CHAPTER 8

Visual Word Recognition in Skilled Adult Readers

Michael J. Cortese and David A. Balota

1 Introduction

Visual word recognition is the foundation of reading. It is the place where form meets meaning and hence is the basis by which higher order semantic and comprehension processes take place. Although evidence suggests that reading is an interactive process, one must be able to recognize a word before one can reliably integrate its meaning into a coherent message. The importance of this process is exemplified by the amount of research that has been conducted on word processing in cognitive psychology and related fields. Word recognition research has been central to developments in cognitive neuroscience (e.g., Frost et al., 2005; Pugh et al., 2005), serial versus parallel processing (e.g., Coltheart and Rastle, 1994; Cortese, 1998; Weekes, 1997), attention (e.g., Neely, 1977; Zevin and Balota, 2000), educational practices (e.g., Harm, McCandliss, and Seidenberg, 2003), connectionism (e.g., Coltheart et al., 2001; Plaut et al., 1996), and much more. In this light, the word can be viewed as important to developments in cognitive psychology as the cell has been to developments in the biological sciences (see Balota, 1994).

In this chapter, we begin by providing an overview of the standard tools employed in visual word recognition research. Next, we discuss some general theoretical issues and controversies, although there are other chapters in this volume dedicated to further specification of these issues. In addition, although we will touch on some evidence from neuroscience and neuropsychology, a more detailed discussion of this topic is covered by Sandak et al. in this volume. The major portion of this chapter will be dedicated toward reviewing the major factors that have been identified in adult word recognition performance, along with some recent methodological developments. Finally, we discuss some continuing controversies covered in the literature and summarize the chapter.

2 Tools of the trade

Although there are many tasks used to measure word recognition, the lexical decision and the naming tasks remain the workhorse tasks in this area (Balota et al., 2004). In the lexical decision task, the subject decides as quickly as possible if a letter string is a word or not and indicates his/her decision by pressing a designated button. In the naming task, the participant simply reads aloud a word as quickly and accurately as possible. In addition, priming paradigms have utilized both naming and lexical decision tasks. In the priming paradigm, two letter strings (i.e., words and/or nonwords) are presented sequentially, and participants either name or make a lexical decision to the second letter string, with the two stimuli varying on some dimension such as relatedness (e.g., *consider dog–cat* versus *pen–cat*).

Although lexical decision and naming are the major tools in this area, there are clearly other important measures. For example, measuring eye movements such as gaze and first fixation durations on a given word has been a very useful tool. These measures may be the best measures of the processes tied to word recognition while reading, and in general converge with the results from naming and/or lexical decision performance (cf. Schilling, Rayner, and Chumbley, 1998). Another useful task involves identifying visually degraded words in which words are briefly presented and often forward and/or backward masked by characters (e.g., Tan and Perfetti, 1999). This task has been viewed as providing an indicant of early visual word processing (but see Broadbent, 1967; Catlin, 1973). Other useful tasks include category verification (i.e., subjects are given a category name followed by a potential exemplar, and they decide if the exemplar is a member of the target category), relatedness judgment (i.e., subjects decide if two words are related or not), and rhyme judgments (i.e., subjects decide if two words rhyme or not). It is important to note that all tasks are likely to engage task-specific processes, and also all have some overlapping processes. Hence, no task is process pure (see Jacoby, 1991). Task differences indicate that one should be cautious in using only one task as a microscope into the processes involved in visual word recognition (see Grainger and Jacobs, 1996; Jacobs et al., 1998).

3 General theoretical issues

With the previously mentioned tools in hand, it is important to consider some of the important theoretical issues that have shaped this field. The initial attempts to capture word recognition processes involved two distinct classes of models; activation models (Morton, 1969) and search models (e.g., Forster, 1976). In order to glean some understanding of these models, consider the word frequency effect (i.e., the finding that high-frequency words produce faster and more accurate responses than low-frequency words across a wide range of tasks). Because of the robustness of this effect, most researchers consider this as a starting point for any viable modeling endeavor (see Forster, 2004; Murray and Forster, 2004).

In Morton's classic logogen model (1969), frequency is coded in the resting level activations of the logogens (i.e., word recognition devices). Due to the frequency of exposure, high-frequency words have higher resting level activations than low-frequency words. Thus, logogens for high-frequency words need less stimulus driven bottom-up (and contextually driven top-down) activation than low-frequency words in order to reach their threshold for identification. In search models (e.g., Forster, 1976; Forster and Murray, 2004; Rubenstein, Garfield, and Millikan, 1970), the lexicon is ordered by frequency and searched serially. An initial perceptual analysis defines an orthographic bin (i.e., a likely candidate set that could match the stimulus) that is searched in a frequency-ordered manner. Therefore, high-frequency words will be located in the lexicon before low-frequency words (there are also phonological bins for auditory information and syntactic/semantic bins to accommodate context effects). There are also hybrid models (e.g., Becker, 1979; Paap, Newsome, and McDonald, 1982; Taft and Hambly, 1986) that include characteristics of both activation and search models. For example, Becker posits initial activation processes that define both sensory and semantic search sets. The target stimulus is then compared to the search sets via a frequency-ordered search process.

A third class of models involves a connectionist approach. The current interest in connectionist models of word recognition can be traced back to the interactive activation (IA) models proposed by McClelland and Rumelhart. (1981; Rumelhart and McClelland, 1982; also see Paap et al., 1982). These models were a logical extension of Selfridge's (1959; also see Selfridge and Neisser, 1960) pandemonium model of letter recognition and also incorporated aspects of Morton's logogen model. Selfridge hypothesized that letters were initially analyzed in terms of their visual features that communicated with letter-level representations which in turn communicated with a decision component. The pandemonium model was important because it was one of the first computational models of pattern recognition, and it also benefited from the temporal contiguity with evidence that was accumulating from neuroscience. Specifically, there was accumulating evidence that specific neurons appeared to code primitive features (e.g., horizontal lines, vertical lines, intersections) which could then serve as the building blocks for pattern recognition (e.g., Hubel and Weisel, 1962; 1968).

The IA model (see Figure 8.1) consists of feature detectors, letter detectors, and word detectors. Representations at and between these different levels are connected by facilitatory (arrowed lines) and/or inhibitory (knobbed lines) pathways. When processing a letter string, letter-level representations activate word-level representations via facilitatory connections. More interesting, letter-level representations are reinforced via top-down activation from the word level. Also, activated representations inhibit competing representations within and between levels so that, eventually, only the appropriate representation reaches its threshold. For example, when presented with *book*, its word-level representation becomes active

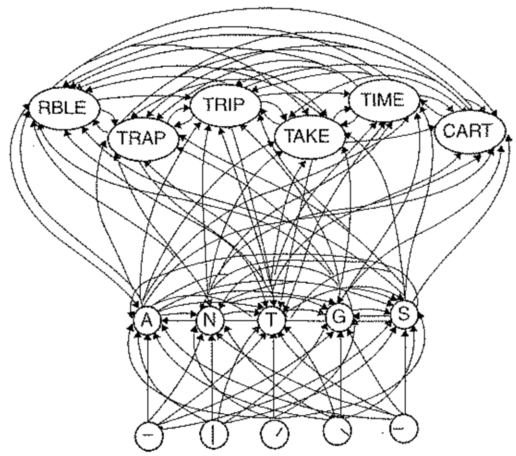

Figure 8.1. The interactive activation (IA) model of McClelland and Rumelhart (1981). Connections marked with arrows denote facilitative connections, and connections marked with circles denote inhibitory connections. This figure was reprinted from McClelland and Rumelhart, "An interactive activation model of context effects in letter perception: Part 1. An account of basic findings." *Psychological Review*, 88, 375–407, 1981, American Psychological Association, reprinted with permission.

across time, and this activation eventually inhibits the word-level representations for *cook*, *boom*, *bock*, etc. Similar inhibitory processing occurs at the feature and letter levels.

Originally, a promising aspect of the IA model was that it could account for the word superiority effect (i.e., the phenomenon that letters are more easily recognized when they occur in words than when they occur in nonwords; cf. Fine, 2001; Reicher, 1969). The explanation is based on the notion of cascadic processing (see Ashby, 1982; McClelland, 1979). Specifically, while information is accumulating in the system, an activated representation does not need to reach its response threshold before it can influence the activation of other representations. Instead, information continuously flows in a bidirectional manner between levels (i.e., among features, letters, and words). Thus, for words, letter-level representations receive bottom-up activation from the feature level and top-down activation from the word level.

