brain, Behaviour, and Cognitive Science Society, Halifax, NS, Canada, June.

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* (pp. 121–135). 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., & MacLeod, C. M. (1990a). Visual word recognition: A dissociation of lexical and semantic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 862–869. **AQ17**

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

Binder, J. R., McKiernan, K. A., Parsons, M. E., Westbury, C. F., Possing, E. T., Kaufman, J. N., & Buchanan, L. (2003). Neural correlates of lexical access during visual word recognition. *Journal of Cognitive Neuroscience*, *15*, 372–393.

Bjork, E. L., & Estes, W. K. (1973). Letter identification in relation to linguistic context and masking conditions. *Memory and 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, and Cognition*, *13*, 582–594.

Bodner, G. E., & Masson, M. E. J. (1997). Masked repetition priming of words and nonwords: Evidence for a nonlexical basis for priming. *Journal of Memory and Language*, *37*, 268–293.

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

Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory & Language*, *50*, 456–476.

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.

Borowsky, R., & Besner, D. (1993). Visual word recognition: A multistage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 813–840.

Borowsky, R., & Besner, D. (2005). Parallel distributed processing and lexical–semantic effects in visual word recognition: Are a few stages necessary? *Psychological Review* (in press). **AQ18**

Bradley, D. (1979). Lexical representations of derivational relations. In: M. Aranoff, & 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.

Brown, G. D. A., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and naming latency. *Memory & Cognition*, *15*, 208–216.

Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review*, *8*, 531–544.

Burgess, C., & Livesay, K. (1998). The effect of corpus size in predicting RT in a basic word recognition task: Moving on from Kucera and Francis. *Behavior Research Methods, Instruments, & Computers*, *30*, 272–277.

Carello, C., Turvey, M. T., & Lukatela, G. (1992). Can theories of word recognition remain stubbornly nonphonological? In: R. Frost, & L. Katz (Eds), *Orthography, phonology, morphology, and meaning* (pp. 211–226). Amsterdam: North-Holland.

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, and 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.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of the orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, *23*, 857–871.

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

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

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.

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

Chateau, D., & Jared, D. (2003). Spelling-sound consistency effects in disyllabic word naming. *Journal of Memory and Language*, *48*, 255–280.

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 and Cognition*, *12*, 590–606.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., & Michel, F. (2000). The visual word form area. Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123*, 291–307.

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 : 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.

Coltheart, M., & Rastle, K. (1994). Serial processing in reading alour: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1197–1220.

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

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.

Cortese, M. J. (1998). Revisiting serial position effects in reading aloud. *Journal of Memory and Language*, *39*, 652–666.

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

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

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

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.

Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Mapping*, *5*, 329–340.

Damian, M. F. (2001). Congruity effects evoked by subliminally presented primes: Automaticity rather than semantic processing. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 154–165.

Davis, & Castles, A., & Iakovidis, E. (1998). Masked homophone and pseudohomophone priming in children and adults. *Language and Cognitive Processes*, *13*, 625–651. **AQ19**

Davis, C. J. (1999). *The self-organizing lexical acquisition and recognition (SOLAR) model of visual word recognition* Doctoral dissertation, University of New South Wales, Sydney, New South Wales, Australia, 1999. *Dissertation Abstracts International*, *62*, 594. Available on the Web at www.maccs.mq.edu.au/~colin/thesis.zip.

Davis, C. J., & Bowers, J. S. (2004). What do letter migration errors reveal about letter position coding in visual word recognition? *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 923–941.

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 and Cognition*, *10*, 358–370.

De Houwer, J., Hermans, D., & Eelen, P. (1998). Affective and identity priming with episodically associated stimuli. *Cognition and Emotion*, *12*, 145–169.

Dehaene, S., Naccache, L., Le Clec, H. G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P. F., & Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, *395*, 597–600.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321.

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.

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

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In: A. Miyake, & P. Shah, (Eds), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). London, UK: Cambridge Press.

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

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin*, *125*, 777–799.

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

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*, *15*, 115–141.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.

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.

