CHAPTER TWO

Attentional control of lexical processing pathways during word recognition and reading

David A. Balota
Department of Psychology, Washington University, USA

Stephen T. Paul
Psychology Department, Mississippi State University, USA

Daniel H. Spieler
Department of Psychology, Stanford University, USA

The focus of the present chapter is on the primary meaning bearing element in reading, i.e. the word. There are certainly many different aspects of words that play crucial roles during word recognition (see Balota, 1994 and Henderson, 1982 for reviews). In the present chapter, we have decided primarily to emphasise research contributing to our understanding of *five* principal factors that have been shown to modulate word identification performance. Specifically, we will review some of the current word recognition research examining the influences of orthography, phonology, and meaning, along with syntactic- and discourse-level context effects. The first three factors are important in that each factor has been shown to affect processing of words both in isolation and in linguistic contexts, and each has yielded impressive amounts of data and controversy. The latter two factors are somewhat different in that discourse-based syntactic and semantic information do not contribute to isolated word recognition, but, of course, are fundamental in our use of words in the vast majority of language processing contexts.

In our discussion of each of the five factors, we will expand on some methodological and stimulus issues that we believe are crucial to our understanding of research in the area. For example, as outlined by Rayner and Pollatsek (1989), there are issues associated with the experimental paradigms involved in word recognition research that may not generalise to normal reading. Obviously, we rarely find ourselves reading with the tachistoscopic staccato frequently employed in word recognition research. Nor is it usual to have each word of a text masked, or presented via rapid serial visual presentation (RSVP) methods. In addition, unless the text is particularly dense with unfamiliar or misspelt words, we don't normally find ourselves performing lexical decisions as we progress through paragraphs of a text in a book. We will emphasise in the present chapter that perhaps the most critical difference between isolated word recognition and normal reading is the direction of attention to relatively distinct processing pathways.

The present chapter involves three sections. In the first section, we shall provide an overview of the processing pathway approach to lexical processing, which will be the unifying theme across aspects of this chapter. The second section will provide a brief overview of some of the theoretical approaches to word recognition and hence provide a foundation for interpreting the empirical literature that will be reviewed. The third and major section will provide a review of the five previously mentioned factors, with special emphasis on the manner in which the processing pathway perspective will help elucidate our understanding of the manner in which these factors influence word processing across a variety of tasks.

ATTENTIONAL SELECTION AND LEXICAL PROCESSING

One of the intriguing aspects of lexical processing is the multiplicity of internal representations and processes which may be used when a reader encounters a string of letters. For example, the deep semantic form of analysis which a reader of a classic Russian novel emphasises is quite different from the unusual mixture of orthographic and lexical information required when the same individual attempts to solve a crossword puzzle. Although there is clearly overlap in the nature of cognitive operations involved in each of these situations, the fact that a single individual may excel in such distinct situations highlights the importance of the highly flexible nature of the human language processing system.

When considering the types of codes that are available for stimuli, there are few stimuli that have such diverse codes as words. For example, a visually presented word can be queried at a number of quite distinct levels: Does it have the letters E and S (orthography)? Does it rhyme with the word SAVE (phonology)? Does it represent an animate object (semantics)? Is it a noun (syntax)? We believe that one of the key features of the human language parser is its flexibility in engaging each of these processing pathways based on the current task demands. We would argue it is precisely this flexibility that needs to be taken into consideration in developing adequate models of word recognition.

Unfortunately, the flexible nature of processing has not always played a crucial role in theories of lexical processing. A theory that incorporates processing flexibility would seem to be necessary in order to provide a complete picture of word processing in humans. Part of the problem may be the implicit notion that experiments are performed primarily to elucidate underlying cognitive architectures. The problem with the architecture metaphor is that it lends an image of a relatively static set of constraints that remain constant within and across experimental paradigms. In fact, in reviewing the work in word recognition, it appears that a fundamental goal has been to identify which codes or pathways are obligatorily processed upon lexical presentation. Consider, for example, the work on the processing of word meaning. Researchers have argued from semantic priming studies (Neely, 1977) and Stroop studies (Stroop, 1935) that meaning is automatically accessed when a word is visually presented. Thus, there appears to be automatic and autonomous access of meaning. However, as we shall see, even this fundamental observation about lexical processing can come under the influence of attentional selection and such results appear to depend upon the processing pathways that subjects select as a function of task demands. Indeed, we will entertain the possibility that there are no obligatory processing pathways engaged for words, and that the available evidence suggests that the cognitive system has the remarkable ability to select distinct processing pathways in response to a given task demand, thereby minimising the role of alternative unselected pathways. The ability to shift between reading a novel and solving a crossword puzzle or analysing a word for meaning versus orthography/phonology are a few examples of such flexibility. With this in mind, perhaps a better metaphor in which to couch experimental studies of word recognition is a quest for the specification of a cognitive "toolbox" in which, for any given task, only a subset of the available tools are brought to bear. Thus, in the spirit of using "the right tool for the right task", subtle changes in experimental context may lead individuals to employ different cognitive tools to accomplishing the goals of a given task.

In the context of the present chapter, the "cognitive tools" are psychological pathways (see also Posner, 1978) which are devoted to the processing of particular forms of information. These processing pathways might be analogous to neural pathways, such as the ventral and dorsal visual pathways, which are devoted to the processing of different forms of visual information (Ungerleider & Mishkin, 1982). Indeed, part of the motivation behind the emphasis on processing pathways is the involvement of different regions of the brain in processing different dimensions of a single stimulus. For example, in processing words, Petersen, Fox, Posner, Mintun, and Raichle (1990) have demonstrated through the use of positron emission tomography that vastly distinct areas of the brain are engaged when participants watch visually presented words (occipital), read aloud visually presented words (temporal), and generate verbs to nouns (frontal). The role of attention is to modulate processing along multiple pathways in a manner consistent with optimal performance of a task. The degree to which a pathway contributes to performance is continuously valued and dependent upon task instructions, strategies the subject might adopt, stimulus contexts, and a host of other factors.

A processing pathway approach emphasises two points which should be kept in mind throughout the present chapter (and possibly others in this volume). The first point concerns the frequently discrepant and often contradictory findings in the literature. A particular experimental design, combined with the stimuli used and the instructions given to individual participants, will result in the individual choosing some subset of available processing pathways and weighing the information from these pathways in a particular fashion. The combined contribution of information from each of these pathways will yield some pattern of performance. The highly flexible control which the attentional system has on the implicated pathways suggests that a different set of instructions, different lexical processing task, or even changes in list configuration could lead to a differential involvement of processing pathways, and hence, a different pattern of performance. Obviously, careful experimental design and task analyses must be applied to reveal differential involvement of processing pathways and to interpret the results of particular experiments.

The second point which we feel deserves emphasis is that theories and models of language processing cannot afford to ignore the central role played by attention. Researchers in the area of visual word recognition have formulated a number of elegant models of word processing in which attentional influences on processing have typically been ignored or have fallen outside the purview of the model. Clearly such simplifications by omission are to some extent necessary in any research endeavour. However, neglect of the role of attentional processes may not always serve to simplify the theoretical questions. In particular, such neglect runs the risk of asking comparatively inflexible models to account for data from highly flexible cognitive systems. In this light, we believe that an adequate model of visual word recognition must reflect the type of attentional selection that occurs during reading. Therefore, we believe that it is paramount to keep in mind the goal standard of developing a model of lexical processing in which task demands direct attentional selection to meaning level information.

Of course, in order to appreciate the role of attentional selection in lexical processing, we must first review the elegant theoretical and empirical work that has been developed in the lexical processing literature. Thus, we now turn to a brief overview of some of the extant theoretical perspectives on lexical processing.

OVERVIEW OF THEORETICAL PERSPECTIVES

In lieu of attempting to provide an overview of the rich theory that has developed in word recognition research, we will simply highlight a few theoretical perspectives that provide a foundation for the later discussion of the empirical literature. The goal here is simply to mention a few of the approaches to word processing, not to provide a detailed evaluation of any single model. In pursuit of this goal, we have selected Forster's search model (1976), the Seidenberg and McClelland (1989) parallel distributed processing (PDP) model, and Coltheart's (1978) dual-route model. At the onset, one should note that a goal of this modelling endeavour is to develop a task-independent model of word recognition. It is precisely this assumption that we believe will need to be substantially modified by the constraining influence of attentional selection.

Forster's serial search model

In general, serial search models (e.g. Becker, 1979, 1980; Norris, 1986; Paap, Newsome, McDonald, & Schvaneveldt, 1982; Taft & Hambly, 1986) propose that, based on a preliminary visual analysis of a stimulus, an assortment of lexical possibilities becomes available. The input stimulus is compared with each member of some candidate set, one at a time, until a match is found. The search set is typically assumed to be organised so that more frequent words are checked before less frequent words. Thus, search models have an easy way to handle a principal finding in word recognition literature, i.e. low-frequency relatively uncommon words are processed more slowly and less accurately than high-frequency relatively more common words (see Balota & Chumbley, 1984, 1985, 1990; Balota & Spieler, 1999; Monsell, Doyle, & Haggard, 1989, for a discussion of the role of word-frequency in word recognition tasks).

Forster's (1976) autonomous search model includes a number of distinct access bins to the master lexicon. These access bins correspond to orthographic codes (reading), phonological codes (speech perception), and semantic/syntactic codes (speech production). Most of the work by

Forster and colleagues has dealt with the orthographic access file. The notion is that when a word is visually presented, an orthographically defined bin is created and then searched according to frequency of occurrence until a match is found. Then, based on a pointer from the matched item, the reader can access the plethora of lexical information available with that item in the mental dictionary.

Although much of Forster's work has dealt with the orthographic access file, other search models, such as Becker's (1979) verification model nicely demonstrate why one might wish to hypothesise other access routines, such as one based on semantic access files. Becker used the serial search framework to account for another principal finding in word recognition, the semantic priming effect. The semantic priming effect refers to the finding that readers are faster and more accurate to process a visually presented word (DOG) when it follows a semantically related word (CAT) compared to when it follows a semantically unrelated word (PIN). Becker simply argues that in addition to an orthographically defined search bin, subjects also have a semantically defined bin, which is a generated set of target candidates made available based on the meaning of the prime item. These related candidates are compared with the stimulus item before unrelated items, and hence, one finds a semantic priming effect.

The emphasis in Forster's original search model is on an autonomous word recognition device that is relatively uninfluenced by attentional control. In fact, such a model was quite consistent with modular views of lexical processing, in which visual lexical presentation drives the search through orthographic bins independent of attentional control. However, the primary data base that was used to test aspects of this model where speeded naming and lexical decision performance; both are tasks that place relatively minimal load on meaning processing, at least compared to reading comprehension. Moreover, work by Glanzer and Ehrenreich (1979) and Becker (1980), both advocates of the serial search framework, have demonstrated that the two fundamental aspects of word recognition, word frequency effects and semantic priming effects, can be modulated by probability manipulations across different lists. Interestingly both Glanzer and Ehrenreich and Becker have suggested that list probability manipulations influence the strategic control of different types of search processes. Thus, these studies are quite consistent with the processing pathway approach in suggesting that attentional control can modulate the manner in which access files are searched.