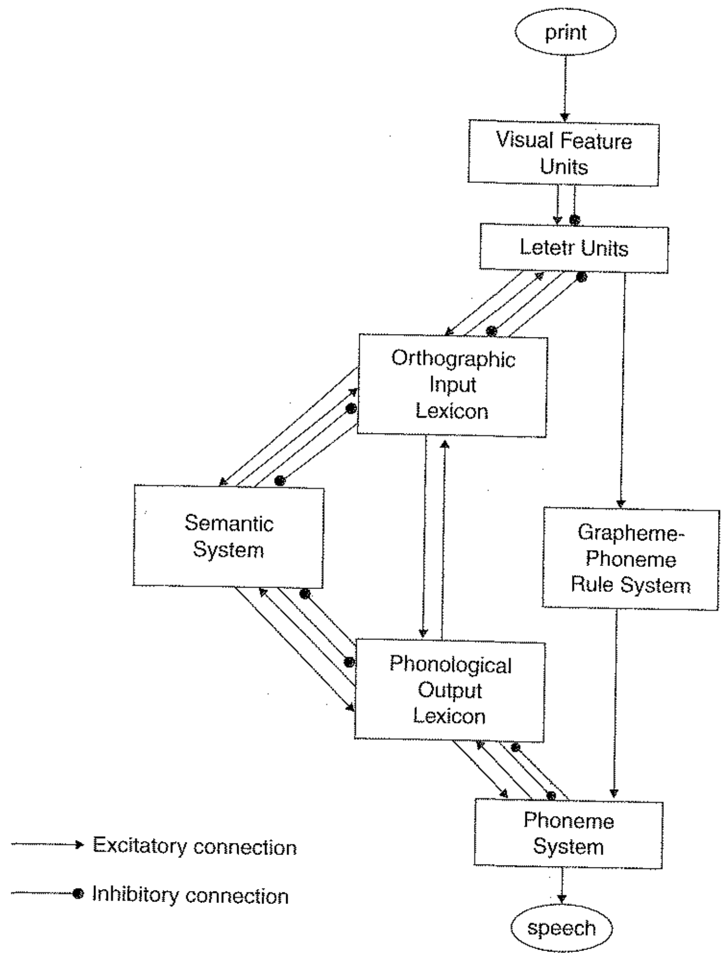

Figure 8.2. The dual-route cascaded (DRC) model of Coltheart et al. (2001). Connections marked with arrows denote facilitative connections, and connections marked with circles denote inhibitory connections. This figure was reprinted from Coltheart et al., "DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204–56, 2001; American Psychological Association, reprinted with permission.

Although the original IA model is important, it primarily dealt with letter recognition performance. More recent models have extended this model to capture word recognition performance. For example, an important model that includes an IA component is the dual-route cascaded (DRC) model (e.g., Coltheart et al., 2001). In the DRC model (see Figure 8.2), two routes are used to process words: a lexical route and a sublexical route. The lexical route is a parallel processor that contains an orthographic and phonological representation for each word in the reader's vocabulary, and has some similarity to the IA model. The sublexical route is a serial processor (working from left to right) that employs a set of grapheme-to-phoneme conversion (GPC) rules to convert letter strings into phonological representations. A grapheme consists of one or more letters that symbolizes a single phoneme (e.g., *champ* consists of the graphemes *ch*, *a*, *m*, and *p*).

Considerable evidence supports the DRC perspective. For example, consider the performance by skilled readers in naming irregular words and nonwords. An irregular word (e.g., *pint*) has a pronunciation that violates GPC rules. Applying GPC rules to *pint* would yield a pronunciation that rhymes with *mint*. Therefore, a correct reading of *pint* appears to require the lexical route. In contrast, a correct reading of nonwords (e.g., *blask*) requires the sublexical route because nonwords are not represented in the lexicon. In addition, there is evidence from different types of acquired dyslexia. Specifically, individuals with surface dyslexia (e.g., Patterson and Behrmann, 1997) are relatively good at pronouncing nonwords but have difficulty naming low-frequency irregular words, often regularizing these items. These individuals apparently have an intact sublexical route but a disabled lexical route. In contrast, individuals with phonological dyslexia (e.g., Funnell, 1983) have difficulty naming nonwords, but are relatively accurate at naming both irregular and regular words. These individuals apparently have an intact lexical route but a disabled sublexical route. This double dissociation between surface and phonological dyslexics originally was viewed as strong evidence for a dual-route model (but see Patterson et al., 1996; Plaut, 1999 for a more recent discussion of this evidence).

The idea that two routes are necessary to pronounce both irregular words and nonwords was brought into question in 1989, when Seidenberg and McClelland introduced their parallel distributed processing (PDP) model of word recognition. The *parallel* feature of the model means that the processing of different units at a given level occurs simultaneously. The *distributed* feature of the model means that each unique word is associated with a unique pattern of activation across a common set of units that are used to process all words. The Seidenberg and McClelland model consisted of a network of simple processing units, including an orthographic input layer and a phonological output layer. All of these input and output units were connected to all the units of a hidden layer. An important initial advantage of these models is that instead of hard wiring the models to capture

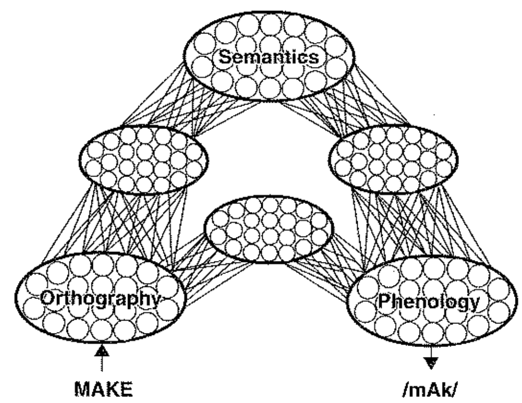

Figure 8.3. The parallel-distributed-processing (PDP) model of Plaut et al. (1996). Reprinted from *Brain and Language*, 52, by Plaut, "Relearning after damage in connectionist networks: Toward a theory of rehabilitation" (1996) with permission from Elsevier.

behavior, the models actually learn via a back propogation algorithm. Specifically, the weights connecting different units are adjusted via exposure to the language in a frequency dependent manner. Weights are adjusted such that the error prone output from the model early during the development is gradually adjusted so that it is more likely to match the desired output in the future. The interesting observation is that the Seidenberg and McClelland model consisted of a single route (from orthography to phonology) that could generate correct pronunciations for both irregular words and nonwords without a lexicon or a set of sublexical rules. However, due to some difficulties with nonword generalization (c.f. Besner et al., 1990) and advances involving recurrent networks (i.e., networks that produce phonological outputs over time), the model was modified by Plaut and colleagues (Plaut et al., 1996) and remains the foremost challenger to the traditional dual-route approach (Figure 8.3).

The Plaut model consists of sets of grapheme units, phoneme units, and semantic units. A layer of hidden units mediates associations between each level of representation, and hence the Plaut model is considerably more complex than the original Seidenberg and McClelland model. As noted, the Plaut model was a recurrent network that eventually settled into a steady state and hence predicted response latencies directly, instead of producing an error measure, which was the output from the original Seidenberg and McClelland model. When recognizing a word, its corresponding grapheme units become activated, and this activation is propagated throughout the network. Knowledge of spelling-to-sound relationships are again contained in the values of weighted connections linking units in the network. The Plaut model demonstrated how to produce codes for regular and irregular words as well as nonwords and eliminated some of the apparent deficiencies with the Seidenberg and McClelland model.

There remains considerable controversy between the connectionist and dual-route models of word recognition (see, for example, Coltheart et al., 2001; Seidenberg, 2005 for recent discussions). In addition, additional models have come online such as the Zorzi, Houghton, and Butterworth (1998) model which combines aspects of the DRC and PDP models. A multiple readout model by Jacobs and colleagues (1998) is also quite important because it emphasizes both task-specific operations and task-general operations in lexical processing. The Ans, Carbonnel, and Valdois (1998) model also has extended this area by providing a computational approach to recognizing multisyllabic words, a deficiency in the previous models dedicated to processing monosyllabic stimuli. Norris (2006) has recently developed a model based on Baysean principles taking into account the probability that a stimulus maps onto a given word given prior probabilities. Also, Perry, Ziegler, and Zorzi (2007) have developed a connectionist model that retains the positive features of earlier models while eliminating many of their weaknesses. As one can see, there is a rich set of theoretical constructs used to capture visual word recognition. Now that we have provided an introduction to the some of the key theoretical issues, we shall turn to the empirical findings that models will need to capture.

4 What variables have been uncovered?

4.1 *The frequency effect*

One of the most robust findings in the literature is that high-frequency words (e.g. *book*) are recognized more quickly and accurately than low-frequency words (e.g. *boom*). In fact, in the large-scale study conducted by Balota and colleagues (2004), word frequency was one of the strongest predictors of performance. In this study, while the word frequency effect was strong for both the naming and lexical decision tasks, the effect was much larger in the lexical decision task.

As noted earlier, although the word frequency effect would appear to be a finding easy to accommodate in models, each model of visual word recognition appears to take a different approach, including thresholds (e.g., Coltheart et al., 2001), weights of connections (e.g., Seidenberg and McClelland, 1989), and locations in frequency-ordered search bins (Murray and Forster, 2004). Others have attempted to argue that task-specific operations contribute to the word frequency effect. For example, consider the lexical decision task. Low-frequency words are more similar to the nonwords than high-frequency words on a familiarity dimension and so are more difficult to accept as a word. Balota and Chumbley (1984) and Balota and Spieler (1999) have argued that this difficulty engages additional analytic processing (also see Besner, 1983), whereas Ratcliff and colleagues (Ratcliff et al., 2004) have recently argued that this will slow the drift rate in a diffusion model. Clearly, the word frequency effect is a classic example of an intuitively simple effect that has been central to theoretical developments in the word recognition literature, and remains a central focus of research.