Feldman, L. B., & Basnight-Brown, D. (2005). The role of morphology in visual word recognition: Stem senses and semantic richness. In: E. Grigorenko, A. Naples (Eds), *Single word reading: Cognitive, behavioral and biological perspectives.* Mahwah, NJ: Lawrence Erlbaum, in press. **AQ20**

Fiebach, C. J., Friederici, A. D., Muller, K., & von Cramon, D. Y. (2002). fMRI evidence of dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience*, *14*, 11–23.

Fiez, J. A., Balota, D. A., Raichle, M. E., & Petersen, S. E. (1999). Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron*, *24*, 205–218.

Fiez, J. A., & Petersen, S. E. (1998). Neuroimaging studies of word reading. *Proceedings of the National Academy of Science*, *95*, 914–921.

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

Forster, K. I. (1976). Accessing the mental lexicon. In: R. J. Wales, & E. C. T. Walker (Eds), *New approaches to language mechanisms* (pp. 257–287). Amsterdam: 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. (1998). The pros and cons of masked priming. *Journal of Psycholinguistic Research*, *27*, 203–233.

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

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

Forster, K. I., & Davis, C. (1991). The density constraint on form-priming in the naming task: Interference effects from a masked prime. *Journal of Memory and Language*, *30*, 1–25.

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 696–713.

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

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. (Ed.). (1984a). *Handbook of psycholinguistics*. San Diego: Academic Press. **AQ21**

Gernsbacher, M. A. (1984b). 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. **AQ22**

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.

Gold, B. T., Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Snyder, A. Z., Salat, D. H., Fischl, B., Dale, A. M., Morris, J. C., & Buckner, R. L. (2005). Differing neuropsychological and neuroanatomical correlates of abnormal reading in early-stage semantic dementia and dementia of the Alzheimer type. *Neuropsychologia*, *43*, 833–846.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, 251–279.

Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2005). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology (submitted). **AQ23**

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. (1992). Orthographic neighborhoods and visual word recognition. In: R. Frost (Ed.), *Orthography, phonology, morphology, and meaning* (pp. 131–146). Oxford, England: North-Holland.

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

Greenwald, A. G., Abrams, R. L., Naccache, L., & Dehaene, S. (2003). Long-term semantic memory versus contextual memory in unconscious number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 235–247.

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.

Harm, M. W., & Seidenberg, M. S. (2004) Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*, 662–720.

Harm, M., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*, 491–528.

Hauk, O., & Pulvermüller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology*, *115*, 1090–1103.

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.

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

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

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

Herbster, A. N., Mintun, M. A., Nebes, R. D., & Becker, J. T. (1997). Regional cerebral blood-flow during word and non-word reading. *Human Brain Mapping*, *5*, 84–92.

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.

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1331–1356.

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and synonymy effects in lexical decision, naming, and semantic categorization tasks: Interactions between orthography, phonology, and semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 686–713.

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

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.

Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese*. Hillsdale, NJ: Erlbaum.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Science*, *79*, 2554–2558.

Howard, I. (1972). *A directional theory of rule application in phonology.* Dissertation at Massachusetts Institute of Technology. AQ24

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

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, *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., & Taylor, D. E. (1982). Automatic phonological priming in visual word recognition. *Memory and Cognition*, *10*, 576–590.

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, UK: Erlbaum.

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

Ischebeck, A., Indefrey, P., Usui, N., Nose, I., Hellwig, F., & Taira, M. (2004). Reading in a regular orthography: An fMRI study investigating the role of visual familiarity. *Journal of Cognitive Neuroscience*, *16*, 727–741.

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

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 and 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, and Cognition*, *18*, 483–491.

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin* (in press). **AQ25**

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

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.

Kahan, T. A., Neely, J. H., & Forsythe, W. J. (1999). Dissociated backward priming effects in lexical decision and pronunciation tasks. *Psychonomic Bulletin & Review*, *6*, 105–110.

Kawamoto, A., Kello, C., Jones, R., & Bame, K. (1998). Initial phoneme versus whole word criterion to initiate pronunciation: Evidence based on response latency and initial phoneme duration. *Journal of Experimental Psychology:Learning, Memory, and Cognition*, *24*, 862–885.

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 & 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. *The 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 and Cognition*, *18*, 289–298.

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

Kessler, B., & Treiman, R. (2003). Is English spelling chaotic? Misconceptions concerning its irregularity. *Reading Psychology*, *24*, 267–289.

Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language*, *47*, 145–171.