Coltheart's (1978) dual-route model

This model is motivated by a logical analysis of the problem facing a reader of English. Namely, that while there exists a degree of consistency of mapping spelling to sound correspondences (e.g. MINT) there are many words which violate these "standard" spelling to sound rules (e.g. PINT). In order to address such problems, Coltheart (1978) suggested that there may be two routes through which a word may be named (see Coltheart, Curtis, Atkins, & Haller, 1993, for a more recent computational version). One route entails a direct look-up of the pronunciation in the lexicon (i.e. the lexical route). Presumably, any word that has been learned by the reader is stored in memory along with its correct pronunciation. Whenever the word is subsequently encountered during reading, the visual form is used to access the lexical entry for the word. Once accessed, the correct pronunciation of the word may be retrieved directly from memory. This direct route is frequency modulated and is quite similar to Morton's (1970) logogen model. The second route computes the pronunciation based on general letter-to-sound rules (i.e. the assembled route). Use of this route is necessary at least for cases in which unfamiliar or new words are encountered, as well as to account for the relative ease with which non-words (pronounceable letter strings such as BLANT and PLATAMARG) may be pronounced (see, however, Marcel, 1980 for an alternative view of non-word naming). Because this route makes use of fairly consistent rules for translating groups of letters into sounds, it will deliver incorrect pronunciations for words that do not comply with these spelling-to-sound rules (e.g. "pint" as a rhyme for "lint"). In this way, the model nicely accounts for the frequency by regularity interaction (e.g. Seidenberg, Waters, Barnes, & Tannenhaus, 1984). Specifically, low-frequency words that have inconsistent spelling to sound correspondences (e.g. PINT) produce slowed response latencies compared to low-frequency words that have consistent spelling-to-sound correspondences (e.g. LINK). Because high-frequency words have a relatively fast lexical route, there is relatively little competition from the slower sublexical route and so one finds little or no consistency effect for high frequency stimuli (e.g. SAME vs. HAVE).

One of the more powerful lines of support for the dual-route model has come from studies of acquired dyslexics, wherein there is evidence of a double dissociation between the two processing pathways. Specifically, for one type of acquired dyslexic (surface dyslexics) there appears to be a breakdown in the lexical processing pathway. Hence, these individuals are fine at pronouncing non-words and regularly spelled words. However, when confronted with an irregularly spelled word these individuals are likely to regularise it, i.e. pronounce BROAD such that it rhymes with the non-word BRODE (see Marshall & Newcombe, 1980; Shallice, Warrington, & McCarthy, 1983). On the other hand, there is a second class of acquired dyslexics, phonological dyslexics, who appear to have an intact lexical route but an impaired phonological route. These individuals can pronounce irregular words and other familiar words that have lexical representations; however, when presented with a non-word that does not have a lexical representation, there is a considerable breakdown in performance (Patterson, 1982; Shallice & Warrington, 1980). Of course, the question that we would be interested in is the extent to which the output from these different processing pathways are under attentional control. In fact, Balota and Ferraro (1993, 1996) have reported evidence that populations that appear to have a breakdown in attentional control (individuals with senile dementia of the Alzheimer's type) appear to produce leakage from the non-selected pathway. Specifically, for a task that demands the lexical processing pathway (naming) there is an increased influence of the sublexical route, as reflected by regularisation errors (see also Patterson, Graham, & Hodges, 1994), whereas, for a task that demands the sublexical processing pathway (rhyme judgements) there is an increased influence of the lexical route, as reflected by word-frequency effects.

Seidenberg and McClelland's (1989) PDP model

An important alternative to the classic dual-route framework is the Seidenberg and McClelland parallel distributed processing approach to word recognition. In this model, information concerning features, letters, and words is distributed across an array of elements, as opposed to being represented "locally" by single units (e.g. McClelland & Rumelhart, 1981). The model consists of three layers of distributed units. In the original model there were 400 orthographic input units that were connected to 200 hidden units, which in turn were connected to 460 phonological units. The weights (connection strengths) between the input units and the hidden units and the weights between the hidden units and the phonological units do not initially represent any organised mapping (i.e. starting weights are given random values). During training, the model is presented an orthographic string which produces some output that initially is very dissimilar to the correct output. The output is compared with the correct output and the weights are then adjusted, via the back propagation algorithm, in order to gradually reduce the discrepancy between the correct pronunciation and the observed pronunciation.

The model was trained on a total of 2884 unique orthographically represented words. The probability of training a given word was monotonically related to its estimated frequency in the language. Thus, high-frequency words had a greater influence on training than low-frequency words. The result of this training regime yielded an impressive ability of the model to encode many aspects of the statistical regularity of the mapping of orthography to phonology even in a language with such variable mapping as English. In fact, Seidenberg and McClelland nicely demonstrated that the model can capture the frequency by regularity interaction that is found in human behaviour with the same set of stimuli. Moreover, the model did a reasonable job in predicting naming performance on items that it was not trained on and also on non-words (see, however, Besner, Twilley, McCann, & Seergobin, 1990).

Interestingly, when the Seidenberg and McClelland model first introduced the model it was couched in terms of a single-route alternative to the dual-route account. However, recent discussions have acknowledged the contributions of multiple information processing pathways (see Plaut, McClelland, Seidenberg, & Patterson, 1996). Perhaps indicating progress in the word recognition literature, it appears that the question has shifted from a single/dual-route dichotomy to a question of what are the relative contributions of different sources of information to performance. For example, do certain words place particular emphasis on visual analyses (bare and bear) or phonological analyses (e.g. lead as in the metal or what a leader does). The central goal at this stage might be to determine the "division of labour" across multiple processing pathways (Seidenberg & Harm, 1995). We would argue here that the division of labour will not be a static aspect of the processing architecture of the model, but rather will ultimately depend upon attentional direction that is driven by task goals. For example, the influence of semantics might be much larger if attention is directed towards understanding message-level information in reading than if it is directed more towards phonology in a task such as naming or orthography in a task such as lexical decision. Thus, we would suggest that even within a connectionist system, there needs to be a role of attentional modulation of the overall influence of the various components (see Cohen, McClelland, & Dunbar 1990, for an example of the role of attentional selection on the output from individual modules within a connectionist framework).

REVIEW OF THE LITERATURE

As noted earlier, our goal in this section was to simply sketch a few of the major processing assumptions that are available in models of word recognition. We believe that it is quite important to note here that the development of each of these models is based on data obtained in large part from studies of isolated word recognition. Hence, we shall now turn to the extant literature regarding the available evidence concerning the five targeted variables (orthography, phonology, meaning, syntax, and semantics) for this chapter. The organisation of each section involves first providing an overview of the data that have been instrumental in developing the extent models of word processing and then reviewing the relevant literature from on-line measures of reading performance (e.g. eye-fixation duration data) to determine if in fact the findings generalise to situations where attentional selection drives processing pathways relevant to message level information, the major goal of reading.

Isolated word recognition

Obviously the visual/orthographic make-up of a word remains an important "initial" influence on identification whether the word occurs in a context or is encountered in isolation. For instance, it is this visual/orthographic form that supplies the visual system with information that determines where eye fixations will occur during reading. Rayner (1979) showed that readers tend to fixate between the first and middle letters of words, regardless of word length. Before an accurate eye movement can be made from a preceding word to a given location within the next word, readers must have some information about the word shape when it initially appears in the parafovea.

A second source of information that can be extracted from a word and still be considered independent of surrounding context is phonology. A word's meaning may change as a function of its contextual setting, whereas its pronunciation is relatively context-independent with the exception of certain types of ambiguous words such as WIND, BASS, TEAR, DOVE, and LEAD (i.e. heteronyms). In fact, some have argued that the phonology of a word provides a more reliable source of information for recognition than either orthography or meaning (e.g. Van Orden, 1991). Therefore, although this chapter deals primarily with reading rather than pronouncing words, it will be important to discuss recent evidence suggesting that phonology plays an important role in visual word recognition.

Finally, there are issues regarding whether meaning-level information (e.g. context availability, number of meanings, concreteness/imageability) can influence identification (see Balota, Ferraro, & Connor, 1991 for a review). To the extent that the activation of meaning-level information assists in the recognition of a word, it may be possible to determine whether the process of recognising a word likely involves the simultaneous extraction of meaning or whether meaning activation is a by-product of (i.e. follows) word recognition. This issue has received recent attention with regard to the number of meanings ("NOM") effect, in which words having multiple meanings such as BANK are identified more quickly than words having fewer meanings (Kellas, Ferraro, & Simpson, 1988; Millis & Button, 1989; Rubenstein, Garfield, & Millikan, 1970). Despite the fact that a word is usually read for meaning, most models of word recognition (e.g. those mentioned previously) do not consider the possible early contribution of its meaning to recognition. Of course, this will be a central issue in our discussion of this literature.

Orthographic neighbourhoods. In English, there are 26 unique letters (ignoring capitalisation alternatives) which can be used in various numbers and combinations to produce words or word-like forms. Assuming reasonable restrictions on length, the number of possibilities is almost limitless. However, in English there are constraints on what combinations can occur among groups of letters in order to be considered word-like. Even with rules governing how these letters can or cannot occur (or are likely to occur) together, there are many possibilities. What is interesting, then, is that we are able to quickly recognise or identify a randomly presented word in isolation. Somehow, we are capable of isolating a single word from the population of stored possibilities. Intuitively, it seems unlikely that recognition could occur as rapidly as it does during reading if we had to search through memory checking each word with all stored words one at a time until a match is found. However, matches *are* obviously found during reading; therefore, there must be some way of speeding the word retrieval process.

One possibility for speeding a memory search would be to organise memory so that words used more frequently are evaluated before less frequently used words. Searching for most words typically encountered during reading will not require an exhaustive or extended search and so access will be relatively quick on the average. This organisation captures the finding that high-frequency words are identified more quickly than low-frequency words (Balota & Chumbley, 1984). Of course, as noted previously, each of the models presented has no difficulty capturing the word-frequency effect.

Another way to make word retrieval more efficient would be somehow to capitalise on the visual similarities among words. Havens and Foote (1963) were one of the first studies to demonstrate that the orthographic similarity of a given word to other words can modulate early aspects of word processing. They showed that words such as CASE, that under tachistoscopic exposures may be confused with a number of similar words (e.g. CANE, BASE, CARE, etc.), were identified less accurately than words less likely to be confusable with alternative words such as OVER. Interestingly, this outcome was found to be independent of word frequency. These results would appear to support the notion that there is parallel activation of a set of orthographically related words, and that word identification involves the pruning of orthographically similar candidates as perception of the stimulus unfolds across time.