4.1.1 FAMILIARITY AND SUBJECTIVE FREQUENCY

While objective frequency counts provide good estimates of the frequency of occurrence of words in print, another measure is to have participants rate the subjective familiarity of the stimulus on a numeric scale ranging, for example, from one indicating no familiarity to seven indicating extremely high familiarity. While subjective frequency would be expected to relate to theoretical models in much the same way as objective frequency, this measure may be better because the standard printed frequency counts (e.g., Kučera and Francis, 1967) do not take into consideration spoken word frequency or how often one produces a word through speech or writing. Gernsbacher (1984) argued that objective word frequency estimates are less reliable for low-frequency words than high-frequency words. She noted that *boxer*, *icing*, and *joker* have the same objective frequency value (according to Kučera and Francis, 1967) as *loire*, *gnome*, and *assay*. Fortunately, there are now more extensive word frequency databases than the Kučera and Francis norms (e.g., Baayen, Piepenbrock, and Rijn, 1993; Zeno et al., 1995). However, it is still the case that these researchers typically use the frequency of words in print as their primary measure of frequency. Hence, some researchers still argue that subjective familiarity ratings are a better measure of sheer exposure to a word.

However, familiarity is difficult to define, and familiarity ratings may be influenced by extraneous variables. Standard instructions for familiarity ratings tend to be vague and may encourage the use of other types of information. In fact, Balota, Pilotti, and Cortese (1999) 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.'s (1999) participants rated 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. They found that these ratings were less influenced by meaningfulness than the Toglia and Battig (1978) familiarity ratings. Therefore subjective frequency ratings may be more appropriate than traditional familiarity ratings because they are less influenced by semantic factors. In a recent study, Balota et al. (2004) found that the subjective ratings were predictive of lexical and naming performance above and beyond objective word frequency, length, neighborhood size, spelling-to-sound consistency, and so forth.

4.2 *Age of acquisition*

Recently, researchers have been concerned with the degree to which the age that one acquires a word is related to performance. A number of reports claim that age of acquisition (AoA) influences word recognition performance (e.g., Brown and Watson, 1987; Cortese and Khanna, 2007; Monaghan and Ellis, 2002; Morrison and Ellis, 1995). The intriguing argument here is that early acquired words might provide a special role in laying down the initial representations that the rest of the lexicon is built upon (e.g., Steyvers and Tennenbaum, 2005). 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 (for a review, see Juhasz, 2005). 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. Therefore, it may prove difficult to tease apart these correlated factors. The second issue is whether or not AoA should be considered an outcome variable (Zevin and 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. They focus on frequency trajectory. *Frequency trajectory* 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 (i.e., the sum of frequency over time) in naming. They found little evidence for frequency trajectory whereas cumulative frequency had a marked effect on performance.

4.3 *Orthographic length*

Effects of orthographic length have proven to be theoretically important (see Coltheart et al., 2001). For example, in naming, Weekes (1997) reported that nonwords produced a large length effect whereas words did not. Balota et al. (2004) found evidence that there was a much larger length effect for low-frequency words than high-frequency words. Note that although Weekes did not find an effect of length for words, the pattern reported by Balota et al. is consistent with the pattern (albeit nonsignificant) found by Weekes. Interestingly, individuals with semantic dementia show exaggerated length effects compared to healthy controls for regular consistent words (Gold et al., 2005).

These findings are important for two reasons. First, the DRC model predicts the length by lexicality interaction reported by Weekes whereas the PDP model has difficulty accounting for this result. The sublexical route that is mainly responsible for nonword pronunciation is a serial processor, and the lexical route that is mainly responsible for word processing is a parallel processor. Hence, length effects should be larger for nonwords that rely on the sublexical route and also for individuals with semantic dementia who rely more on the sublexical route to name words aloud, because their semantic/lexical route is impaired. In contrast, the PDP model processes both words and nonwords via the same parallel architecture. In order to account for greater length effects in nonwords, one must posit that the window available for parallel processing is somehow smaller in nonwords than words or that each letter in a nonword requires more computational resources than each letter in a word. Interestingly, New et al. (2006) recently reported an analysis on a large data set and found a quadratic relation between length and lexical decision latencies, that is, short words produced a negative correlation between length and lexical decision latencies, whereas long words produced a positive correlation. This pattern may in part be due to a preferred lexical window size based on the most common length of the words readers experience, which are of moderate length.

4.4 *Regularity and consistency*

In many studies, people are slower and less accurate to name irregular words than regular words (e.g., Baron and Strawson, 1976; Gough and Cosky, 1977). As noted, an *irregular word* can be defined as one whose pronunciation violates GPC rules (e.g., *pint*). In the DRC model, when reading an irregular word, the lexical and sublexical routes produce conflicting information to the phonemic output system, that is, the lexical route produces the correct pronunciation for *pint*, and the sublexical route produces the regularized pronunciation (rhymes with *hint*). Hence, there is either a slowdown in response latency or an increase in error rates. In contrast, when processing a regular word (e.g., *punt*), each route produces the same output such that a quick and accurate pronunciation can be made.

In PDP models, regularity effects result from the adjustment of weighted connections during learning. For example, *int* in *mint*, *tint*, *hint*, and so forth, is pronounced /Int/. Therefore, in these words, weights are adjusted so that *int* yields /Int/. However, when exposed to *pint*, weight changes occur that lead to the /aint/ pronunciation. Although *pint* will be learned, the connections will be weaker than for a regular word (e.g., *punt*), and these weaker connections produce a slower reaction time.

Note, however, that the word *pint* is irregular at two levels. First, it can be considered irregular because it violates GPC rules (i.e., /I/ is usually pronounced as in *stick*, *lid*, and *dish*). Second, it is irregular (i.e., inconsistent) at the rime level because all other words with the *int* rime, pronounce it as /Int/. Many irregular words are (rime) inconsistent which has led to a confounding of these two variables (i.e., GPC regularity and rime consistency), but they are separable dimensions.

Several studies have demonstrated an effect of rime consistency that is independent of GPC regularity (e.g., Glushko, 1979; Jared, McRae, and Seidenberg, 1990). These studies generally examined regular words (defined by GPC rules) containing consistent (e.g., *spoon*) and inconsistent rimes (e.g., *spook* is inconsistent because of *book*, *took*, etc.). A number of other studies have distinguished regularity from consistency by crossing the two factors factorially (e.g., Andrews, 1982; Cortese and Simpson, 2000; Jared, 1997, 2000). These studies are important because the DRC and the PDP models make contrasting predictions regarding the relative influence of these two factors. The DRC model predicts a large effect of regularity and a small effect of consistency. In contrast, the PDP model predicts a large effect of consistency and a small effect of regularity. The results of these studies have generally found that, in words, rime consistency has a larger influence on latencies and errors than GPC regularity. The PDP model simulates these results well whereas the DRC model does not. Moreover, in studies employing many words, Treiman and colleagues (Kessler, Treiman, and Mullennix, 2002; Treiman et al., 1995) have found that consistency at the rime level is a better predictor of naming performance than consistency at the grapheme-to-phoneme level. Interestingly, Andrews and Scarratt (1998) found that nonword reading is more affected by grapheme-to-phoneme consistency than by rime-level consistency. It is quite possible that subjects may rely on different types of information when pronouncing a set of nonwords. In any case, the procedures used to pronounce nonwords may ultimately be quite useful to the understanding of how subjects bring to bear spelling to sound correspondences stored in the lexicon to name novel stimuli. However, it should be noted that Zevin and Seidenberg (2005) have recently argued that the consistency effects reported in nonword naming tasks appear to be more consistent with the PDP perspective than the DRC perspective.

4.4.1 POSITION OF IRREGULARITY EFFECT

Notice that 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*). Interestingly, contemporary models make different predictions about the position of irregularity effect. According to the DRC model, words with early GPC violations are more prone to sublexical interference than words containing later inconsistencies/irregularities because the sublexical route is a serial processor. In contrast, the PDP model processes words in parallel, and so does not predict a position of irregularity effect.

A number of studies have reported a position of irregularity/inconsistency effect (e.g., Coltheart and Rastle, 1994; Cortese, 1998; Rastle and Coltheart, 1999). Although some of these studies have been criticized on methodological grounds (c.f. Cortese, 1998; Zorzi, 2000) the effect appears to be real. That is, latencies are longer for words containing early inconsistencies than words containing later inconsistencies. These results appear to support the DRC model and appear to be more problematic for the PDP model.

4.4.2 FEEDBACK CONSISTENCY EFFECTS

Heretofore, when we have been considering regularity and consistency effects, we have been considering feedforward effects from the orthography of the word to the phonology. For example, *pint* is inconsistent because most words with the *int* orthography produce phonologies that rhyme with *hint*. *Feedback consistency* refers to the likelihood of a given phonological form being spelled in a given manner. For example, the rime in the word *tone* is feedback inconsistent because the /on/ phonological pattern can also be produced via the spelling patterns *own* as in *grown*, and *oan* as in *moan*. As one might guess, many words are both feedforward inconsistent and feedback inconsistent. Stone, Vanhoy, and Van Orden (1997) were the first to decouple feedforward consistency from feedback consistency and observed effects of both variables in lexical decision performance. In addition, Balota et al. (2004) found reliable and equivalent effects of feedback consistency in both lexical decision and naming performance, whereas Ziegler, Montant, and Jacobs (1997) found for French stimuli that the feedback consistency effects were larger in lexical decision than in naming performance. The influence of feedback consistency is quite important theoretically because it suggests a type of resonance in route to the response such that consistent phonological forms provide feedback onto the orthographic patterns (see for example, Pexman, Lupker, and Jared, 2001). However, this area is still controversial, since Peereman, Content, and Bonin (1998) have argued that the feedback consistency effects in French are eliminated when familiarity is controlled (also see Kessler, Treiman, and Mullenix, 2007).