Kessler, B., Treiman, R., & Mullennix, J. (2005). Feedback consistency effects in single-word reading (in press). **AQ26**

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

King, J. W., & Kutas, M. (1998). Neural plasticity in the dynamics of human visual word recognition. *Neuroscience Letters*, *224*, 61–64.

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.

Kinoshita, S. (2000). The left-to-right nature of the masked onset priming effect in naming. *Psychonomic Bulletin and Review*, *7*, 133–141.

Kinoshita, S., & Lupker, S. J. (2002). Effects of filler type in naming: Change in time criterion or attentional control of pathways? *Memory and Cognition*, *30*, 1277–1287.

Kinoshita, S., & Lupker, S. J. (2003). Priming and attentional control of lexical and sublexical pathways in naming: A re-evaluation. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *29*, 405–415. **AQ27**

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

Kinoshita, S., Lupker, S. J., & Rastle, K. (2004) Modulation of regularity and lexicality effects in reading aloud. *Memory and Cognition*, *32*, 1255–1264.

Koriat, A. (1981). Semantic facilitation in lexical decision as a function of prime-target association. *Memory and 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.

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

Kunde, W., Kiesel, A., & Hoffman, J. (2003). Conscious control over the content of unconscious cognition. *Cognition*, *88*, 223–242.

Kutas, M., & Federmeier, K. D. (2000): Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science*, *4*, 463–470.

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

Lambert, A. J., Payne, B. K., Jacoby, L. L., Shaffer, L. M., Chasteen, A. L., & Khan, S. K. (2003). Stereotypes as dominant responses: On the "social facilitation" of prejudice in anticipated public context. *Journal of Personality and Social Psychology*, *84*, 277–295.

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

Lee, Y., Binder, K. S., Kim, J., Pollatsek, A., & Rayner, K. (1999). Activation of phonological codes during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 948–964.

Lewis, M. B., Gerhand, S., & Ellis, H. D. (2001). Re-evaluating age-of-acquisition effects: Are they simply cumulative-frequency effects? *Cognition*, *78*, 189–205.

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

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

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 and Cognition*, *14*, 95–103.

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

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.

Lukatela, G., & Turvey, M. T. (1994). Visual lexical access is initially phonological: 1. Evidence from associative priming by words, homophones, and pseudohomophones. *Journal of Experimental Psychology: General*, *123*, 107–128.

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 and 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.

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.

Marr, D. (1982). *Vision.* San Francisco, CA: W. H. Freeman.

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 Ltd.

Massaro, D. (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.

Massaro, D. W., & Cohen, M. M. (1994). Visual, orthographic, phonological, and lexical influences in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1107–1128.

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* (pp. 233–263). Hillsdale, NJ: Erlbaum.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 3–23.

Masson, M. E. J., & Isaak, M. I. (1999). Masked priming of words and nonwords in a naming task: Further evidence for a nonlexical basis for priming. *Memory and Cognition*, *27*, 399–412.

Mayall, K., & Humphreys, G. W. (1996). Case mixing and the task-sensitive disruption of lexical processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 278–294.

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 pentagram 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 pentagram 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 and 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.

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

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

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 and 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.

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

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.

Monaghan, J., & Ellis, A. W. (2002). What, exactly, interacts with spelling-sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *28*, 183–206.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). 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.

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 116–153.

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

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

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

Naccache, L., & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition*, *80*, 215–229.

Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory and 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., & Kahan, T. A. (2001). Is semantic activation automatic? A critical re-evaluation. In: H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Suprenant (Eds), *The nature of remembering: Essays in honor of Robert G. Crowder*, (pp. 69–93). Washington, DC: American Psychological Association.

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). New York: Academic Press.

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

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

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms*. Retrieved June 1, 2000, from the University of South Florida Web site: http://www.usf.edu/FreeAssociation/.

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2005). Re-examining word length effects in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin and Review* (in press). AQ28

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

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

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana, IL: The 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 and Psychophysics*, *8*, 273–275.

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

Palmer, E. D. (2003). *Neural mechanisms underlying single word reading: An fMRI investigation*. Unpublished doctoral dissertation, Washington University at St. Louis, St. Louis, MO.