Of course, one might be concerned about the possibility that individuals are using partial information to guess what the stimulus word is under the degraded presentation conditions of the Havens and Foote study. Thus, Coltheart, Davelaar, Jonasson, and Besner (1977) extended the Havens and Foote observation in a seminal study of lexical decision performance (speeded word/non-word discriminations). Coltheart et al. (1977) evaluated lexical decision tasks ("LDT") latencies to both words and non-words that varied in how many English words could be produced by replacing a single letter with some other letter of the alphabet (but without changing letter positions). Coltheart et al. called this the N of a letter string, and it has since become known as the neighbourhood of a word or non-word. It was found that words which had many neighbours were identified as quickly as words with few neighbours. Non-words, on the other hand, were more quickly identified as such when they had fewer neighbours than when they had many neighbours. Because there was no influence of neighbourhood size for words, the authors concluded that access to a word occurred directly.

Coltheart et al.'s failure initially to find neighbourhood size effects with words is inconsistent with more recent accumulation of evidence that neighbourhood size *can* affect word processing performance (e.g. Andrews, 1989, 1992; Grainger, 1990; Pugh, Rexer, Peter, & Katz, 1994). However, even within this body of research, results are inconsistent with words having large neighbourhoods producing inhibition under some circumstances (e.g. lexical decision in Grainger, 1990), and facilitation under other conditions (e.g. lexical decision and naming for low-frequency words, Andrews, 1989; see Andrews, 1997, for a recent review).

Grainger's view (Grainger, O'Regan, Jacobs, & Segui, 1989) with regard to neighbourhood effects is that neighbourhood frequency, rather than neighbourhood size, is what drives response latencies. Controlling for bigram frequency, Grainger et al. (1989) demonstrated with both lexical decision and visual gaze durations (sum total of all fixations on a word during reading) that performance was slowed on words that had at least one neighbour of a higher frequency than the stimulus itself. Orthographic neighbourhood size was not found to affect performance. Grainger (1990) further demonstrated that the neighbourhood frequency effect had opposite results for lexical decision and naming tasks. Controlling neighbourhood size, Grainger found that, as neighbourhood frequency increased, performance was slowed for lexical decisions but speeded for naming. Although Andrews (1992) did not directly manipulate neighbourhood frequency, she did examine bigram frequency and concluded that it had no effect on either lexical decision or naming performance. In accounting for the apparent discrepancy between the effects of neighbourhood frequency (i.e. inhibition) and neighbourhood size (i.e. facilitation) in LDT, Andrews (1992) proposed that Grainger et al.'s

(1989) manipulation of neighbourhood size was not strong enough to detect neighbourhood size effects (ranging only from 2.2 to 7.9 neighbours compared with her 3.4 to 12.0 range). Andrews proposed that the neighbourhood frequency effects observed when neighbourhood size was restricted (Grainger, 1990), along with her results showing neighbourhood size effects, indicates that it is likely that both processes affect word recognition. However, Andrews also argued that because the inhibitory effects due to frequency have only been demonstrated with LDTs, it might be that the locus of such effects is in the decision stage of the task, rather than a reflection of earlier affects on lexical access. This conclusion has received more recent support by Pugh, Rexer, Peter, and Katz (1994), who suggested that the inhibitory effects of neighbourhood frequency observed by Grainger (1990) were likely due to strategic, or post-access processes associated with the LDT. Interestingly, Perea and Pollatsek (1998) have recently reported data from an eye-tracking study which suggests that the frequency of neighbours influencing relatively late processes in reading (i.e. regressions and fixation durations on the subsequent word). These results are also consistent with a post access locus of neighbourhood frequency.

It is important to note that the facilitatory effects observed for increases in neighbourhood density have proved to be troublesome for serial search models of word recognition. For example, according to Forster's autonomous search model, if neighbourhoods are collected together in single search bins ordered by frequency, then, on average, a serial search would result in slower identification times for large neighbourhood items. This is because more time would be required to search larger bins (dense neighbourhoods) than smaller bins (sparse neighbourhoods). This is precisely the opposite pattern found by Andrews (1992) and Pugh, Rexer, Peter, and Katz (1994). On the other hand, if a letter string produces parallel activation of many neighbours then one might expect larger neighbourhoods to produce overall more familiarity than smaller neighbourhoods; that this might drive a "word" response in lexical decisions, and possibly slow a non-word response (see Balota & Chumbley, 1984; Besner, 1983; Besner & Swan, 1982, for familiarity-based interpretation of lexical decision performance). Such a pattern would not necessarily extend to reading performance wherein the task demands should direct attention to message level information that requires a unique identification of the target word. Thus, we shall now turn to studies of on-line reading and orthographic neighbourhood effects.

Bohemier (1994) conducted a series of experiments that were designed to directly address the role of orthographic neighbourhoods across naming, lexical decision, and on-line reading tasks. In the first three experiments, Bohemier reported evidence of neighbourhood density effects in both naming and lexical decision tasks. More importantly, for the present chapter is the pattern observed in Exp. 4: Bohemier found that there was no evidence of a density effect when the same stimulus words were embedded in short sentences and eye-fixations were measured. In fact, Bohemier cites the results from three eye-tracking studies that have failed to find neighbourhood density effects in on-line measures of reading (Bohemier, 1991; Bohemier & Inhoff, in prep., Exps. 3 and 4; Grainger et al., 1989, Exp. 2). Hence, it appears that, when attentional selection is directed towards message-level pathways, one finds a reduced influence of a code (reflected by orthographic neighbourhood effects) that appears quite useful in isolated word processing tasks. Thus, the nature of the word recognition architecture appears to be highly influenced by where attention is directed. In making lexical decisions, overall activation of many candidates is likely to lead to increased familiarity thereby increasing the ease of making lexical decisions. Also, it is possible that when there are many candidates that converge on a similar spelling to sound correspondence, naming latencies will be facilitated as observed by Andrews (1992). Both of these benefits may be somewhat minimised when attention is directed towards message-level pathways and there is a clear demand to integrate the unique word into on-going comprehension processes, as in reading text.

Phonology. Much of our linguistic experience (especially during language acquisition) is with the sounds of words. Our experience with how words sound is obviously critical to learning to read (Perfetti, Beverly, Bell, & Hughes, 1987). Moreover, there is considerable evidence that phonological recoding is a critical component of reading comprehension (see Rayner & Pollatsek, 1989 for a review), and hence there is little debate concerning the importance of phonological coding during reading. The question that has received considerable attention is whether phonological information plays an early role in word identification, and hence we can simplistically divide this literature into two camps: First, according to the phonology-is-necessary camp, all word identification must go through some transformation of orthography to phonology for identification to occur. There are a number of proponents of this view that have varying commitments to this extreme position (e.g. Lukatela & Turvey, 1994a,b; Perfetti, Bell, & Delaney, 1988; Rayner, Sereno, Lesch, & Pollatsek, 1995; Van Orden & Goldinger, 1994). Second, according to the multiple route camp, there is a number of distinct routes in early lexical processing; one that is primarily mediated by orthography to phonology transformations and one that is more directly available from the mapping of a visual stimulus onto a lexical representation. Of course, the classic model within this camp is the dual-route model (Coltheart, 1978; Coltheart et al., 1993) reviewed earlier. We would simply argue that the role of phonology in word recognition depends upon the direction of attention to distinct processing pathways. The extent to which phonology is demanded by task goals will at least in part modulate the role of phonological processing. We shall now turn to a review of such evidence.

Interestingly, there are a number of findings that are quite consistent with the role of attention (modulated by task demands) in the degree of influence of the phonological processing route in word recognition. First, consider the classic consistency effect that is observed in naming performance. As noted earlier, readers are relatively slow to name low-frequency words that have inconsistent spelling patterns (e.g. PINT) compared to low-frequency words that have consistent spelling patterns (e.g. PARK). However, one might ask that because the response is overt naming, possibly there is an increased reliance on phonological information. Consistent with this argument, one finds relatively little influence of spelling to sound correspondence when one uses the same stimuli in a lexical decision task (see Andrews, 1982; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). This would be expected according to the processing pathway perspective because overt naming may place a heavier load on a phonological pathway compared to lexical decision, in which there may be more emphasis on a visually based pathway.

An interesting set of experiments by Monsell, Patterson, Graham, Hughes, and Milroy (1992) has recently indicated that, even in naming performance, one can find strategic control of different processing routes. Monsell et al. were motivated in part by an observation by Midgley-West (1979). In the Midgley-West study, subjects were asked to name a series of 24 non-words, followed by an irregularly pronounced word. The interesting pattern was that subjects often regularised the last word (e.g. if the last word was WOLF, the subjects would name it as if it was pronounced like GOLF after naming 24 non-words). The argument is that, because non-words primarily demand the sublexical spelling to sound pathway, participants decreased their reliance on the lexical pathway and hence regularised the irregular word. Thus, these results are quite consistent with the notion that there is attentional control of processing pathways even in naming. Monsell et al. (1992) extended this finding to situations where subjects either received lists of only non-words, only exception words, or mixed lists of non-words and exception words. Subjects named exception words faster and made fewer regularisations when they were not also prepared to name non-words. Thus, Monsell et al. argued that subjects controlled, via attentional direction, the role of the sublexical spelling to sound pathway. Although there has been an alternative interpretation of the Monsell et al. results (see Lupker, Brown, & Colombo, 1997), there also is converging evidence across a number of distinct paradigms and languages for attentional control of these pathways (e.g. Baluch & Besner, 1991; Frost, Katz, & Bentin, 1987; Paap & Noel, 1991; Pugh, Rexer, Peter, & Katz, 1994; Simpson & Kang, 1994; Zevin & Balota, 1999).

Of course, there are other data that have been taken as strong support for an early role of phonology in lexical processing. Consider the important work of Perfetti et al. (1988). They reported the results from a series of masking studies in which a stimulus word was presented (MAIL) and then followed by one of three different types of mask: a graphemically and phonologically unrelated mask (e.g. FLEN), a graphemically related and phonologically unrelated mask (e.g. MARL), or a graphemically related and phonologically related mask (e.g. MAYL). The results clearly indicated that performance was best in the graphemically and phonologically related condition, which was better than in the graphemically related and phonologically unrelated condition, which in turn was better than the graphemically and phonologically unrelated condition. Thus, there appears to be some benefit from a phonologically related mask above and beyond simple visual similarity. Interestingly, however, Verstaen, Humphreys, Olson, and d'Ydewalle (1995) have recently reported evidence that, even under these conditions, there may be strategic control of the output from these different pathways. Specifically, Verstaen et al. found that the influence of phonology was eliminated under conditions in which all target items were homophones. Verstaen et al. argued that subjects modulated, via attentional control, the influence of the phonological pathway, minimising its role when the target stimuli could not be discriminated based on sound. Clearly, it appears that the role of phonological information in early aspects of lexical processing is not independent of attentional control.

A more important question concerns the role of spelling to sound correspondences in a task in which attention is directed towards message-level information, as in reading. There is some evidence of an early role for phonology during reading in a study by Pollatsek, Lesch, Morris, and Rayner (1992). Pollatsek et al. used the parafoveal preview paradigm. In this study, individuals' eye movements were monitored while reading a sentence. In the critical sentences, a homophonic word (e.g. reins) or a visually similar word (e.g. ruins) was replaced with a homophone (e.g. rains). Pollatsek et al. found that fixation duration was shorter on the target when it was preceded by the homophone word compared to the visually similar word (see also Folk & Morris, 1995).