4.5 *Orthographic neighborhood size*

Neighborhood size (i.e., N) refers to the number of words that can be derived from a target word by changing one letter while preserving the other letters and their positions in the word (see Coltheart et al., 1977). For example, *back* has the neighbors *sack, buck, bask,* and so forth. In the DRC model (Coltheart et al., 2001), inhibitory connections between word-level representations inhibit words with large neighborhoods while facilitatory connections between the lexicon and the phonemic output system and between the lexicon and the letter unit input system facilitate responses to words with large neighborhoods. Therefore, the effect of N will depend on the actual parameter settings in the model (c.f. Coltheart et al., 2001). In PDP models, because the same representations are used to process all words, network connections will be strong for words sharing similar representations. Therefore, when words are consistent with regard to their spelling-to-sound correspondences, there tends to be less error in orthographic and phonological systems for high-N words than low-N words, and this characteristic facilitates responses (Sears, Hino, and Lupker, 1999).

Coltheart et al. (1977) were the first to examine this factor, and since that seminal study, N has been at the focus of considerable research (for a review see Andrews, 1997). Although the results have been somewhat mixed, a few conclusions can be made. First, there is typically a facilitative effect of N on response latencies that tends to be relatively larger in naming than lexical decision (Balota et al., 2004). Second, increasing N in nonwords increases lexical decision latencies (Coltheart et al., 1977). Third, the list context modulates the effect of N. For example, Johnson and Pugh (1994) found facilitatory effects of N when illegal nonwords served as distracters and inhibitory effects when legal nonwords are used. Therefore, as Andrews notes, it is important to consider task-specific characteristics when interpreting effects of N. Fourth, in naming, N interacts with word frequency such that low-frequency words produce larger effects of N than high-frequency words (Andrews, 1992). Finally, in lexical decision, Balota et al. (2004) found that for younger adults N also interacted with frequency such that having many neighbors facilitated lexical decision latencies for low-frequency words and inhibited decision latencies for high-frequency words. Within a DRC framework, one might argue that high-frequency words are more sensitive to the lexical-level inhibition, whereas low-frequency words are more sensitive to the sublexical facilitation. Clearly further work is needed to understand the complex pattern of N effects across tasks, list contexts, and other variables such as word frequency.

4.6 *Phonological neighborhood size*

Interestingly, recent work by Yates and colleagues (Yates, Locker, and Simpson, 2004) indicates that phonological neighborhood size (i.e., the number of words that can be constructed from a target word by changing one phoneme while preserving the other phonemes and their positions in the word) also facilitates lexical decision performance independently of orthographic neighborhood size. Yates et al. note that in previous work on orthographic neighborhood effects, orthographic neighborhood size is often confounded with phonological neighborhood size. This finding suggests that word recognition models need to accommodate early influences of phonology on recognition (also see Ziegler and Perry, 1998). In addition, these findings along with those reported originally by Abramson and Goldinger (1997) and more recently by Lukatela and colleagues (Lukatela et al., 2004) suggest that visual and auditory word recognition may engage the same phonological processes. Specifically, Lukatela et al. found that lexical decision times were longer for words that, when spoken, produce a long vowel (e.g., *plead*) than those associated with a short vowel length (e.g., *pleat*).

4.7 *Morphological decomposition*

For a more thorough discussion of morphological processing, see chapters in this volume by Diependaele, Grainger and Sandra, and Woolams and Patterson. However, because of the importance this topic has for word recognition, we briefly review some of the literature here as well.

Traditionally, the *morpheme* has referred to the basic unit of meaning in a language (Hockett, 1966). Many words are made up of more than one morpheme. For example, *jumped* consists of the free morpheme *jump* and the bound morpheme *ed*. Taft and Forster (1975; 1976) proposed that readers decompose words into their constituent morphemes when recognizing them. Root morphemes are then used to access their polymorphemic relatives. Evidence for this perspective comes from studies reporting an effect of root frequency when the overall frequency of words has been controlled (e.g., Taft, 1979a; 1979b). In addition, equivalent long-term priming of roots (e.g., *jump*) from relatives (e.g., *jumped*) and the roots themselves (e.g., *jump*, Stanners et al., 1979) suggests that the root has been accessed during the processing of the more complex relative. It is important to note that these morphemic effects are not due to letter overlap (e.g., Lima, 1987). For example, Lima (1987) reported that while *arson* does not facilitate the recognition of *son*, *dishonest* does facilitate the recognition of *honest*.

In years, research on morphological decomposition has taken on new theoretical significance. This is, in part, due to the fact that the PDP perspective has emerged as a viable general theory of language processing and because PDP models do not possess distinct morphemic representations (e.g., Plaut and Gonnerman, 2000; Rueckl et al., 1997). According to the PDP perspective, morphemic effects emerge from interactions among orthography, phonology, and semantics (Gonnerman, Seidenberg, and Anderson, 2007). Support for this view comes from a recent cross-modal lexical decision study by Gonnerman et al. (2007), who found that the degree of facilitation for visually presented targets was a function of the semantic and phonological overlap found in prime–target pairs regardless of morphemic overlap. For example, *sneer* was just as effective of a prime for *snarl* as *teacher* was for *teach*. In contrast, pairs of items that were more weakly related (e.g., *lately–late*) produced less facilitation.

However, a study by Rastle, Davis, and New (2004) suggests that morphological decomposition does not rely on semantic relationships. In their lexical decision study, target words were preceded by a briefly presented (forty-two ms) masked prime that maintained both a semantic and morphological relationship with the target (e.g., *cleaner–clean*), an apparent morphological relationship only (e.g., *corner–corn*), and a nonmorphological relationship (e.g., *brothel–broth*). Rastle et al. found equivalent priming for targets preceded by primes that appeared to have a morphological relationship with the target regardless of the semantic relationship. Thus, it appears that decomposition is not dependent on 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).

4.8 *Pseudohomophone effects*

Pseudohomophones are nonwords that are homophonic with real words (e.g., *brane*). Pseudohomophones are an important stimulus tool because they allow researchers to study the influence of phonology in accessing meaning. With words, meaning can theoretically be accessed via orthography or phonology, but this is probably not the case with pseudohomophones. That is, upon encountering *brane*, there is a high probability that the reader has not seen this letter sequence before. Therefore, if subjects are faster at naming *brane* than *brone* (e.g., McCann and Besner, 1987) or slower at rejecting *brane* than *brone* in lexical decision (e.g., Rubenstein, Lewis, and Rubenstein, 1971), then there is evidence that meaning has been accessed via phonology. Experiments on pseudohomophones have yielded exactly these results, along with a number of additional findings that have important implications for word recognition models (see Reynolds and Besner, 2005 for a review).

In the DRC model, a pseudohomophone will activate a lexical representation due to interactive activation from the phonemic output system via the sublexical route. If the response is naming, the lexical representation reinforces the phonemic output (thus facilitating the response), and if the response is a lexical decision, the latency is increased due to the increased lexical activation. In PDP models (e.g., Harm and Seidenberg, 2004), pseudohomophones activate semantics whereas regular nonwords do not. The activation of semantics will facilitate a naming response by providing additional input into phonology and inhibit a lexical decision because meaning cannot be used to distinguish the pseudohomophone from a word.

4.9 *Semantic characteristics of the word*

For a more thorough discussion of semantic memory, see the chapter by Cree and Armstrong in this volume. At the onset, one should note that semantic effects are larger in lexical decision than naming (Balota et al., 2004). When one considers the nature of the tasks, this finding is not surprising. In lexical decision, because the task is to distinguish between meaningful word stimuli and less meaningful nonword stimuli, participants tend to direct attention to meaning-level information. In contrast, naming only requires one to convert print into phonology, thus meaning is not required to perform the task. Because naming does not require the access of meaning, finding a true semantic effect provides evidence of an interactive word recognition system.

One variable that has sparked interest in the field is *imageability*. Although imageability effects are larger in lexical decision than naming, there are reports of imageability effects in naming. For example, Strain, Patterson, and Seidenberg (1995) crossed imageability and spelling-to-sound consistency in a word naming study for low-frequency words (see Experiment 2). They found that the consistency effect was reduced for words that had highly imageable referents (e.g., *comb* versus *corpse*) compared to words with lowly imageable referents (e.g., *caste* versus *clause*). Recently, however, some researchers have argued that this imageablity effect was confounded with age of acquisition (i.e., AoA or the age at which a person acquires a word, see Monaghan and Ellis, 2002).