Palmer, E. D., Brown, T. T., Petersen, S. E., & Schlaggar, B. L. (2004). Investigation of the functional neuroanatomy of single word reading and its development. *Scientific Studies of Reading*, *8*, 203–223.

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* (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.

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*, 181–192.

Peereman, R., Content, A., & Bonin, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory and Language*, *39*, 151–174.

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

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1989). Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience*, *1*, 153–170.

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.

Pexman, P. M., Lupker, S. J., & Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 139–156.

Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. In: Cognitive Science Society (Eds), *Proceedings of the 17th annual conference of the Cognitive Science Society* (pp. 37–42). Hillsdale, NJ: Erlbaum.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, *12*, 765–805.

Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review*, *107*, 786–823.

Plaut, D. C., & Gonnerman, L. M. (2000) Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes*, *15*, 445–485.

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

Plourde, C. E., & Besner, D. (1997). On the locus of the word frequency effect in visual word recognition. *Canadian Journal of Experimental Psychology*, *51*, 181–194.

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

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

Pollatsek, A., Perea, M., & Carreiras, M. (2005). Does conal prime CANAL more than cinal? Masked phonological priming effects in Spanish with the lexical decision task. *Memory & Cognition*, *33*, 557–565.

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.

Price, C. (2004). The functional neuroanatomy of reading. In: R. S. J. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan, C. J. Price, S. Zeki, J. Ashburner, & W. Penny (Eds), *Human brain function* (2nd ed., pp. 547–562). San Diego, CA: Elsevier.

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

Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Constable, R. T., Skudlarski, P., Fulbright, R. K., Bronen, R. A., Shankweiler, D. P., Katz, L., Fletcher, J. M., & Gore, J. C. (1996). Cerebral organization of component processes in reading. *Brain*, *119*, 1221–1238.

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

Quinlan, P. T. (2003). Visual feature integration theory: Past, present, and future. *Psychological Review*, *129*, 643–673.

Rapp, B. C. (1992). The nature of sublexical orthographic organization: The bigram trough hypothesis examined. *Journal of Memory & Language*, *31*, 33–53.

Rastle, K., & Coltheart, M. (1999). Lexical and nonlexical phonological priming in reading aloud. *Journal of Experimental Psychology: Human Perception and Perception*, *25*, 461–481.

Rastle, K., & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory & Language*, *42*, 342–364.

Rastle, K., & Coltheart, M. (2006). Is there serial processing in the reading system; and are there local representations? S. Andrews (Ed.). *All about words: Current issues in lexical processing.* Psychology Press, to appear. **AQ29**

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, *11*, 1090–1098.

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

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, *111*, 159–182.

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 and 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.

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

Reynolds, M., & Besner, D. (2005a). Basic processes in reading: A critical review of pseudohomophone effects in naming and a new computational account. *Psychonomic Bulletin and Review*, *12*, 622–246.

Reynolds, M., & Besner, D. (2005b). Contextual control over lexical and sub-lexical routines in reading aloud. *Psychonomic Bulletin and Review* (in press). **AQ30**

Rissman, J., Eliassen, J. C., & Blumstein, S. E. (2003). An event-related fMRI investigation of implicit semantic priming. *Journal of Cognitive Neuroscience*, *15*, 1160–1175.

Rodd, J. M. (2004). The effect of semantic ambiguity on reading aloud: A twist in the tale. *Psychonomic Bulletin & Review*, *11*, 440–445.

Roget, P. M. (1911). *Roget's thesaurus of English words and phrases*. Available from Project Gutenberg, Illinois Benedictine College, Lisle, IL.

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

Rueckl, J. G., Mikolinski, M., Raveh, M., Miner, C., & Mars, F. (1997). Morphological priming, fragment completion, and connectionist networks. *Journal of Memory and Language*, *36*, 382–345.

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.

Rumsey, J. M., Horwitz, B., Donahue, C., Nace, K., Maisog, J. M., & Andreason, P. (1997). Phonologic and orthographic components of word recognition: A PET–rCBF study. *Brain*, *119*, 739–759.

Sandra, D., & Taft, M. (Eds). (1994). *Morphological structure, lexical representation and lexical access: A special issue of Language and Cognitive Processes*. Hove, UK: Erlbaum.

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

Schiller, N. O. (2004). The onset effect in word naming. *Journal of Memory and Language*, *50*, 477–490.