Inhoff and Topolski (1994) examined fixations while subjects were reading words that either had consistent or inconsistent spelling to sound correspondences. These words were embedded in short neutral sentence contexts. The results indicated that there was no evidence of neighbourhood consistency on either first fixation data or gaze durations and only small effects of spelling to sound regularity on first fixations, despite the fact that this same set of stimuli yielded quite striking consistency and regularity effects in speeded naming performance. These results suggest that the demands of the naming task may in fact exaggerate the role of phonological codes, and that such effect of spelling to sound consistency effects may not be strong in the early analysis of words in on-line reading as the strong versions of mandatory phonological processing would demand. This is not to say that there is no role for phonology early in the processing of lexical information during reading. Rather, the point is that phonological influences appear to be attenuated when attention is directed towards message-level processing pathways. This contrasts with processing during single word naming studies in which attention is directed toward the production of a phonological output (e.g. Lukatela & Turvey, 1994a,b). We believe that further work is necessary to determine whether the same factors that strongly modulate naming performance, and have had a striking influence in theory development, also produce a strong influence in on-line measures of reading (see Folk & Morris, 1995, for one such study).

In the previous two sections we have reviewed evidence that addresses the influence of both orthographic and phonological information on the routes to word recognition. Clearly, these two types of information have dominated work in word processing. We shall now turn to a third factor that has been relatively ignored in word processing. That is, the extent to which the ease of accessing meaning for a word influences the speed to recognise that word.

Meaning. A word's meaning(s) must be retrieved from memory to understand the message conveyed by the text. It would appear that some translation of an auditory or visual presentation of a word is necessary in order to locate the meaning of the word in memory. That is, the process appears to require that the lexical or phonological status of a word be determined prior to meaning activation. This intuition, is captured in the notion of a modular word recognition system (e.g. Fodor, 1983; Forster, 1976).

Fodor (1983) elaborated a modular view of processing whereby the various processes associated with language use are delegated to several modules. According to this view, for example, accessing a representation of a word in the mental lexicon proceeds the same whether it is encountered in a context or in isolation (presuming that the quality of the inputs are equivalent). No influences external to the lexicon will affect the speed with which the word is located. Once it is located, however, the lexical processor provides the lexical information to the next module in the series for processing (e.g. context integration, elaboration). It is at these later modules that aspects of a word such as its syntax, phonology, and meaning may be observed to affect performance. The modular position contrasts with a strong interactive and cascadic processing view (McClelland, 1987; McClelland & Rumelhart, 1981) in which information is shared among operations throughout the recognition process (including meaning). In this case, a word's meaning(s) could be useful for recognition processes because lower level information about the input will become available to higher level processes prior to identification. These higher level processes (e.g. at the meaning-level) will provide useful information to lower level processes (e.g. lexical level) thereby aiding identification or lexical access.

The extent to which the meaning-level characteristics of a word have been shown to influence isolated word recognition has varied (see Balota et al., 1991 for a review). For example, the concreteness of a word's meaning has been shown to influence lexical decision in that more-concrete words have an advantage over less-concrete words (e.g. Day, 1977; Kroll & Mervis, 1986). However, there is also evidence suggesting that concreteness may *not* affect lexical decision (e.g. Richardson, 1976) when potentially confounding factors (e.g. familiarity, context availability) have been controlled (Schwanenflugel, Harnishfeger, & Stowe, 1988). It is noteworthy that these potentially confounding factors are themselves likely to be correlated with semantic characteristics of a word. Thus, it is therefore not clear from this literature that meaning-level influences on isolated word recognition are non-existent. Interestingly, Strain, Patterson, and Seidenberg (1995) have recently reported evidence that the imageability of a word (abstract versus concrete) can produce an influence on word naming for words that are relatively low in frequency and spelling to sound regularity. Specifically, subjects were slower to name exception words with abstract meanings (e.g. scarce) compared to abstract regular words (e.g. scribe) or highly imageable exception words (e.g. soot). This pattern is quite intriguing and suggests that there may be a trade-off between high-level semantic information and the clarity of the translation between orthography and phonology, precisely as an interactive architecture might predict (see also Cortese, Simpson, & Woolsey, 1997).

The concreteness or imageability of a word seems intuitively to tap some semantic aspect of meaning which probably varies from word meaning to word meaning. In addition to this semantic variable, a second variable that has received considerable attention in the literature is the number of meanings that are available for a word. For example, the word "bank" can refer to a number of distinct unrelated meanings (something an aeroplane does in a turn; the land immediately adjacent to a river; a financial institution). It is possible that if there is increased semantic support for both phonological and orthographic information then number of meanings may lead to facilitated lexical decision and naming performance. There have been a number of demonstrations of such effects in both word naming and lexical decisions (Balota et al., 1991; Fera, Joordens, Balota, & Ferraro, 1992; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas et al., 1988; Millis & Button, 1989; Rubenstein et al., 1970; Rubenstein, Lewis, & Rubenstein, 1971). Unfortunately, however, there have also been a number of concerns raised about this observation. For example, Clark (1973) found that the number of meanings ("NOM") effect observed by Rubenstein et al. (1970) in their analysis by subjects was not significant when analysed by stimuli. In addition, a later study by Gernsbacher (1984) demonstrated that the NOM effect observed by Jastrzembski (1981; Jastrzembski & Stanners, 1975) could be attributed to an uncontrolled variable confounded with number of meanings (i.e. familiarity). As noted earlier, it is unclear what dimensions of a stimulus are tapped in unspeeded familiarity judgements, and so it is possible that familiarity ratings actually include meaning-level dimensions of a stimulus. Moreover, the NOM effect observed by Kellas et al. (1988) was obtained using stimuli that were controlled along many dimensions, including familiarity, in lexical decision performance.

A recent series of experiments by Borowsky and Masson (1996) replicated the NOM effect in lexical decision performance under conditions in which the non-words were orthographically legal, and failed to find a NOM effect with orthographically illegal non-words. In addition, Borowsky and Masson did not find any consistent effect of number of meanings in naming performance (although there was an effect found in Exp. 1 by subject means). Borowsky and Masson accounted for the general discrepancy between their results (failure to show NOM effects in naming) and previous naming results showing NOM effects (Fera et al., 1992) as possibly being due to a differential sensitivity of the subject populations to factors confounded with number of meanings (e.g. word length, neighbourhood density). When Borowsky and Masson matched these possible differences between ambiguous and unambiguous items, there was no trace of a NOM effect in naming (Exp. 2).

As Borowsky and Masson (1996) pointed out, NOM effects rely on between-stimuli comparisons of performance, and so will always be vulnerable to the possibility that observed differences are due to some uncontrolled variable that happens to vary between stimulus sets (e.g. the familiarity account of the NOM effect proposed by Gernsbacher, 1984). Borowsky and Masson proposed that matching stimuli on as many extraneous variables as possible, as well as partialling out such influences using multiple regression techniques, will maximise the validity of the NOM effect. Because the number of meanings metric typically involves either counting meanings in a dictionary, or asking subjects to provide an estimate of number of meanings, the final meaning tally associated with individual items represents a fairly continuous, but potentially skewed dimension. Simply splitting such items into two groups (ambiguous and unambiguous) for subsequent analyses does not seem ideal for representing the nature of potential relationships between NOM and other variables of interest. Indeed, this seems especially problematic if items classified as unambiguous according to such splits are actually ambiguous. This is illustrated more concretely in that a word such as "force" can be classified as ambiguous by Millis and Button (1989) but classified as unambiguous by Borowsky and Masson (1996). Although regression techniques were employed by Borowsky and Masson (1996), it is possible that the range of NOM across the 64 ambiguous and 64 unambiguous items used in this study was too variable to provide a sensitive test of meaning effects in naming.

We have examined the NOM effect more closely using the full set of 360 stimuli (see Kellas et al., 1991), which have a fairly large range for rated number of meanings. Based on the concerns raised by Borowsky and Masson (1996), we used regression analyses to examine contributions of variables known to have been (or may have been) confounded with number of meanings. Most importantly, we obtained the settling times from a recent model of word naming (Plaut, McClelland, Seidenberg, & Patterson, 1996) to partial out the regularities of various spelling to sound correspondences. The results of this study indicated that there were clear number of meaning effects that occur in word naming beyond what can be accounted for by frequency number of neighbours, and the settling times obtained from the Plaut et al. model, $t(352) = 4.09$, $P < .01$. Obviously, this pattern suggests either that (1) the Plaut et al. model does not capture some components of orthographic to phonological computations involved in naming (see Balota & Spieler, 1998; Besner & Bourassa, 1995; Seidenberg & Plaut, 1998; Spieler & Balota, 1997, for further discussion), or (2) a level that takes into account number of meanings needs to be added to the model.

It appears then that there may be some relatively small beneficial effect of NOM on word naming and a stronger effect on lexical decision performance. Within the present chapter, we are interested in what sorts of processing dimensions engaged by both naming and lexical decision might produce such a facilitatory effect of polysemous words. With respect to the lexical decision task, one might argue that because words and nonwords can be discriminated on the amount of meaning that is available, subjects may direct attention to the overall activation available in meaning-level pathways, thereby producing the beneficial effect of words with multiple meanings. Turning to naming performance, it is possible that multiple meanings that converge on the same phonological output may have a stabilising force for that phonological output.

If the previous accounts of NOM effects in naming and lexical decision are correct, then one might ask what would be expected if one examined the influence of NOM in reading. Based on the processing pathways approach, one might expect just the opposite influence of NOM in on-line reading measures. Specifically, because in reading attention is directed towards message-level processing pathways, one cannot integrate two unrelated meanings simultaneously with a previous context and so one might expect a slowdown in resolving ambiguity. Interestingly, this is precisely what Duffy, Morris, and Rayner (1988) and Rayner and Frazier (1989) have found. In both of these studies, gaze durations were longer on ambiguous words compared to unambiguous words in neutral context. When the number of meanings available for a word can increase either the ability to discriminate words from non-words (lexical decision) or possibly provide converging evidence for a phonological representation (naming), one finds evidence of a benefit from the availability of multiple meanings, whereas, when attention is directed towards message level information (and hence meaning selection must occur), one finds some decrement in performance.

It is important to note here that a number of recent models of word processing have placed considerable importance on the difference in the NOM effects in naming, lexical decision, and reading performance. For example, Borowsky and Masson (1996) have interpreted such effects (assuming no effect of NOM in naming) within Masson's (1991) distributed model of semantic priming and word processing. Borowsky and Masson argued that such effects may simply fall from different criteria used across naming, lexical decision, and gaze durations. In addition, Kawamoto, Farrar, and Kello (1994) have presented a connectionist model that also has been directed towards handling the opposite effects of number of meanings in lexical decision and in gaze durations. Kawamoto et al. suggest that such differences naturally follow from differences in monitoring information from different sets of units in the network. In fact, the notion of a monitor directed towards output of different processing modules is quite consistent with the notion of the current processing pathway approach. We believe that an important distinction that may eventually need to be made in this area is whether the results are best captured by differences in criteria or in monitoring the output from different modules; both of these approaches do not change the parameters influencing the activation across units within a module. The alternative possibility is that there is a module that maintains task demands and this module actually can control a gain parameter that modulates the activation processes within the relevant and irrelevant processing modules ("pathways" in keeping with the present jargon). This is the approach taken in the architecture described in the Cohen, Dunbar, and McClelland (1990) model of Stroop performance.