In a recent study of lexical decision and naming performance for 2,428 monosyllabic words, Balota and colleagues (2004) found evidence for semantic influences on performance. The basic finding that semantic factors influence lexical decision performance more than naming was clear. However, semantic factors such as imageability and connectivity (i.e., the degree of semantic clustering between a word and other words) were shown to influence naming above and beyond standard lexical and sublexical variables. We note, however, that due to a lack of information regarding AoA for these words, AoA was not included as a predictor variable. Subsequent analyses were performed by Cortese and Khanna (2007) on the same data set. When AoA was included as a predictor variable, imageability no longer accounted for unique variance in naming latencies. Specifically, AoA accounted for unique variance in reaction times for both naming and lexical decision, whereas imageability's influence was limited to lexical decision. It is entirely possible that imageability effects could be due to AoA (or alternatively trajectory frequency, e.g., Zevin and Seidenberg, 2002; 2004).

One intriguing idea about semantic structure that may relate to word recognition performance is the small-world structure described by Steyvers and Tenenbaum (2005; also see Buchanan, Westbury, and Burgess, 2001 for a similar approach). According to Steyvers and Tenenbaum, a relatively small set of concepts serves as a communication hub 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 (e.g., Nelson, McEvoy, and Schreiber, 1998) with other words may be processed more quickly than words characterized by sparse connections. In the analyses conducted by Balota et al., connectivity as defined by Nelson et al. was, indeed, related to performance in both naming and lexical decision (albeit more in lexical decision) above and beyond standard sublexical and lexical variables.

5 Priming/context effects

Heretofore, we have focused on isolated word recognition. However, words are most commonly processed in the context of other words. Although there are separate chapters in this book devoted to sentence processing, we will briefly describe the work that has employed the priming paradigm. In this paradigm, a prime word is presented and followed by a target word that is responded to. Varying the relationship between the prime and target has been instrumental in demonstrating the types of codes activated by the prime used in route to lexical access. For example, the prime and target may be orthographically related (*couch–touch*), phonologically related (*much–touch*), or semantically related (*feel–touch*). Because of space limitations, we only touch upon some of the major themes in this area. For a more detailed discussion of this literature, see Neely (1991), Hutchison (2004), McNamara (2005), and Kinoshita and Lupker (2003).

5.1 *Orthographic priming effects*

One approach to identifying the access code in word recognition is the masked orthographic priming paradigm developed by Evett and Humphreys (1981, also see Forster, Mohan, and Hector, 2003; Humphreys, Besner, and Quinlan, 1998; and Ziegler, Ferrand, and Jacobs, 2000). In this paradigm, subjects are briefly presented two letter strings that are preceded and/or followed by pattern masks. Subjects typically are unable to consciously identify the primes, and hence these effects reflect early access processes. 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 the original Evett and Humphreys study: First, subjects were better at identifying the second letter string when it shared letters with the first letter string even though these shared letters appeared in different cases (e.g., *lert–lost*). Second, this effect occurred even when the prime items were nonwords, but only when the naming task is employed (Sereno, 1991); one finds little evidence of masked priming for nonwords in the lexical decision task (Forster, 1987). Furthermore, eyetracking studies by Rayner, McConkie, and Zola (1980) have also provided compelling evidence that case independent orthographic codes can be used to access words in the parafovea while reading (see Balota and Rayner, 1991 for a review).

5.2 *Phonological priming studies*

There has been some debate concerning the mandatory role of phonological codes in word recognition (see Frost, 1998 for an excellent review). In an early study, Evett and Humphreys (1981) used the masked priming paradigm and found priming for pairs that were orthographically and phonologically related (e.g., *bribe–tribe*) compared to pairs that were orthographically related but phonologically unrelated (e.g., *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 (e.g., *trail*). More recently, in a masked priming study conducted in Spanish, Pollatsek, Perea, and Carreiras (2005) found that lexical decisions to targets were facilitated by phonologically consistent primes when as little as sixty-six ms separated the onset of the prime from the onset of the target. Note also that there is evidence of phonological priming in the parafoveal priming paradigm, examining eye movements during the reading of text. Specifically, Pollatsek et al. (1992) found that previews that were homophonic with targets (e.g., *site–cite*) facilitated performance (both in pronunciation latencies and fixation durations) compared to nonhomophonic previews controlled for orthographic similarity (e.g., *cake–sake*). Again, this pattern would appear to support a role for phonology as an access code (also see Lee et al., 1999).

5.3 *Semantic priming effects*

The semantic (associative) priming paradigm has been thoroughly investigated, and began with a seminal study by Meyer and Schvaneveldt (1971). They found that subjects were faster to make lexical decisions to pairs of related words (e.g., *cat–dog*) than pairs of unrelated words (e.g., *pen–dog*). This robust effect appeared 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 appear to be critical to higher level comprehension. We note that semantic/associative priming effects occur not only in standard lexical decision and naming tasks (see Hutchison et al., 2007, for a recent large-scale study), but they also occur cross-modally (i.e., when an auditory prime precedes a visually presented target; cf. Holcomb and Anderson, 1993)

Semantic priming has been a topic of rich empirical and theoretical debate. For example, one might ask if the effect is truly "semantic," that is, reflects similarity in semantic features or category membership, such as *dog* and *cat*, or if it primarily reflects associative relationships among items (e.g., *rat* and *cheese*). Two recent reviews of this topic appear to reach quite different conclusions. Lucas (2000) indicated that there were indeed semantic effects in priming, but Hutchison (2004) concluded that a simple associative account could handle much of this literature. One of the findings that Hutchison focuses on is mediated priming (e.g., Balota and Lorch, 1986; McNamara and Altarriba, 1988). *Mediated priming* refers to priming across intervening nonpresented concepts, that is, from *lion* to *stripes*. Of course, there is very little semantic overlap between *lion* and *stripes*, but there is an associative relationship via *lion* to *tiger* to *stripes*.

A second area where semantic priming has been central concerns *threshold priming*. In threshold priming experiments, the prime item is presented so briefly and patterned masked that subjects are presumably unaware of its presence. Initial experiments reported semantic priming effects under conditions in which subjects apparently can no longer make presence/absence decisions about the prime item (e.g., Balota, 1983; Fowler et al., 1981; Marcel, 1983). This initial work indeed was criticized because of the threshold setting procedures (see Cheesman and Merikle, 1984; Holender, 1986; Merikle, 1982), and the nature of an objective identification threshold still is debated today

(Dosher, 1998; Greenwald and Draine, 1998). The important point here is that one can obtain semantic priming effects under highly degraded situations. This paradigm has been extended to cognitive neuroscience domains (see Dehaene et al. 2005) and social psychology (see Ric, 2004).

A third area of interest regarding priming effects is *backward priming*. There are two types of backward priming effects. First, priming has been reported when the prime is presented after the target (see Balota, Boland, and Shields, 1989; Kiger and Glass, 1983). In this type of experiment, the target that requires a response appears prior to the prime, and then the prime appears afterward, but prior to the response. This finding falls naturally from a cascadic framework in which partial activation is released from representations before such representations have reached threshold (see earlier discussion of the McClelland and Rumelhart model). A second type of backward priming entails direction relations; for example, one finds priming from *boy to bell* in the lexical decision task (c.f. Koriat, 1981; Seidenberg et al., 1984), but not typically in the pronunciation task, at least at long (Stimulus onset synchrony) SOAs. This would appear to support the notion that subjects check back from the target to the prime to bias their lexical decisions. Interestingly, one can find this type of backward priming for noncompounds (e.g., *baby–stork*) at short SOAs (see Kahan, Neely, and Forsythe, 1999) even in naming. This may actually be related to the first type of backward priming, suggesting that when there is close temporal proximity in the presentation of the prime and target, both forward and backward priming effects can be observed, even in a task that does not encourage backward checking such as speeded naming performance (see Hutchison, 2004 for further discussion).

Regarding the theoretical developments, we will only list a few of the mechanisms used to accommodate semantic priming. One of the most popular mechanisms still is some variant of spreading activation theory. 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 and Loftus, 1975). When a node in memory becomes activated via stimulus presentation or via internal direction of attention, activation spreads from that node along associative pathways to nearby nodes. Thus, the reason that subjects are faster to recognize *dog* when it follows *cat*, compared to when it follows *pen*, is because the underlying representation for these two words is 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* has been preactivated and hence needs less stimulus information to surpass threshold.

One of the most compelling studies in support of an automatic spreading activation mechanism comes from Neely's (1977) dissertation. In this study, participants were given category primes and instructions regarding what to expect when a given category was presented. For some categories, subjects were told to expect category exemplars designated by that category (i.e., when you receive the category *bird*, think of *birds*), however, for other categories, subjects received instructions to shift from the designated category to a new category (i.e., when you receive the prime *body*, think of types of building parts). Neely also manipulated the time to process the prime before the target was presented. Amazingly, Neely found full-blown semantic priming at the short prime–target interval that was totally independent of the instructions, that is, there was equal priming from *bird* to *robin* and *body* to *arm* in the previous examples. However, at the long prime–target interval, priming reflected only the subjects' expectancies. For example, Neely found equivalent priming for *bird* to *robin* and *body* to *door*, and the priming observed for *body* to *arm* was equivalent to a totally unrelated condition such as *body* to *maple*. Neely claimed that automatic spreading activation produced priming at the short prime–target interval, and the long prime–target interval reflected a second independent attentional mechanism (also see, Balota, Black, and Cheney, 1992; Favreau and Segalowitz, 1983).