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

Schlaggar, B. L., Brown, T. T., Lugar, H. M., Visscher, K. M., Miezin, F. M., & Petersen, S. E. (2002). Functional neuroanatomical differences between adults and school-age children in the processing of single words. *Science*, *296*, 1476–1479.

Schomaker, L., & Segers, E. (1999). Finding features used in the human reading of cursive handwriting. *International Journal on Document Analysis and Recognition*, *2*, 13–18.

Schomaker, L. R. B., & Van Galen, G. P. (1996). Computer modeling of handwriting. In: A. F. J. Dijkstra, & K. J. M. J. De Smedt (Eds), *Computational psycholinguistics: AI and connectionist models of human language processing* (pp. 386–420). London: Taylor & Francis.

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 & Language*, *27*, 499–520.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 876–900.

Seidenberg, M., & Plaut, D. C. (1998). Evaluating word reading models at the item level: Matching the grain of theory and data. *Psychological Science*, *9*, 234–237.

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. (2005). Connectionist models of word reading. *Current Directions in Psychological Science* (in press). AQ31

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., Plaut, D. C., Petersen, A. S., McClelland, J. L., & McRae, K. (1994). Nonword pronunciation and models of word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1177–1196.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. (1984a). 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. (1984b). Pre- and post-lexical loci of contextual effects on word recognition. *Memory and Cognition*, *12*, 315–328.

Selfridge, O. G. (1959). Pandemonium: A paradigm for learning. In: *Proceedings of a symposium on the mechanization of thought processes* (pp. 511–526). London: H. M. Stationery Office. AQ32

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: Indiana University Linguistics Club.

Selkirk, E. O. (1982). The syllable. In: H. van der Hulst, & N. Smith (Eds), *The structure of phonological representations* (Part II pp. 337–383). Dordrecht: Foris.

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.

Sereno, S. C., & Rayner, K. (1992). Fast priming during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 173–184.

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. *The Quarterly Journal of Experimental Psychology*, *35A*, 111–138.

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

Simos, P. G., Breier, J. L., Fletcher, J. M., Foorman, B. R., Castillo, E. M., & Papanicolaou, A. C. (2002). Brain mechanisms for reading words and nonwords: An integrated approach. *Cerebral Cortex*, *12*, 297–305.

Simpson, G. B., & Kang, H. (1994). The flexible use of phonological information in word recognition in Korean. *Journal of Memory & Language*, *33*, 319–331.

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.

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science*, *8*, 411–416.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in normal older adults and individuals with senile dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 461–479.

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., & Painton, S. (1979a). Memory representation for prefixed words. *Journal of Verbal Learning and Verbal Behavior*, *18*, 733–743. AQ33

Stanners, R. F., Neiser, J. J., Hernon, W. P., & Hall, R. (1979b). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior*, *18*, 399–412.

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

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, *30*, 34–78.

Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, *64*, 1358–1368.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*, 41–78.

Stolz, J. A., Besner, D., & Carr, T. H. (2005). Semantic priming is robust but unreliable: Automaticity, strategic control, and interpersonal stability in the activation and evaluation of semantic knowledge. *Visual Cognition* (in press). AQ34

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language*, *36*, 337–359. AQ35

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1140–1154.

Strain, E., Patterson, K., & Seidenberg, M. S. (2002). Theories of word naming interact with spelling-sound consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 207–214.

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

Taft, M. (1979b). Lexical access via an orthographic code: The basic orthographic syllabic structure (BOSS). *Journal of Verbal Learning and Verbal Behavior*, *18*, 21–39.

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.

Tan, L. H., & Perfetti, C. A. (1998). Phonological codes as early sources of constraint in Chinese word identification: A review of current discoveries and theoretical accounts. *Reading and Writing: An Interdisciplinary Journal*, *10*, 165–200.

Tan, L. H., & Perfetti, C. A. (1999). Phonological activation in visual identification of Chinese two-character words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 382–393.

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

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. Hillsdale, NJ: Erlbaum.

Treiman, R. (1989). The internal structure of the syllable. In: G. Carlson, & M. Tanenhaus (Eds), *Linguistic structure in language processing* (pp. 27–52). Dordrecht, the Netherlands: Kluwer.