Contextual constraints in word recognition

Heretofore, we have reviewed evidence that effects of traditional variables that have been investigated in isolated word processing tasks such as lexical decision and naming can sometimes have quite different influences when attention is directed to message-level information such as in on-line reading measures. We shall now turn to a discussion of the results from studies that directly address the influence of context on word processing. Obviously, it is relatively rare that words are read in isolation. When processing words in context, there are at least two major types of information that play an important role: syntax and message-level semantic information. However, there is considerable controversy concerning how and when such information plays a role. There are at least three ways in which discourse-level information would likely be useful to reading: Syntactic and semantic information may serve to facilitate *initial* word processing, *subsequent* integration processes, or *both.* If these context-based sources of information are useful for initial word processing, it becomes less clear whether examinations of isolated word recognition generalise to normal reading processes. On the other hand, if word recognition proceeds uninfluenced by sentence-level (or more complex) sources of information, then it would suggest that discourse information is only important for later or higher level processes. If both early recognition and late comprehension processes are influenced by syntactic and semantic information, then examinations of either isolated word recognition or of discourse comprehension alone are incomplete.

Syntactic contextual constraint. It seems reasonable (and many theories of syntactic processing hold) that a word's syntactic classification is determined very rapidly in order for it to be integrated with the grammatical structure of a sentence. This view suggests that, at least under some conditions, syntactic/contextual information may actually be useful in the early aspects of word processing.

The early investigations of the role of syntax on initial word recognition processes involved minimal syntactic priming manipulations along with lexical decision and naming responses as the dependent measures. For example, Goodman, McClelland, and Gibbs (1981) examined syntactic priming in a lexical decision task. Part of the reason for examining relatively impoverished syntactic (priming) contexts was that the authors wanted to reduce the likelihood that their results could be interpreted as semantically based rather than syntactically based (a possible interpretation of the results of Fischler & Bloom, 1979 and Schuberth & Eimas, 1977). It is difficult, for example, to manipulate the syntactic constraint of a sentence without also affecting semantic constraint. Goodman et al. reasoned that the use of carefully constructed minimal priming contexts could serve to address this potential problem if a syntactic prime made the target syntactically predictable but not semantically predictable (e.g. whose-planet; the-tree; he-sent).

Goodman et al. (1981) demonstrated effects of both syntax and semantics on lexical decision times (see also Wright & Garrett, 1984). When semantic and syntactic prime conditions were blocked, words were identified more quickly when they were syntactically or semantically primed (e.g. FRUIT-APPLE, YOUR-OVEN) than when they were incongruently primed (e.g. THREAD-APPLE, THEY-OVEN). However, when semantic and syntactic primes varied from trial to trial, priming effects were observed only for semantic primes. Goodman et al. concluded that subjects' utilisation of syntactic information in the blocked but not the mixed conditions indicated that syntactic priming is likely to be the result of conscious or strategically driven processes (cf. Posner & Snyder, 1975). On the other hand, because semantic effects were observed regardless of changes in stimulus presentation (mixed versus blocked), Goodman et al. hypothesised that semantic effects result from automatic structural characteristics of the lexical processing system. Consider these results within the processing pathway approach. It is possible that in the blocked conditions, subjects were able to selectively focus attention to information most beneficial to performance (syntactic elements in the HE-SENT conditions and semantic information in the FRUIT-APPLE conditions). Under mixed conditions, semantic information apparently wins out over syntactic information, possibly because more information is typically conveyed via semantic routes than syntactic routes in relatively impoverished conditions (two words) such as those used by Goodman et al. (1981).

Seidenberg, Waters, Sanders, and Langer (1984) extended the results of Goodman et al. (using the identical stimuli) by contrasting performance on the LDT with naming performance (see also Sereno, 1991). Syntactic effects were observed only for the LDT, which was interpreted as support for the view that syntactic processing is under strategic control. Although the conclusions among these studies tend to converge, it is worth considering whether single-word syntactic primes actually reflect the full engagement of syntactic analyses that are necessary when reading sentences. Studies which have examined syntactic influences using sentence contexts have shown syntactic effects which seem to support the hypothesis that one-word contexts do not fully engage syntactic processes (e.g. Nicol & Swinney, 1989; West & Stanovich, 1986; Wright & Garrett, 1984). Unfortunately, one must temper the conclusions of these studies because it is possible that some of the effects were due to a type of post-lexical checking process, described earlier, that is engaged by the demands of the lexical decision task. Thus, we again need to look at converging evidence from on-line measures of reading when attention is directed to message-level information.

The influence of syntax on early word processing in reading was addressed in a classic study by Frazier and Rayner (1982). In this study subjects were presented one of the following two sentences:

(1) *Since Jay always jogs a mile this seems like a short distance to him.*
(2) *Since Jay always jogs a mile seems like a short distance to him.*

Frazier and Rayner (1982) were interested in, among other strategies, a syntactic parsing strategy called late closure. According to late closure, readers attempt to attach new items to the phrase or clause currently being processed. The critical issue in these two sentences is that in (1) the subject can attach the word *mile* as the object of *jogs* (consistent with late closure), whereas in (2) subjects may attempt to attach *mile* to *jogs* but then when they encounter the word *seems* it is syntactically incorrect and so subjects will need to reparse the sentence. Interestingly, Frazier and Rayner found that subjects increased their first fixation on the word *seems* in (2) compared to (1). Given that eye fixations on the target word were on the order of 250 ms, it appears that subjects detected the incorrect syntactic interpretation very early on in processing, thereby suggesting a strong influence of syntactic constraint on early lexical processing. This of course is quite reasonable given the importance of the direction of attention to message-level information, and the immediate loss of such information if the sentence is misparsed.

The power and independence of the syntactic parser on lexical processing in reading was also elegantly demonstrated in a set of experiments by Britt, Perfetti, Garrod, and Rayner (1992), Ferreira and Clifton (1986), Rayner, Carlson, and Frazier (1983), and Rayner, Garrod, and Perfetti (1992). In general, these studies indicate that there is very little influence of contextual constraint on the initial stages of strong syntactic preferences. For example, Rayner et al. (1983) found that readers tend to favour a particular syntactic structure when given an ambiguous initial clause such as, "The performer sent the flowers...". Here, readers tend to initially process this as a simple active clause which would be inconsistent with the grammatically acceptable reduced-relative completion, "...*was very* pleased with herself." Rayner et al. (1983) found that readers slowed down their reading considerably at the point of disambiguating of the sentence (*was very*) for the reduced-relative version of the sentence compared with the consistent simple-active version of the sentence (i.e. "...and *was very* pleased with herself." These findings are in agreement with either the view that the influence of certain types of syntactic processes act prior to pragmatic considerations or independent of higher level semantic operations (Fodor, 1983; Forster, 1976, 1979; Oden & Spira, 1983; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982).

Unfortunately, there has been considerable recent controversy concerning the influence of contextual constraints on syntactic parsing. For example, there have been a number of studies that have demonstrated an influence of context on reading times for reduced relatives (MacDonald, 1994; Pearlmutter & MacDonald, 1992; Trueswell & Tanenhaus, 1991; Trueswell, Tanenhaus, & Garnsey, 1994). In an attempt to reconcile the differences across these two sets of studies, MacDonald, Pearlmutter, and Seidenberg (1994) argued that some of this discrepancy may be accounted for by differences in the frequency of the morphological forms of ambiguous verbs. It appears that when context effects were found on initial syntactic parsing, verbs were used that made the reduced relative a more viable option, whereas when context effects were not found it was more likely that verbs were used that did not have a reduced relative reading as a likely option. MacDonald et al. argue from these results that the reader will use whatever information that is available to constrain a particular interpretation. This again suggests a non-static processor of information that takes into account multiple sources of information that are relevant to accomplish the goals of the task. Possibly, the more important question is not whether the syntactic parser acts independent of contextual constraints but how does the reader modulate the activation amongst a number of relevant processing pathways that are necessary to achieve the goal of reading, i.e. abstract message-level information. We shall now turn to a second type of constraint which for simplicity we label as semantic contextual constraint.

Semantic contextual constraint. At the onset of this subsection, it is important to note that we shall use the term "semantic" contextual constraint rather loosely here, and shall be discussing a number of distinct levels of constraint in this section. Possibly, a better label would be non-syntactic interword constraints. However, because of the prominent role semantic priming research has played in this area, we will retain this label, and attempt to indicate the precise nature of the constraint in the reviewed literature.

The influence of semantic context on word recognition has been a central area of research within a number of distinct areas of cognitive psychology. For example, a considerable amount of work in this area has dealt with semantic priming effects (see Neely, 1991 for a review). As noted earlier, the semantic priming effect refers to the finding that subjects are faster and/or more accurate to make a response to a target stimulus when it follows a related prime compared to when it follows an unrelated prime. There has been considerable evidence in this literature that the nature of priming effects are quite dependent upon the task constraints (see Balota & Lorch, 1986; Neely, 1991; Sereno, 1991; Stanovich & West, 1983). This is quite consistent with the processing pathway approach because the tasks determine what sorts of pathway are most useful for making a given response.

In order to illustrate the powerful role of task constraints in interpreting priming effects, we shall briefly review a set of experiments by Balota and Paul (1996). This study addressed the issue of whether two associatively related primes produce additive, underadditive, or superadditive effects on target word processing. Balota and Paul were interested in the influence of multiple primes, because in natural sentence processing there are multiple constraints (semantic, syntactic, global constraints) on target word processing. In this study, there were two major classes of stimuli that differed with respect to whether the primes converged on the same semantic representation of the target (e.g. LION-STRIPES-TIGER) or diverged onto distinct semantic representations (e.g. MUSIC-KIDNEY-ORGAN). The results of the first five experiments yielded clear additive effects of related primes (i.e. the beneficial effect of two related primes was the algebraic sum of the independent effects of each of the single primes). This pattern was found in both naming and lexical decision tasks, across varying stimulus onset asynchronies, and for degraded targets. More importantly, the additive effects were observed equally for conditions in which the primes converged onto the same semantic representation (e.g. LION-STRIPES-TIGER) and conditions in which the primes diverged onto distinct semantic representations (e.g. MUSIC-KIDNEY-ORGAN). Because there was no influence of the type of semantic representation that was being primed (i.e. no difference between convergent and divergent semantic conditions), Balota and Paul suggested that the locus of the observed additive effects is most consistent with the notion that the subjects were primarily relying on lexical processing pathways in both the naming and lexical decision tasks. In the sixth experiment, subjects' attention was directed to select a semantic processing pathway by requiring a relatedness judgements between the target word and the primes. Now, the results differed dramatically between convergent and divergent target conditions. Specifically, for targets in which the primes converged onto the same semantic representation there were clear additive effects of the primes; however, for targets that included primes that diverged onto distinct semantic representations, there was no additional benefit of two primes related to different semantic interpretations of the target over a single prime. These results suggest that when the priming task demands selection of meaning-level pathways, one will observe very different patterns of data compared to when the task primarily emphasises lexical processing pathways. This pattern of dissociation across different types of tasks that place more or less emphasis on lexical level representations compared to meaning-level representations is quite consistent with the data reviewed previously on meaning-level effects on isolated word processing. Specifically, in naming and lexical decision, there appears to be a benefit of multiple-meaning representations, whereas, in reading, there appears to be a deleterious effect when multiple meanings are available. A similar pattern is found in Balota and Paul's multiple priming studies when attention is directed towards meaning level representations via the relatedness judgements. The question that we now turn to is what is the influence of different types of message-level constraint in our target task, reading.