In addition to distinctions between automatic and attentional mechanisms underlying semantic priming effects, there have been many attempts to model such effects (e.g., Becker, 1980; Forster, 1981; Norris, 1986). Most recently, researchers have developed computational models of semantic priming. The problem here is to specify the nature of the underlying semantic/associative representations. One approach has been to model priming in terms of featural overlap between the meanings of the primes and targets (c.f. Cree, McRae, and McNorgan, 1999) or by a temporal contiguity between semantic features of the prime and phonological features of the target (Masson, 1995; Plaut and Booth, 2000). The notion is that when the prime is presented, a set of distributed features is activated, and the extent to which these features overlap with the target modulates the observed priming effects. An alternative approach is Ratcliff and McKoon's (1988) compound cue model. The notion here is that the prime and target serve as a compound cue that is compared to traces already stored in memory, with related cues producing higher familiarity values than unrelated cues. Although each of these models has intriguing components, it is still likely that no single model of priming will be able to handle the rich diversity of this literature, and as Neely (1991) has argued, it is likely that multiple mechanisms will need to be postulated to account for the breadth of priming effects.

6 Recent methodological developments for constraining theories

6.1 *Factorial designs versus large-scale item analyses*

Historically, researchers have 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 potential limitations to this approach (for a discussion of these issues see Balota et al., 2004; Culter, 1981). More recently, researchers have examined word recognition performance for a large set of words (Balota and Spieler, 1998; Besner and Bourassa, 1995; Kessler, Trieman, and Mullennix, 2002; Spieler and Balota, 1997; Trieman et al., 1995). As noted, Balota et al. (2004) examined lexical decision and naming performance in younger and older adults for 2,428 words. Multiple regression techniques were utilized in order to obtain estimates of the unique variance attributable to a set of predictor variables, and these researchers were able to account for .49 of the variance in lexical decision performance and .50 of the variance in naming performance. This is a multifold increase over current computational models (see Balota and Spieler, 1998 and Seidenberg and Plaut, 1998 for a discussion of the pros and cons of this comparison). This outcome was obtained despite the success these computational models have had in accounting for performance at the factor level (but see Perry et al., 2007). 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 forty thousand words (Balota et al., 2007). The English Lexicon Project website (http://elexicon.wustl.edu) provides a comprehensive data set that researchers can access, via a powerful search engine, performance measures and item characteristics.

In addition to these large-scale behavioral databases, there are also large-scale analyses of the contexts in which words occur in natural language databases. An example of this is the work by Steyvers and Tenenbaum (2005) on the small-scale semantic networks described earlier. This work has been recently reviewed by Cree (2005), and is very exciting because it provides a computational approach of

formance have been
oach has been useful,
ial limitations to this
ission of these issues
; Culter, 1981). More
have examined word
nce for a large set of
Spieler, 1998; Besner
Kessler, Trieman, and
ler and Balota, 1997;
). As noted, Balota
lexical decision and
n younger and older
Multiple regression
ed in order to obtain
que variance attrib-
dictor variables, and
able to account for
lexical decision per-
e variance in naming
multifold increase
itional models (see
98 and Seidenberg
scussion of the pros
parison). This out-
lespite the success
nodels have had in
nance at the factor
l., 2007). The large-
es provide another
constraint in the
cal approaches to
recently, Balota and
d naming and lex-
or over forty thou-
2007). The English
e (http://elexicon.
omprehensive data
access, via a pow-
ormance measures

large-scale behav-
re also large-scale
s in which words
age databases. An
work by Steyvers
) on the small-
described earlier.
ently reviewed by
exciting because
onal approach of

grounding semantics, via the contexts in which words co-occur.

6.2 *Distributional analyses*

Typically, in word recognition experiments, one compares the mean response latency across several conditions to determine if the predictions generated by an experimental hypothesis are correct or not. However, researchers have long noted that means of conditions are only one estimate available from performance. For example, in the 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 reaction 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 (see Spieler, Balota, and Faust, 1996 for a replication of this pattern). These researchers have fit reaction time distributions to ex-Gaussian functions, but other functions such as the Weibul or ex-Wald could also accomplish the same goals. As theories become more precise regarding the item-level performance, there should be an increased level of sophistication regarding the predictions concerning the underlying reaction time distributions. Balota and Spieler (1999) found that frequency and repetition influenced these parameters differently depending on the task (however, see Andrews and Heathcote, 2001). Ratcliff et al. (2004) have recently used reaction time distributions to more powerfully test a diffusion model of lexical decision performance. We anticipate that the precision of reaction time distribution analyses will be critical in the discrimination of available models (see Balota et al., 2008 for a review).

6.3 *Neuroimaging*

Models of word recognition have been constrained by findings in the neuropsychological literature. For example, early versions of the dual-route model were designed inductively to accommodate certain dyslexia subtypes. More recent findings in the neuroimaging literature have also been important and will continue to influence future theoretical developments (also see Sandak et al., this volume). Some early findings in this literature include those reported by Petersen and colleagues in a positron emission tomography (PET) study (Petersen et al., 1989). For example, Petersen et al. found that passively viewing words was associated with activation of the occipital lobes, reading words aloud was associated with temporal activation, and generating verbs from nouns was associated with frontal lobe activation. Also in a PET study, Fiez and colleagues (Fiez et al., 1999) linked lexicality, frequency, and spelling-to-sound consistency to specific brain regions. Interestingly, spelling-to-sound consistency and lexicality was associated with the activation of an area in the left frontal lobe, suggesting that this region may be involved in orthographic-to-phonological translation. In addition, the primary motor cortex in both hemispheres was associated with greater activation when processing inconsistent words (e.g., *pint*) than consistent words (e.g., *punt*) suggesting that motor production is affected by consistency. Also, processing low-frequency words was associated with activation of a region in the left temporal lobe and supplementary motor area. Finally, more recent studies (as described by Sandak et al., this volume) have attempted to associate these and other regions to dual-route models and parallel distributed processing models. Sandak and colleagues conclude that activation patterns found in the brain are consistent with an interactive PDP framework that divides labor among various codes. Interestingly, the division of labor within this framework depends on the relative acquisition of skill at different levels.

7 Continuing controversies

7.1 *The magic moment*

A reasonable question that one might ask is what does it mean to recognize a word? In other words, is there a "magic moment" (e.g., Balota, 1990) that corresponds to a discrete point in time when a word has been recognized? The answer to this question may depend on the model of word recognition one uses to explain the process. The magic moment for models that contain discrete lexical entries for words (e.g., Coltheart et al., 2001) would occur when a threshold for identification has been reached. In contrast, in models containing distributed representations (e.g., Plaut et al., 1996; Seidenberg and McClelland, 1989), the magic moment is more difficult to discern. In distributed models, there are no separate lexical representations for each individual word and no threshold to be achieved. Therefore, one might posit that the magic moment occurs when the network settles into a stable pattern of activation and that the subject may modulate what "stable" means depending on task, list context, and other variables, that is, there is no single magic moment. In this light, it may be better to consider what the trigger is for a given response in a given task rather than to speculate about a task independent magic moment.

7.2 *Phonological codes in silent reading*

The extent to which readers use phonological codes to access meaning in silent reading has been somewhat controversial (see Frost, 1998 for a review). Representing one side of the issue are those who argue for the mandatory use of phonology in silent reading (e.g., Frost, 1998; Lukatela and Turvey, 1991; Van Orden, 1987). In contrast, some researchers posit that phonological codes play more or less of a role in performance depending upon the nature of the stimulus and/or the ability of the reader. For example, Jared and Seidenberg (1991) and Jared, Levy, and Rayner (1999) argued that phonological codes are used more for infrequent words, and with sufficient exposure, phonology plays a relatively small role. Rather, semantics can be accessed rather directly from orthography. Thus, all reading researchers agree that phonology is used to access meaning at least on a partial basis, and some claim that phonology is mandatory for reading all words.

The role of phonology in reading is quite critical because it influences how educators teach reading in elementary school (also see Nation, this volume). Some advocates of a whole word approach suggest that in learning to read, children can use context to uncover the meaning and most often the phonology of the visual stimulus. The whole word instructional approach exposes children to whole words, their pronunciations, and meanings, both in and out of sentential and discourse context. In contrast, phonics-based approaches emphasize that knowledge of the relationships between graphemes and phonemes is fundamental to skilled reading. The evidence on this issue indicates that phonics-based instruction is a more effective strategy than whole word techniques that do not include phonics (for reviews see Adams, 1990; Rayner et al., 2001; and Snow, Burns, and Griffin, 1998).

7.3 *Attentional control of processing pathways and time criterion*

One question that has received considerable recent interest is the extent to which the lexical processing system adapts to the current processing demands. For example, one might expect different processing of text when proofreading, comprehending, or checking for grammaticality. Indeed, virtually every theory of word recognition posits multiple ways of accessing or computing the phonological code from print. For example, in dual-route models, 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, phonology can be computed directly from orthography or indirectly via semantics. The issue addressed here concerns the extent to which attention to a processing pathway can be biased by the experimental operations in a given study. For example, are there procedures that will bias the reader to rely more on the lexical or sublexical pathway within a dual-route framework? This is important because it brings into question the modularity of the lexical processing system (see Fodor, 1983). One way to examine this issue is to present words that place different demands on the lexical and sublexical pathways. For example, nonwords should bias the sublexical pathway and low-frequency exception words should bias the lexical pathway, since the sublexical pathway would lead to regularization errors, that is, pronouncing *pint* such that it rhymes with *hint*. In an early study, 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. This supports the notion that the 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 and Coltheart, 1999; Reynolds and Besner, 2005; Simpson and Kang, 1994; Zevin and Balota, 2000).