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). Hove, UK: Erlbaum.

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

Treiman, R., Kessler, B., & Bick, S. (2002). Context sensitivity in the spelling of English vowels. *Journal of Memory and Language*, *47*, 448–468.

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

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

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, *6*, 171–178.

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

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 perception of objects. *Cognitive Psychology*, *14*, 107–141.

Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *NeuroImage*, *16*, 765–780.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory and 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.

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

West, R. F., & Stanovich, K. E. (1986). Robust effects of syntactic structure on visual word processing. *Memory and Cognition*, *14*, 1040–1112.

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

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

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

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

Wijk, A. (1969). Regularized Inglish. The only practicable solution of the English spelling reform problem. In: W. Haas (Ed.), *Alphabets for English*. Manchester: Manchester 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.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, *72*, 262–274.

Yates, M., Locker, L., & Simpson, G. (2004). The influence of phonological neighborhood on visual word perception. *Psychonomic Bulletin and Review*, *11*, 452–457.

Yates, M. (in press). Phonological neighbors speed visual word processing: Evidence from multiple tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* **AQ36**

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Science.

Zevin, J. D., & Balota, D. A. (2000). Priming and attentional control of lexical and sublexical pathways during naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 121–135.

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language*, *47*, 1–29.

Zevin, J. D., & Seidenberg, M. S. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory and Cognition*, *32*, 31–38.

Zevin, J. D., & Seidenberg, M. S. (2005) Simulating consistency effects and individual differences in nonword naming (submitted). **AQ37**

Ziegler, J. C., Ferrand, L., Jacobs, A. M., Rey, A., & Grainger, J. (2000). Visual and phonological codes in letter and word recognition: Evidence from incremental priming. *Quarterly Journal of Experimental Psychology*, *53A*, 671–692.

Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language*, *37*, 533–554.

Ziegler, J. C., & Perry, C. (1998). No more problems in Coltheart's neighborhood: resolving neighborhood conflicts in the lexical decision task. *Cognition*, *68*, 53–62.

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1131–1161.

AUTHOR QUERY FORM

ELSEVIER

Handbook of Psycholinguistics

JOURNAL TITLE:	**HP-TRAXLER**
ARTICLE NO:	**Ch009**

Queries and / or remarks

Query No	Details required	Author's response
AQ1	Noble (1960), Mayall, Humphreys, & Olson (1997), Jared (1997), Coltheart et al. (1997), Stone, Vanhoy, and Van Orden (1998), Zorzi (1999), Rastle & Coltheart (2005), Andrews & Heathcote (2002), Balota et al. (1999), Forster & Chen (1996), Johnson & Pugh, 1994, Andrews (1997), Sears, Hino, & Lupker (1999), Rayner (1998), Balota & Abrams (1995), Petersen, Fox, Posner, Mintun and Raichle (1988), Seidenberg & Plaut (1998) are not listed in the reference list.	
AQ2	In Schomaker & Van Galen (1996), Van Galen is spelt as 'Van Galen' and 'van Galen'. Please confirm the correct spelling.	
AQ3	Is 'GPC' well-known abbreviation? If not, pleaseexpand at the first instance.	
AQ4	In Kessler, Treiman, & Mullennix (2005), Mullennix is spelt as 'Mullennix' and 'Mullenix'. Please confirm the correct spelling.	
AQ5	In Ans et al. (1998), Ans is spelt as 'Anns' and 'Ans'. Please confirm the correct spelling.	
AQ6	In Gonnerman, Seidenberg, & Andersen (2005), Andersen is spelt as 'Andersen' and 'Anderson'. Please confirm the correct spelling.	
AQ7	In Baayen & Schreuder (2003), Schreder is spelt as 'Schreuder' and 'Schreder'. Please confirm the correct spelling.	