Hess, Foss, and Carroll (1995) discussed three fundamental models that have evolved to account for effects of sentence context on word processing. According to the first view, associative priming accounts of context effects typically rely on automatic spreading activation from one word in the mental lexicon to all other related words. Because the spread of activation is assumed to be automatic, there is no mechanism for preventing multiple meanings of ambiguous words, for example, from becoming activated even when the ambiguous word is embedded in a biasing context such as, "We carved our initials into the *bark*" (Seidenberg et al., 1982; Swinney, 1979). As Hess et al. point out, a clear advantage of this model is that it can handle context effects independently of the contextual setting in which words occur (i.e. single- or multiple-word primes). According to the second view, there is a direct influence of discourse-level representations on lexical processing (e.g. Auble & Franks, 1983; Foss & Ross, 1983; Hess et al., 1995; Sharkey, 1990; Sharkey & Sharkey, 1992; Vu, Kellas, & Paul, 1998). Here, the notion is that context effects in reading are a reflection of the extent to which a given target word fits a higher level representation that has been extracted from the earlier text. Finally, there are hybrid models in which context effects are argued to result from multiple sources within a text. One source is a fast-acting automatic spread of activation from a single word of a text to related words within the lexicon. The other source occurs later and represents a more global process of integration.

The available evidence from on-line reading measures suggests that there are clear influences of contextual constraint during reading. For example, consider two early studies by Ehrlich and Rayner (1981) and

Zola (1984). Both studies found clear effects of *predictability* of a target word on fixation durations. Of course, there is some question about where in the system such effects of predictability are occurring. For example, in the Zola study subjects were presented with a strong contextual adjective (e.g. *buttered* before *popcorn* versus *adequate* before *popcorn*). As noted previously, the effect of predictability may be due to more local intralexical priming or possibly due to higher level discourse congruity effects.

Given that there appear to be clear influences of constraint on word recognition, let us now consider three distinct ways in which researchers have attempted to delineate the locus of such constraint, i.e. more intralexical effects versus higher level text integration effects. For example, Schustack, Ehrlich, and Rayner (1987) manipulated both global contextual constraint (distance of previous mention of the target item in the text) and local contextual constraint (via a preceding associatively related verb). The results indicated that in naming performance there was only an effect of the local context, with no influence of global constraint. On the other hand, when one considers on-line reading measures, one finds clear influences of both variables on eye fixation measures. Schustack et al. interpreted this pattern as indicating that the effect of local information in both naming performance and as reflected in eye initial fixations on target words may reflect a type of intralexical priming; however, the more global discourse effect of distance on eye-fixations is more likely to reflect integration processes that are not engaged in speeded naming. Thus, based on this pattern it appears that the hybrid model may be most applicable to the reading situation.

Morris (1994) also attempted to distinguish between local and global context effects during reading. In the Morris study, readers were presented sentences such as *The gardener watched as the barber trimmed the moustache*. Here, one can see that there are two words related to *trimmed* (i.e. *gardeners trim hedges* and *barbers trim hair*). A change in the sentence such as *The gardener who watched the barber trimmed the moustache*, results in a mismatching of the message-level context and the lexical relationships among the words. According to a lexical view of context effects, fixation times on "*moustache*" should not be affected by message-level shifts as long as the lexical components have not been altered (as in the previous sample sentences). A discourse model, on the other hand, requires that the inconsistency between what *gardeners trim* and the message of the second sentence (*a gardener trimming a moustache*) will increase fixation times while readers attempt to resolve the apparent conflict. Morris found that fixation times on *moustache* were significantly shorter when it was presented in a sentence that contained the words *barber*, *gardener*, and *trimmed* but only when the overall sentence context was consistent with the *barber trimming the moustache.* No facilitation was observed when changes in the context made *barber* inconsistent with *moustache trimming*. Although these results tend to imply that it is overall context (discourse) which affects word processing, Morris also found *lexical* influences on word processing in which gaze durations on the critical verb in the sentence (e.g. *trimmed*) were unaffected by the overall message. Specifically, when either *barber* or *gardener* appeared alone in the sentence (the other replaced by a neutral word such as *person*), gaze durations were shorter compared to when both were replaced by neutral words. Thus, based on these on-line measures of reading, it appears that both types of constraint are powerful determiners of lexical integration (see also Duffy, Henderson, & Morris, 1989; O'Seaghdha, 1989).

The second approach to determine the locus of the influence of predictability on word recognition versus integration processes is to orthogonally manipulate variables related to the two factors. This was done in a study by Balota, Pollatsek, and Rayner (1985) in which predictability of a target word was crossed with the visual similarity of a parafoveal preview word. For example, participants would read sentences such as the following: *Since the wedding was today, the baker rushed the wedding cake/pies to the reception.* The high-constraint target was *cake* whereas the low-constraint target was *pies.* In addition, when subjects were fixated before a critical boundary (within the second occurrence of the word *wedding* in this example), there were a number of distinct types of previews that were crossed with predictability of the target (e.g. identical previews, *cake*, *pies*; visually similar previews *cahc picz*; visually dissimilar preview, *bomb*). The results indicated that there was an interaction between contextual constraint and parafoveal preview. Specifically, subjects apparently benefited more from visually related parafoveal information for high constraint targets (cake) than for low-constraint targets (pies). Because the extraction of parafoveal information should be relatively early in word processing, one might argue that the influence of predictability in this case is on word identification processes. In further support of this observation, subjects were also more likely to skip the high-constraint targets than the low-constraint targets, again supporting a relatively early influence of context. A similar pattern was reported by Schustak, Ehrlich, and Rayner (1987). Finally, it should be noted that predictability also appeared to have had a later effect in the Balota et al. (1985) study because there was some evidence of a spill-over effect of predictability onto fixations after the target word was processed. This later effect may be more likely to reflect an influence at global text integration processes than more local processing constraints (i.e. lexical information).

The third approach to distinguishing between the influence of local and global contexts is to examine the influence of context on reading polysemous words (e.g. *mint*, *bear*, *top*, etc.). Because ambiguous words have multiple possible semantic interpretations in isolation, embedding them in a constraining context and probing for meaning activation (e.g. using naming or lexical decision) provides information about the relative success of the context in determining which meaning is activated (or received the greatest activation). The basic positions have been either that context *can* affect initial meaning activation (context dependent view), or that context *does not* affect initial meaning activation (context independent view). The context-dependent view is more consistent with a discourse view of priming, whereas the context-independent view is in line with lexical accounts of context effects.

There is a relatively large literature examining context effects in resolving ambiguity for polysemous words (e.g. Conrad, 1974; Dooling, 1972; Foss, 1970; Foss & Jenkins, 1973; Holmes, Arwas, & Garrett, 1977; Kellas, Paul, Martin, & Simpson, 1991; Neill, Hilliard, & Cooper, 1988; Oden & Spira, 1983; Onifer & Swinney, 1981; Paul, Kellas, Martin, & Clark, 1992; Schvaneveldt, Meyer, & Becker, 1976; Seidenberg et al., 1982; Simpson, 1981; Swinney, 1979; Swinney & Hakes, 1976; Tabossi, 1988; Till, Mross, & Kintsch, 1988). The results from such studies have typically been viewed as supporting the context-independent (lexical) view of initial processing. As mentioned earlier, this outcome is also consistent with accounts of single prime word recognition (see Simpson, 1984, 1994, for reviews) and so can be captured within similar theoretical frameworks without positing involvement from higher levels of processing (i.e. discourse effects). Although there appears to be converging evidence across different tasks for lexical-level effects, there has also been some evidence to support discourse-level effects (Kellas et al. 1991; Paul et al., 1992; Simpson, 1981; Tabossi, 1988; Tabossi, Columbo, & Job, 1987; Van Petten & Kutas, 1987; Vu et al., 1998).

One possibility that might account for both previous failures and successes to show discourse-level effects relies on the processing pathways approach we have attempted to develop throughout this chapter. Specifically, the extent to which evidence will be found for lexical or discourse influences on word processing depends on the information subjects utilise in order to perform the task. This can presumably be manipulated in a number of different ways. One way we have already considered concerns the particular task subjects are required to perform (e.g. lexical decision, naming, relatedness judgements). A second possibility is that there may be differences in the goals of the reading task. For example, one may approach the materials differently if one is required to later recall the gist of the material versus recognise the details of what was earlier read. In the latter case, it is possible that subjects would be more likely to emphasise lexical-level pathways, whereas in the earlier case it is more likely that individuals will emphasise semantic-level integration-based information. Of course, this is an important possibility when one considers on-line reading measures to examine word processing effects, under different levels of instructions. Again, the notion is that the task demands should drive the pathways that are crucial in achieving the goals of the task.

Another possible means by which subjects may be directed to use different sources of information to perform a task may be particularly subtle. Specifically, a manipulation of contextual constraint might be sufficient to direct subjects to multiple sources of information. If this is the case, then researchers interested in estimating context effects on lexical processing would first be required to consider carefully the appropriateness of their stimuli for assessing context effects (for additional discussions of such concerns see Kellas et al., 1991; Oden & Spira, 1983; Olson, 1970; Tanenhaus & Donnenworth-Nolan, 1984). This point may be made more salient by considering contexts which contain ambiguous words. Such contexts can be constructed with varying levels of constraint. That is, context preceding an ambiguous word can range from making only a single interpretation of the homograph sensible to making no particular interpretation any more sensible than another. When no interpretation is biased by the context, it would not be surprising to find that multiple interpretations of the homograph are initially activated. Hence, this would be interpreted as support for a lexical view of processing. What *would* be surprising from a lexical (or context-independent) view of processing is if a constraining context resulted in *only* the contextually appropriate meaning becoming activated. Consider that, in most experiments, subjects do not typically receive only one or two contexts. Rather, sentence primes are presented, as is also typical of most single word prime studies, one after another as separate and unrelated from one another. It is likely that this is atypical of most common or normal reading situations. However, the degree to which the sentences, as a group, are similar to one another may be readily picked up by subjects. If a substantial proportion of the sentences are not sufficiently constraining, subjects may become biased to process all sentences (whether constraining or not) at a lexical level. This would probably result in findings inconsistent with discourse views of processing.