However, the evidence for route priming has been controversial. Specifically, work by Kinoshita and Lupker (2002; 2003) suggests that much of these findings can be accounted for by a time criterion model. Specifically, there is evidence that participants adopt a time criterion whereby they produce a response at a latency 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 700 ms and a set of high-frequency words produces response latencies on the average of 600 ms. If one now embeds these same words in the context of nonwords that produce an average response latency of 700 ms, the word frequency effect will likely diminish. Specifically, latencies to the low-frequency words will remain the same (because the latencies are quite similar to the nonwords), whereas latencies to the high-frequency words will increase considerably, that is, 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.

Although the evidence suggests that participants do adopt a time criterion based on the difficulty of items within a block, we believe that there is also evidence for pathway control. For example, all of the effects reported by Zevin and Balota (2000) hold even after the response latencies to the context items are partialled out via analyses of covariance. Clearly, however, further work is necessary in this area.

8 Summary

In this chapter, we have described the major tasks employed, the different theoretical perspectives, many of the variables that influence word recognition performance, and some of the continuing controversies. Of course, this overview only provides a glimpse of the vast amount of research that has been conducted on visual word recognition. Although much has been accomplished, there is clearly need for continuing work in clarifying the processes engaged in the seductively simple act of visual word recognition.

References

Abramson, M. & Goldinger, S. (1997). What the reader's eye tells the mind's ear: Silent reading activates inner speech. *Perception & Psychophysics*, 59, 1059–68.
Adams, M. (1990). *Beginning to read*. Cambridge, MA: MIT Press.
Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.

(1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22, 261–95.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition*, 10, 565–75.

(1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 234–54.

(1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review*, 4, 439–61.

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 & Cognition*, 27, 545–55.

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, 1052–86.

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

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

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

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

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

(1994). Visual word recognition: The journey from features to meaning. In Gernsbacher, M. A. (Ed.) *Handbook of psycholinguistics* (pp. 303–48). San Diego: Academic Press.

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, 340–57.

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. http://elexicon.wustl.edu/. Washington University.

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

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, 336–45.

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition (Special Issue)*, 29, 639–47.

Balota, D. A. & Rayner, K. (1991). Word recognition processes in foveal and parafoveal vision: The range of influence of lexical variables. In Besner, D. & Humphreys, G. W. (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*, 8, 238–40.

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

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

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. I., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39, 445–59.

Baron, J. & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 386–93.

Becker, C. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 252–9.

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

(1983). Basic decoding components in reading: Two dissociable feature extraction processes. *Canadian Journal of Psychology*, 37, 429–38.

Besner, D. & Bourassa, D. C. (June 1995). *Localist and parallel processing models of visual word recognition: A few more words*. Paper presented at the annual meeting of the Canadian Brain, Behavior, and Cognitive Science Society, Halifax, Nova Scotia, Canada.

Besner, D., Twilley, L., McCann, R. S., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review*, 97, 432–46.

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

Brown, GDA & Watson, F. L., (1987). First in, first out: Word learning age and spoken word-frequency as predictors of word familiarity and word naming latency. *Memory & Cognition*, 15, 208–16.

Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review*, 8, 531–44.

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, & Cognition*, 23, 857–71.

Catlin, J. (1973). In defense of sophisticated guessing theory. *Psychological Review*, 80, 412–16.

Cheesmann, J. I. & Merikle, P. M. (1984). Priming with and without awareness. *Perception & Psychophysics*, 36, 387–95.

Collins, A. & Loftus, E. (1975). A spreading activation theory of semantic processing. *Psychological Review*, 82, 407–28.

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

Coltheart, M. & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1197–220.

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, 204–56.

Cortese, M. J. (1998). Revisiting serial position effects in reading aloud. *Journal of Memory and Language*, 39, 652–66.

Cortese, M. J. & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical decision performance above and beyond 22 other predictor variables: An analysis of 2340 words. *Quarterly Journal of Experimental Psychology*, 60, 1072–82.

Cortese, M. J. & Simpson, G. B. (2000). Regularity effects in word naming: What are they? *Memory & Cognition*, 28, 1269–76.

Cree, G. S. & Armstrong, B. C. (this volume). Computational models of semantic memory.

Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, 23, 371–414.

Culter, A. (1981). Making up materials is a confounded nuisance: or Will we be able to run any psycholinguistic experiments at all in 1990? *Cognition*, 10, 65–70.

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

Dosher, B. A. (1998). The response-window regression method-Some problematic assumptions: Comment on Draine and Greenwald (1998). *Journal of Experimental Psychology: General*, 127, 311–17.

Draine, S.C. & Grrenwald, A.G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General*, 127, 286–303.

Ellis, A. W. & Monaghan, J. (2002). Reply to Strain, Patterson, and Seidenberg. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 215–20.

Evett, L. J. & Humphreys, G. W., (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology*, 33A, 325–50.

Favreau, M. & Segalowitz, N. S. (1983). Automatic and controlled processes in first- and second-language reading of fluent bilinguals. *Memory & Cognition*, 11, 565–74.

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*, 205–18.

Fine, E. M. (2001). The word superiority effect does not depend on one's ability to identify the letter string as a word. *Journal of Vision*, *1*, 410.

Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology.* Cambridge, Mass.: MIT Press.

Forster, K. I. (1976). Accessing the mental lexicon. In Wales, R. J. & Walker, E. (Eds.) *New Approaches to Language Mechanisms* (pp. 257–87). Amsterdam: North Holland.

(1979). Levels of processing and the structure of the language processor. In Cooper, W. E. & Walker, E. (Eds.) *Sentence processing: Psycholinguistic essays presented to Merrill Garrett.* Hillsdale, NJ: Lawrence Erlbaum Associates.

(1981). Priming and the effects of sentence and lexical context on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology*, *33A*, 465–95.

(1987). Form-priming with masked primes: The best match hypothesis. In Coltheart, M. (Ed.), *Attention and Performance 12: The Psychology of Reading* (pp. 127–46). Hillsdale, NJ: Lawrence Erlbaum Associates.

(2004). Category size effects revisited: Frequency and masked priming effects in semantic categorization. *Brain and Language*, *90*, 276–86.

Forster, K. I., Mohan, K., & Hector, J. (2003). The mechanics of masked priming. In Kinoshita, S. & Lupker, S. J. (Eds.) *Masked priming: State of the art* (pp. 3–17).

Fowler, C. A., Wolford, G., Slade, R., & Tassinary, L. (1981). Lexical access with and without awareness. *Journal of Experimental Psychology: General*, *110*, 341–62.

Frederiksen, J. R. & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception & Performance*, *2*, 361–79.

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

Frost, S. J., Mencl, W. E., Sandak, R., Moore, D. L., Rueckl, J. G., Katz, L., Fulbright, R. K., & Pugh, K. R. (2005). A functional magnetic resonance imaging study of the tradeoff between semantics and phonology in reading aloud. *Neuroreport: For Rapid Communication of Neuroscience Research*, *16*, 621–24.

Funnell, E. (1983). Phonological processes in reading: New evidence from acquired dyslexia. *British Journal of Psychology*, *74*, 159–80.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*, 256–81.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 674–91.

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*, 833–46.

Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General*, *136*, 323–45.

Gough, P. B. & Cosky, M. L. (1977). One second of reading again. In Castellan, J. N., Pisoni, D. B., & Potts, G. R. (Eds.) *Cognitive Theory (vol. 2)*. Hillsdale, NJ: Lawrence Erlbaum.

Grainger, J. & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *103*, 518–65.

Harm M. W., McCandliss, B. D., & Seidenberg, M. S. (2003). Modeling the successes and failures of interventions for disabled readers. *Scientific Studies of Reading*, *7*, 155–82.

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.

Heathcote, A., Popiel, S. J., & Mewhort, D. J., (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*, *109*, 340–7.

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*, 1331–56.

Hockett, C. F. (1966). The problem of universals in language. In Greenberg, J. H. (Ed.) *Universals of Language* (pp. 1–29). Cambridge, MA: MIT Press.

Holcomb, P. J. & Anderson, J. E. (1993). Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes*, 8, 379–411.

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.

Hubel, D. H. & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160, 106–54.

(1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195, 215–43.

Humphreys, G. W., Besner, D., & Quinlan, P. T. (1998). 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 & Cognition*, 10, 576–90.

Hutchison, K. A. (2004). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review*, 10, 785–813.

Jacobs, A. M., Rey, A., Ziegler, J. C., & Grainger, J. (1998). MROM-p: An interactive activation, multiple read-out model of orthographic and phonological processes in visual word recognition. In Grainger, J. & Jacobs, A. M. (Eds.) *Localist connectionist approaches to human cognition* (pp. 147–88). Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30, 513–41.

Jared, D. (1997). Spelling-sound consistency affects the naming of high-frequency words. *Journal of Memory and Language*, 36, 505–29.

(2000). Spelling-sound consistency and regularity effects in word naming. *Journal of Memory and Language*, 46, 723–50.