AUTHOR QUERY FORM

ELSEVIER

Handbook of Psycholinguistics

JOURNAL TITLE:	**HP-TRAXLER**
ARTICLE NO:	**Ch009**

Queries and / or remarks

Query No	Details required	Author's response
AQ8	In Connine et al. (1990), Yelens is spelt as 'Yelen' and 'Yelens'. Please confirm the correct spelling.	
AQ9	In Baayen, Piepenbrock, & van Rijn (1993), van Rijn is spelt as 'van Rijn' and 'Rijn'. Please confirm the correct spelling.	
AQ10	In Carreiras, Perea, & Grainger (1997), Grainger is spelt as 'Grainger' and 'Granger'. Please confirm the correct spelling.	
AQ11	In De Groot (1989), De Groot is spelt as 'De Groot' and 'de Groot'. Please confirm the correct spelling.	
AQ12	In Smith, Shoben, & Rips (1974), Rips is spelt as 'Rips' and 'Ripps'. Please confirm the correct spelling.	
AQ13	In Den Heyer, Sullivan, & McPherson (1987), Den Heyer is spelt as 'Den Heyer' and 'den Heyer'. Please confirm the correct spelling.	
AQ14	In Damian (2001), Damian is spelt as 'Damian' and 'Damien'. Please confirm the correct spelling.	
AQ15	In Den Heyer, Briand, & Dannenbring (1983), Den Heyer is spelt as 'Den Heyer' and 'den Heyer'. Please confirm the correct spelling.	
AQ16	Please provide place of publication in Andrews (2006).	

AUTHOR QUERY FORM

ELSEVIER

Handbook of Psycholinguistics

JOURNAL TITLE:	**HP-TRAXLER**
ARTICLE NO:	**Ch009**

Queries and / or remarks

Query No	**Details required**	**Author's response**
AQ17	Since there are two publications of Besner et al. (1990), as per style, we cross-refer it as '1990a' and '1990b' for tagging purposes. Please go through each citation and indicate by comparing with the list as either 'a' or 'b'.	
AQ18	Please provide Vol. no and page range in Borowsky & Besner (2005).	
AQ19	Please provide initials of the author 'Davis' in Davis et al. (1998).	
AQ20	Please provide complete details in Feldman & Basnight-Brown (2005).	
AQ21	Since there are two publications of Gernsbacher (1984), as per style, we cross-refer it as '1990a' and '1990b' for tagging purposes. Please go through each citation and indicate by comparing with the list as either 'a' or 'b'.	
AQ22	Please provide editor names in Gibson et al. (1963).	
AQ23	Please provide complete details in Gonnerman et al. (2005).	
AQ24	Please provide place of publication in Howard (1972).	
AQ25	Please provide Vol. no and page range in Juhasz (2005).	
AQ26	Please provide complete details in Kessler et al. (2005).	

AUTHOR QUERY FORM

ELSEVIER

Handbook of Psycholinguistics

JOURNAL TITLE:	**HP-TRAXLER**
ARTICLE NO:	**Ch009**

Queries and / or remarks

Query No	Details required	Author's response
AQ27	Since there are two publications of Kinoshita & Lupker (2003), as per style, we cross-refer it as '1990a' and '1990b' for tagging purposes. Please go through each citation and indicate by comparing with the list as either 'a' or 'b'.	
AQ28	Please provide Vol. no and page range in New et al. (2005).	
AQ29	Please provide place of publication in Rastle & Coltheart (2006).	
AQ30	Please provide Vol. no and page range in Reynolds & Besner (2005b).	
AQ31	Please provide complete details in Seidenberg (2005).	
AQ32	Please provide editor names in Selfridge (1959).	
AQ33	Since there are two publications of Stanners (1979), as per style, we cross-refer it as '1990a' and '1990b' for tagging purposes. Please go through each citation and indicate by comparing with the list as either 'a' or 'b'.	
AQ34	Please provide Vol. no and page range in Stolz et al. (2005).	
AQ35	Stone et al. (1997), Sears et al. (1995), Seidenberg (1985), Rayner et al. (1978), Noble (1952), Mayall & Humphreys (1996) are not cited in the text.	
AQ36	Please provide year of publication, Vol. no and page range in Yates (in press).	

AUTHOR QUERY FORM

ELSEVIER

Handbook of Psycholinguistics

JOURNAL TITLE:	**HP-TRAXLER**
ARTICLE NO:	**Ch009**

Queries and / or remarks

Query No	Details required	Author's response
AQ37	Please provide complete details in Zevin & Seidenberg (2005).	
AQ38	In Coltheart et al. (2001), Coltheart is spelt as 'Coltheart' and 'Colheart'. Please confirm the correct spelling.	