One of the most frequently cited examples of context-independent processing may be found in Swinney (1979). Although Swinney presented context stimuli auditorially, subjects responded to visually presented targets (i.e. lexical decision). Ignoring, for the sake of argument, any potential concerns with the task used (after all, lexical decisions were, at the time, widely held to accurately reflect initial processing), there is an aspect of the study that suggests the contexts used may not have been adequately constraining. Consider the Swinney (1979) study in light of a distinction made by Hess et al. (1995) between local and global context, reviewed above (see also Foss & Speer, 1990). Given the following stimulus: *Rumour had it that, for years, the government building had been plagued with problems. The man was not surprised when he found several spiders, roaches, and other bugs in the corner of his room*, it could be argued that there is a clearly biasing local context that favours the "insect" meaning of *bugs*. Indeed, Swinney found that responses to "ant" were facilitated relative to responses to an unrelated word (i.e. "sew"). Swinney also reported, however, that responses to words related to the contextually inappropriate sense of *bugs* (e.g. "spy") were facilitated as well. One might argue that the "spy" sense of the ambiguous word may also be related to the global context which includes government buildings, problems, and covert listening devices. Given that responses to "spy" were no longer facilitated when presented three syllables following the presentation of *bugs*, the conclusion that multiple meanings were initially activated independent of context seems as likely as the alternative conclusion that all meanings relevant to the global and local context of *bugs* were initially facilitated, as a *result* of context. The reason that the "insect" meaning of *bugs* was maintained for a longer duration may be due to a specific interpretation of the *type* of problems which plagued the government building (that is, the "spy" interpretation of "plagued with problems" was made less likely by later context). Although this was only one stimulus of many used by Swinney (1979), it is worthwhile to note that the notion of priming from local and global sources of context has been investigated for unambiguous words (e.g. Foss & Ross, 1983; Hess et al., 1995).

In sum, there appears to be converging evidence for two loci of contextual constraint during reading: one locus is more locally driven and may result from intralexical processes, and a second locus appears to involve a higher level text integration process. Thus, when attention is directed toward message-level processing pathways in reading, it appears that readers do not merely rely on a single source of constraint but rather will capitalise on any source that increases the efficiency of extracting message-level information.

Conclusions

One of the goals of word recognition research is to uncover the processes involved from the journey from visual features to meaning. Clearly, the evidence that has accumulated from isolated word recognition has been highly influential in a number of distinct areas of cognitive science. Moreover, there have been considerable theoretical advances that have led to important insights into the processes involved in language processing.

We believe that now that researchers are armed with this battery of observations and theoretical perspectives, it is time to more closely tune models of word recognition to processes involved in the common use of words in language processing tasks, i.e. as the core building blocks of message-level information. Hence, we have adopted a processing pathway approach in this chapter which emphasises the pathways that are engaged by the reader in pursuit of a given task's goals. In an attempt to illustrate the importance of attentional selection of appropriate processing pathways, we have reviewed a number of the important issues in word recognition research, which included orthographic neighbourhood effects, spelling to sound correspondence effects, number of meaning effects, syntactic context effects, and discourse-level context effects. In each of these sections, we have attempted to demonstrate that one may find quite different patterns of data depending upon the types of goals engaged by a given task. The relevance of certain classes of variables apparently come to the forefront when a given task places high priority on those variables.

Figure 2.1 presents a possible way to conceptualise the manner in which attentional control might influence different characteristics of the lexical processing system. As shown, there are a number of distinct processing modules that deal with computations involved in orthography, phonology, meaning, syntax, and higher level discourse integration. The goals of the task drive attention to relevant processing dimensions. For example, in naming, the attentional control system would increase the influence of the computations between orthography and phonology. On the other hand, the demands of lexical decision performance might place a high priority on the computations between orthographic and meaning level modules. Finally, if the goal of the task is reading comprehension, then attentional control would increase the priority of computations of the syntactic-, meaning-, and discourse-level modules. It is important to note that we do not intend to argue that attention *totally* controls the output of a given set of task relevant processing modules. Rather, we believe that the *degree of influence* of these different processing modules will be affected by attentional control. Moreover, we would argue that the degree of control will be dependent upon such factors as (1) the ability to maintain a representation of experimental task demands across time, and also (2) the prepotent strength of the pathways between the modules. For example, with respect to this later factor, although participants can avoid naming the stimulus word in a Stroop study, it is very difficult, but not impossible (see Besner, Stoltz, & Boutilier, 1997), to completely eliminate the influence from the relatively strong pathways that compute meaning from orthography.

In conclusion, we would like to argue for a flexible lexical processing system in which attentional systems modulate the importance of the

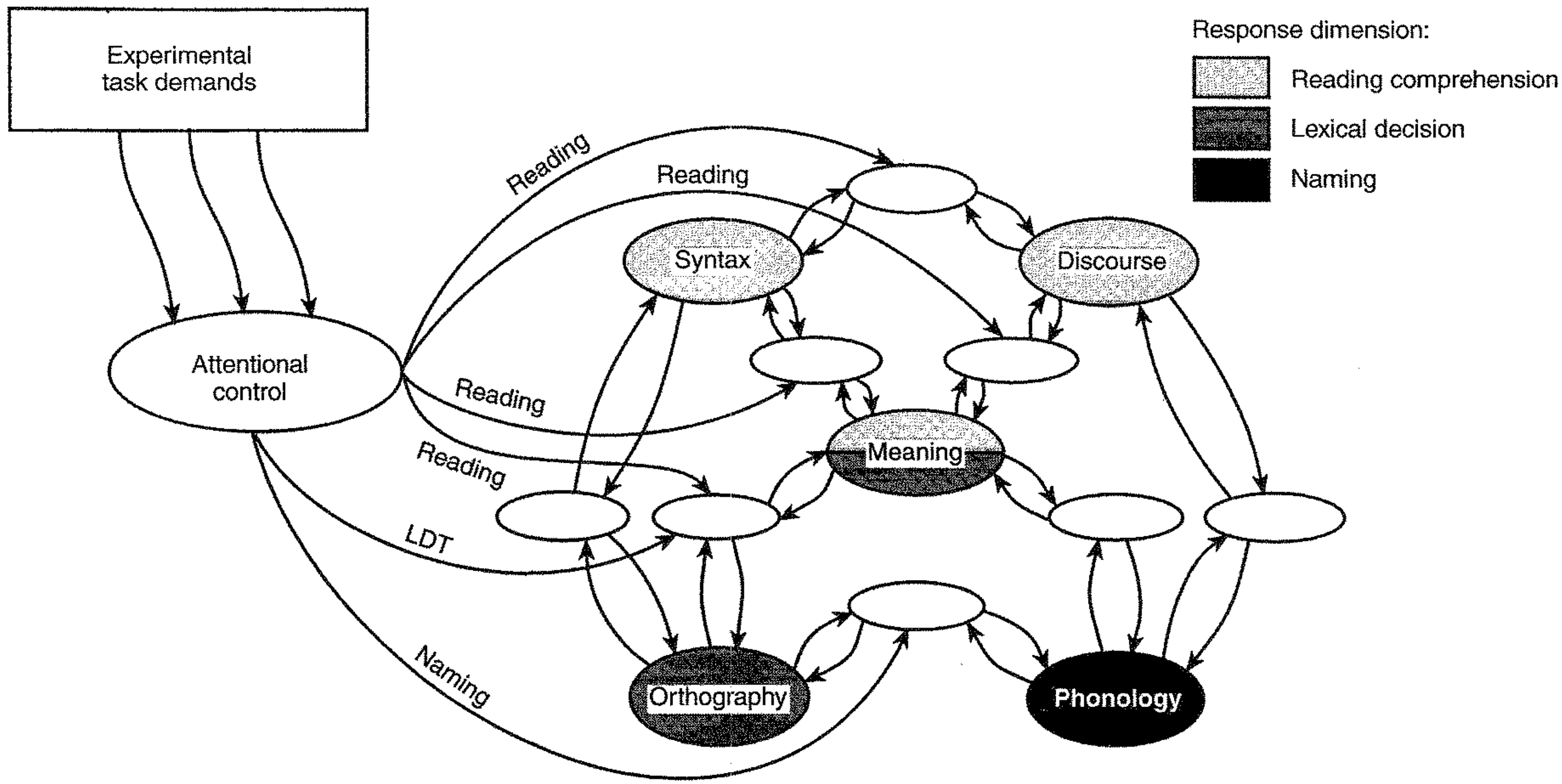

FIG. 2.1 An attentional control framework for lexical processing tasks.

numerous codes available for a given word. Until this flexibility is acknowledged and eventually implemented, we believe that one may be misled by the relevance of a given set of results to the goal of this enterprise, i.e. to understand how words serve as vehicles for communication.

REFERENCES

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

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

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

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

Auble, P.M., & Franks, J.L. (1983). Sentence comprehension processes. *Journal of Verbal Learning and Verbal Behavior, 22*, 395–405.

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

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

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

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

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

Balota, D.A., & Ferraro, F.R. (1996). Lexical, sublexical, and implicit memory processes in healthy young and healthy older adults and in individuals with dementia of the Alzheimer type. *Neuropsychology, 10*, 1–14.

Balota, D.A., Ferraro, F.R., & Connor, L.T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P.J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187–222), Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

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

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

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

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

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

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

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

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

Besner, D. (1983). Basic decoding components in reading: Two dissociable feature extraction processes. *Canadian Journal of Psychology, 37*, 429–438.

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

Besner, D., Stolz, J.A., & Boutilier (1997). The Stroop effect and the myth of automaticity. *Psychonomic Bulletin & Review, 4*, 221–225.

Besner, D., & Swan, M. (1982). Models of Lexical access in visual word recognition. *Quarterly Journal of Experimental Psychology, 34A*, 313–325.

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

Bohemier, G.L. (1991). *Neighborhood and frequency effects on word recognition in a foveal presentation of text.* Unpublished master's thesis, State University of New York at Binghamton, NY.

Bohemier, G.L. (1994). *Neighborhood effects and the role of neighborhood rhyme.* Unpublished dissertation, State University of New York at Binghamton, NY.

Bohemier, G.L., & Inhoff, A.W. (in prep.). *Effects of neighborhood density and dominance on lexical access measured across tasks.*

Borowski, R. & Masson, M.E.J. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 63–85.

Britt, M.A., Perfetti, C.A., Garrod, S., & Rayner, K. (1992). Parsing in discourse: Context effects and their limits. *Journal of Memory and Language, 31*, 293–314.

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

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

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies in information processing* (pp. 151–216). San Diego, CA: Academic Press.

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

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

Conrad, C. (1974). Context effects in sentence comprehension: A study of the subjective lexicon. *Memory and Cognition, 2*, 130–138.