Jared, D., Levy, B. A., & Rayner K. (1999). The role of phonology in the activation of word meanings during reading: Evidence from proofreading and eye movements. *Journal of Experimental Psychology: General*, 128, 219–64.

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

Jared, D. & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General*, 120, 358–94.

Johnson, N. F. & Pugh, K. R. (1994). A cohort model of visual word recognition. *Cognitive Psychology*, 26, 240–346.

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification, *Psychological Bulletin*, 131, 684–712.

Kahan, T. A., Neely, J. H., & Forsythe, W. J. (1999). Dissociated backward priming effects in lexical decision and pronunciation tasks. *Psychonomic Bulletin & Review*, 6, 105–10.

Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language*, 47, 145–71.

(2007). Feedback consistency effects in single-word reading. In Grigorenko, E. J. & Naples, A. J. (Eds.) *Single-word reading: Cognitive, behavioral, and biological perspectives* (pp. 159–74). Mahwah, NJ: Erlbaum.

Kiger, J. L. & Glass, A. (1983). The facilitation of lexical decisions by a prime occurring after the target. *Memory & Cognition*, 11, 356–65.

Kinoshita, S. & Lupker, S. J. (2002). Effects of filler type in naming: Change in time criterion or attentional control of pathways? *Memory & Cognition*, 30, 1277–87.

(2003). Priming and attentional control of lexical and sublexical pathways in naming: A reevaluation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 405–15.

Kinoshita, S. & Woollams, A. (2002) The masked onset priming effect in naming: Computation of phonology or speech-planning? *Memory & Cognition*, 30, 237–45.

Koriat, A. (1981). Semantic facilitation in lexical decision as a function of prime-target association. *Memory & Cognition*, 9, 587–98.

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

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, 948–64.

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

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

Lukatela, G., Eaton, T., Sabadini, L., & Turvey, M. T. (2004). Vowel duration affects visual word identification: Evidence that the mediating phonology is phonetically informed. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 151–62.

Lukatela, G. & Turvey, M. T. (1991). Phonological access of the lexicon: evidence from associative priming with pseudohomophones. *Journal of Experimental Psychology*, 17, 951–66.

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 570–90.

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

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

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

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 108, 287–330.

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

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York: 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, 545–59.

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

Meyer, D. E. & Schvaneveldt, R. W. (1971). Facilitation in recognizing words: Evidence of a dependence upon retrieval operations. *Journal of Experimental Psychology*, 90, 227–34.

Monaghan, J. & Ellis, A. W. (2002). What exactly interacts with spelling-sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 116–53.

Monsell, S., Patterson, K., Graham, A., Hughes, C. H., & Milroy, R. (1992). Lexical and sub-lexical translations of spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 452–67.

Morrison, C. M. & Ellis, E. 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, 116–53.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76, 165–78.

Murray, W. S. & Forster, K.I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review*, 111, 721–56.

Nation, K. (this volume). Decoding, orthographic learning, and the development of visual word recognition.

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, 226–54.

(1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In Besner, D. & Humphreys, G. W. (Eds.) *Basic processes in reading: Visual word recognition* (pp. 236–64), Hillsdale, NJ: Erlbaum..

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Re-examining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review*, 13, 45–52.

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

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

(2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, 105, 761–81.

Paap, K. R. & Johansen, L. S. (1994). The case of the vanishing frequency effect: A retest of the verification model. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1129–57.

Paap, K. R., Newsome, S. L., & McDonald, J. E. (1982). An activation-verification model for letter and word recognition: The word superiority effect. *Psychological Review*, 89, 573–94.

Patterson, K. & Behrmann, M. (1997). Frequency and consistency effects in a pure surface dyslexic patient. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1217–31.

Patterson, K., Plaut, D. C., McClelland, J. L., Seidenberg, M. S., Behrmann, M., & Hodges, J. R. (1996). Connections and disconnections: A connectionist account of surface dyslexia. In Reggia, J., Berndt, R., & Ruppin, E. (Eds.) *Neural modeling of cognitive and brain disorders* (pp. 177–99). New York: World Scientific.

Peereman, R., Content, A. L., & Bonin, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory & Language*, 39, 151–74.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, 114, 273–315.

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, 153–70.

Pexman, P. M., Lupker, S. J., & Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 139–56.

Plaut, D. C. (1996). Relearning after damage in connectionist networks: Toward a theory of rehabilitation. *Brain and Language*, 52, 25–82.

(1999). Computational modeling of word reading, acquired dyslexia, and remediation. In Klein, R. M. & McMullen P. A. (Eds.) *Converging methods in reading and dyslexia* (pp. 339–72). Cambridge, MA: MIT Press.

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, 445–85.

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

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, 148–62.

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, 557–65.

Pugh, K. R., Sandak, R., Frost, S. J., Moore, D., & Mencl, W. E. (2005). Examining reading development and reading disabilities in English language learners: Potential contributions from functional neuroimaging. *Learning Disabilities Research & Practice*, 20, 24–30.

Rastle, K., Davis, M., & New B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, 11, 1090–8.

Rastle, K. & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 482–503.

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging*, 19, 278–89.

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

Rayner, K. (1998). Eye movements in reading and information processing: Twenty years of research. *Psychological Bulletin*, 124, 372–422.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest Monograph*, 2, 31–74.

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

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology*, 12, 206–26.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 81, 275–80.

Reynolds, M. & Besner, D. (2005). Contextual control over lexical and sublexical routines when reading English aloud. *Psychonomic Bulletin & Review*, 12, 113–18.

(2005). Basic processes in reading: A critical review of psuedohomophone effects in naming and a new computational account. *Psychonomic Bulletin & Review*, 12, 622–46.

Ric, F. (2004). Effects of the activation of affective information on stereotyping: When sadness increases stereotype use. *Personality & Social Psychology Bulletin*, 30, 1310–21.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning & Verbal Behavior*, 9, 487–94.

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

Rueckl, J. G., Mikolinski, M., Raveh, M., Miner, C. S., & Mars, F. (1997). *Journal of Memory and Language*, 36, 382–405.

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

Sandak, R. **et al.** (this volume). How does the brain read words?

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, 26, 1270–81.

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.

(1999). Orthographic neighbourhood effects in parallel distributed processing models. *Canadian Journal of Experimental Psychology*, 53, 220–9.

Seidenberg, M. S. (2005). Connectionist models of word reading. *Current Directions in Psychological Science*, 238–42.

Seidenberg, M. S. & McClelland, J. L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review*, 96, 523–68.

Seidenberg, M. S., Petersen, A., MacDonald, M. C., & Plaut, D. C. (1996). Pseudohomophone effects and models of word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 48–62.

Seidenberg, M. S. & Plaut, D. C. (1998). Evaluating word-reading models at the item level: Matching the grain of theory and data. *Psychological Science*, 9, 234–7.

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

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre- and post-lexical loci of contextual effects on word recognition. *Memory & Cognition*, 12, 315–28.

Selfridge, O. G. (1959). Pandemonium: A paradigm for learning. In *Symposium of the mechanization of thought processes*. Proceedings of a symposium held at the National Physical Laboratory, U.K. Her Majesty's Stationary Office.

Selfridge, O. G. & Neisser, U. (1960). Pattern recognition by machine. *Scientific American*, 203, 60–8.

Sereno, J. (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.

Simpson, G. B. & Kang, H. (1994). The flexible use of phonological information in word recognition in Korean. *Journal of Memory and Language*, 33, 319–31.

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

Spieler, D. H. & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science*, 8, 411–16.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Levels of selective attention revealed through analyses of response time distributions. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 506–26.

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

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.

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, 337–59.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1140–54.

(2002). Theories of word naming interact with spelling-sound consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 207–14.

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

Taft, M. (1979b). Recognition of affixed words and the word frequency effect. *Memory & Cognition*, 7, 263–72.

Taft, M. & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14, 638–47.

(1976). Lexical storage and retrieval of polymorphemic and pollysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, 15, 607–20.

Taft, M. & Hambly, G. (1986). Exploring the Cohort Model of spoken word recognition. *Cognition*, 22, 259–82.

Tan, L. H. & Perfetti, C. A. (1999). Phonological and associative inhibition in the early stages of English word identification: Evidence from backward masking. *Journal of Experiment Psychology: Human Perception and Performance*, 25, 59–69.

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

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 Memory and Language*, 124, 107–36.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 50*A*, 439–56.

Yates, M., Locker, L., & Simpson, G. (2004). The influence of phonological neighborhood on visual word perception. *Psychonomic Bulletin & Review*, 11, 452–7.

Zeno, S. M., Ivenz, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator"s word frequency guide*. Brewster, NY: Touchstone Applied Science Associates.

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, 121–35.

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.

(2004). Age of acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory & Cognition*, 32, 31–38.

(2006). Simulating consistency effects and individual differences in nonword naming: A comparison of current models. *Journal of Memory & Language*, 54, 145–60.

Ziegler, J. C., Ferrand, L., & Jacobs, A. M. (2000). Visual and phonological codes in letter and word recognition: Evidence from incremental priming. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 53*A*, 671–92.

(1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language*, 37, 533–54.

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. (2000). Serial processing in reading aloud: No challenge for a parallel model. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 847–56.

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, 1131–61.