Cortese, M.J., Simpson, G.B., & Woolsey, S. (1997). Effects of association and imageability on phonological mapping. *Psychonomic Bulletin & Review, 4*, 226–231.

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

Dooling, D.J. (1972). Some context effects in the speeded comprehension of sentences. *Journal of Experimental Psychology, 93*, 56–62.

Duffy, S.A., Henderson, J.M., & Morris, R.K. (1989). The semantic facilitation of lexical access during sentence processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 791–801.

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

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

Fera, P., Joordens, S., Balota, D.A., & Ferraro, F.R. (1992, November). *Ambiguity in meaning and phonology: Effects on naming*. Poster presented at the annual meeting of the Psychonomic Society, St Louis, MO.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*, 348–368.

Fischler, I., & Bloom, P.A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior, 18*, 1–20.

Fodor, J.A. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.

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

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

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

Foss, D.J. (1970). Some effects of ambiguity upon sentence comprehension. *Journal of Verbal Learning and Verbal Behavior, 9*, 699–706.

Foss, D.J., & Jenkins, C.M. (1973). Some effects of context on the comprehension of ambiguous sentences. *Journal of Verbal Learning and Verbal Behavior, 12*, 577–589.

Foss, D.J. & Ross, J.R. (1983). Great expectations: Context effects during sentence processing. In G.B. Flores d'Arcais & R.J. Jarvella (Eds.), *The process of language understanding* (pp. 169–191). New York: John Wiley.

Foss, D.J. & Speer, S.R. (1990). Global and local context effects in sentence processing. In R.R. Hoffman & D.S. Palermo (Eds.), *Cognition and the symbolic processes: Applied and ecological perspectives* (pp. 115–139). Hillsdale, NJ: Lawrence Erlbaum Associates Ltd.

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

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

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

Glanzer, M., & Ehrenreich, S.L. (1979). Structure and search of the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 18*, 381–398.

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

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

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

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

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

Hess, D.J., Foss, D.J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General*, *124*, 62–82.

Holmes, V.M., Arwas, R., & Garrett, M.F. (1977). Prior context and the perception of lexically ambiguous sentences. *Memory and Cognition*, *5*, 103–110.

Inhoff, A.W., & Topolski, R. (1994). Use of phonological codes during eye fixations in reading and in on-line and delayed naming tasks. *Journal of Memory and Language*, *33*, 689–713.

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

Jastrzembski, J.E., & Stanners, R.F. (1975). Multiple word meanings and lexical search speed. *Journal of Verbal Learning and Verbal Behavior*, *14*, 534–537.

Kawamoto, A.H., Farrar, W.T., & Kello, C.T. (1994). When two meanings are better than one: Modeling the ambiguity advantage using a recurrent distributed network. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1233–1247.

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

Kellas, G., Paul, S.T., Martin, M., & Simpson, G.B. (1991). Contextual feature activation and meaning access. In G.B. Simpson (Ed.), *Understanding word and sentence* (pp. 47–71). Amsterdam: North-Holland.

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

Lukatela, G., & Turvey, M.T. (1994a). Visual lexical access is initially phonological: 1. Evidence from associative priming by words, homophones, and pseudohomophones. *Journal of Experimental Psychology: General3*, *123*, 107–128.

Lukatela, G., & Turvey, M.T. (1994b). Visual lexical access is initially phonological: 2. Evidence from phonological priming by homophones and pseudohomophones. *Journal of Experimental Psychology: General*, *123*, 331–353.

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*, 570–590.

MacDonald, M.C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, *9*, 157–201.

MacDonald, M.C., Pearlmutter, N.J., & Seidenberg, M.S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.

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

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

Masson, M.E.J. (1991). A distributed memory model of context effects in word identification. In D. Besner & G.W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 233–263). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

McClelland, J.L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 3–36). Hove, UK: Lawrence Erlbaum Associates Ltd.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: 1. An account of basic findings. *Psychological Review, 86,* 287–330.

Midgley-West, L. (1979). *Phonological encoding and subject strategies in skilled reading.* Unpublished dissertation, Birkbeck College, University of London.

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

Monsell, S. (1990). Frequency effects in lexical tasks: Reply to Balota and Chumbley. *Journal of Experimental Psychology: General, 119,* 335–339.

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

Monsell, S., Patterson, K.E., Graham, A., Hughes, C.H., & Milroy, R. (1992). Lexical and sublexical translation of spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 180,* 452–467.

Morris, R.K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 92–103.

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

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

Neely, J.H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Neill, W.T., Hilliard, D.V., & Cooper, E.A. (1988). The detection of lexical ambiguity Evidence for context-sensitive parallel access. *Journal of Memory and Language, 27,* 279–287.

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18,* 5–19.

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

Oden, G.C., & Spira, J.L. (1983). Influence of context on the activation and selection of ambiguous word senses. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 35,* 51–64.

Olson, D.R. (1970). Language and thought: Aspects of a cognitive theory of semantics. *Psychological Review, 77,* 257–273.

Onifer, W., & Swinney, D.A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory and Cognition, 9* 225–236.

O'Seaghdha, P.G. (1989). The dependence of lexical relatedness effects on syntactic connectedness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 73–87

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

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

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

Patterson, K.E., Graham, N., & Hodges, J.R. (1994). Reading in dementia of the Alzheimer's type: A preserved ability? *Neuropsychology, 8*, 395–407.

Paul, S.T., Kellas, G., Martin, M., & Clark, M.B. (1992). The influence of contextual features on the activation of ambiguous word meanings. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 703–717.

Pearlmutter, N.J., & MacDonald, M.C. (1992). Plausibility and syntactic ambiguity resolution. In *Proceedings of the 14th annual conference of the Cognitive Science Society* (pp. 498–503). Hillsdale NJ: Lawrence Erlbaum Associates Ltd.

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 767–780.

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

Perfetti, C.A., Beverly, S., Bell, L.C., & Hughes, C. (1987). Phonemic knowledge and learning to read: A longitudinal study of first grade children. *Merrill-Palmer Quarterly, 33*, 283–319.

Petersen, S.E., Fox, P.T., Posner, M.I., Mintun, M., & Raichle, M.E. (1990). Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience, 1*, 153–170.

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

Posner, M.I. (1978). *Chronometric explorations of the mind.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

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

Pugh, K.R., Rexer, K., Peter, M., & Katz, L. (1994). Neighborhood effects in visual word recognition: Effects of letter delay and nonword context difficulty. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 639–648.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8*, 21–30.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior, 22*, 358–374.

Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 779–790.

Rayner, K., Garrod, S., & Perfetti, C.A. (1992). Discourse influences during parsing are delayed. *Cognition, 25*, 109–139.

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

Rayner, K., Sereno, S.C., Lesch, M.F., & Pollatsek, A. (1995). Phonological codes are automatically activated during reading: Evidence from an eye movement priming paradigm. *Psychological Science, 6*, 26–32.

Richardson, J.T.E. (1976). The effects of stimulus attributes upon latency of word recognition. *British Journal of Psychology*, *67*, 315–325.

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

Rubenstein, H., Lewis, S.S., & Rubenstein, M.A. (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.

Schuberth, R.E., & Eimas, P.D. (1977). Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, 73, 27–36.

Schustack, M.W., Ehrlich, S.F., & Rayner, K. (1987). Local and global sources of contextual facilitation in reading. *Journal of Memory and Language*, *26*, 322–340.

Schvaneveldt, R.W., Meyer, D.E., & Becker, C.A. (1976). Lexical ambiguity, semantic context, and visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 243–256.

Schwanenflugel, P.J., Harnishfeger, K.K., & Stowe, R.W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, *27*, 499–520.

Seidenberg, M.S., & Harm, M. (1995, November). *Division of labor and masking in a multi-component model of word recognition*. Paper presented at the 36th annual meeting the Psychonomic Society, Los Angeles, CA.

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

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*, 234–237.

Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, *14*, 489–537.

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

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.

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

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

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

Sharkey, N.E. (1990). A connectionist model of text comprehension. In D.A. Balota, G.B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 487–514). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Sharkey, A.J.C., & Sharkey, N.E. (1992). Weak contextual constraints in text and word priming. *Journal of Memory and Language*, *31*, 543–572.

Simpson, G.B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior*, *20*, 120–136.

Simpson, G.B. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin*, *96*, 316–340.

Simpson, G.B. (1994). Context and the processing of ambiguous words. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* pp. 359–374). San Diego, CA: Academic Press.

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

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

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

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

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–661.

Swinney, D.A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645–659.

Swinney, D.A., & Hakes, D.T. (1976). Effects of prior context upon lexical access during sentence comprehension. *Journal of Verbal Learning and Verbal Behavior, 15*, 681–689.

Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language, 27*, 324–340.

Tabossi, P., Colombo, L., & Job, R. (1987). Accessing lexical ambiguity: Effects of context and dominance. *Psychological Research, 49*, 161–167.

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

Tanenhaus, M.K., & Donnenworth-Nolan, S. (1984). Syntactic context and lexical access. *The Quarterly Journal of Experimental Psychology, 36A*, 649–661.

Till, R.E., Mross, E.F., & Kintsch, W. (1988). Time course of priming for associate and inference words in a discourse context. *Memory and Cognition, 16*, 283–298.

Trueswell, J.C., & Tanenhaus, M.K. (1991). Tense, temporal context and syntactic ambiguity resolution. *Language and Cognitive Processes, 6*, 303–338.

Trueswell, J.C., Tanenhaus, M.K., & Garnsey, S.M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*, 285–318.

Ungerleider, L.G., & Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, M.A. Goodale, & R.J.W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–580). Cambridge, MA: MIT Press.

Van Orden, G.C. (1987). A ROWS is a ROSE: Spelling, sound and reading. *Memory and Cognition, 15*, 181–198.

Van Orden, G.C. (1991). Phonologic mediation is fundamental to reading. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 77–103). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Van Orden, G.C., & Goldinger, S.D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 1269–1291.

Van Petten, C., & Kutas, M. (1987). Ambiguous words in context: An event-related potential analysis of the time course of meaning activation. *Journal of Memory and Language, 26*, 188–208.

Verstaen, A., Humphreys, G.W., Olson, A., & d'Ydewalle, G. (1995). Are phonemic effects in backward masking evidence for automatic prelexical phonemic activation in visual word recognition? *Journal of Memory and Language, 34*, 335–356.

Vu, H., Kellas, G., & Paul, S.T. (1998). Sources of sentence constraint on lexical ambiguity resolution. *Memory & Cognition, 26*, 979–1001.

West, R.F., & Stanovich, K.E. (1986). Robust effects of syntactic structure on visual word processing. *Memory and Cognition*, *14*, 104–112.

Wright, B., & Garrett, M. (1984). Lexical decision in sentences: Effects of syntactic structure. *Memory and Cognition*, *12*, 31–45.

Zevin, J.D., & Balota, D.A. (1999). *Priming and attentional control of lexical and sublexical pathways during naming*. Manuscript submitted for publication.

Zola, D. (1984). Redundancy and word perception during reading. *Perception and Psychophysics*, *36*, 277–284.