BEHAVIORAL AND BRAIN SCIENCES (2003) **26,** 445–526
Printed in the United States of America

The E-Z Reader model of eye-movement control in reading: Comparisons to other models

Erik D. Reichle
Department of Psychology, University of Pittsburgh, Pittsburgh, PA 15260
reichle@pitt.edu

Keith Rayner
Department of Psychology, University of Massachusetts, Amherst, Amherst, MA 01003
rayner@psych.umass.edu

Alexander Pollatsek
Department of Psychology, University of Massachusetts, Amherst, Amherst, MA 01003
pollatsek@psych.umass.edu

Abstract: The E-Z Reader model (Reichle et al. 1998; 1999) provides a theoretical framework for understanding how word identification, visual processing, attention, and oculomotor control jointly determine when and where the eyes move during reading. In this article, we first review what is known about eye movements during reading. Then we provide an updated version of the model (E-Z Reader 7) and describe how it accounts for basic findings about eye movement control in reading. We then review several alternative models of eye movement control in reading, discussing both their core assumptions and their theoretical scope. On the basis of this discussion, we conclude that E-Z Reader provides the most comprehensive account of eye movement control during reading. Finally, we provide a brief overview of what is known about the neural systems that support the various components of reading, and suggest how the cognitive constructs of our model might map onto this neural architecture.

Key words: attention; eye-movement control; E-Z Reader; fixations; lexical access; models; reading; regressions; saccades

1. Introduction

Reading is a complex skill that involves the orchestration of many different stages of information processing. As the eyes move across the printed page, the visual features of the text are converted into orthographic and phonological patterns, which are then used to guide further language processing so that the content of the text can be understood. In this target article, we will compare different models that try to account for how eye movements are controlled in reading. We will not review all of the models that have been proposed to explain various aspects of reading. Instead, we will discuss only those models that have attempted to explain the interface between vision and low-level aspects of language processing; that is, models that specify some combination of the following components of reading: eye-movement control, visuospatial attention, and/or the visual processing of words.[1] Not surprisingly, we will argue that the model that we implemented, E-Z Reader[2] (Reichle et al. 1998; 1999), does a better job of accounting for a wide range of data than does its competitors. However, we will also point out some shortcomings of the model.

The remainder of this article will be organized into five major sections. First, in section 2, we will briefly review some important findings regarding eye movements in reading; within this section we will describe some findings that we believe a model of eye-movement control should be able to accommodate. Second, in section 3, we will provide an overview of the E-Z Reader model, including an updating of the model (E-Z Reader 7). Third, in section 4, we will provide an overview of other models of eye-movement control in reading (including discussions of the pros and cons of the models compared to E-Z Reader). Fourth, in section 5 we will discuss future directions and ways that we intend to extend the E-Z Reader model. In this section, we will also discuss a possible mapping between model components and neurophysiological mechanisms. Finally, we will provide some concluding comments in section 6.

2. Eye movements in reading

Any discussion of models of eye-movement control must begin with a brief overview of eye movements during reading. In this section, we will describe what is known about eye movements during reading as background material. The following topics will be discussed: (1) saccades and fixations; (2) visual acuity; (3) saccade latency; (4) the acquisition of information during eye fixations; (5) perceptual span; (6) parafoveal preview effects; (7) regressions; (8) eye-movement control (where to fixate next and when to

 0140-525X/03 $12.50

move the eyes); (9) measures of processing time. It is not our intention to provide a complete and comprehensive review of each of these topics, as our primary purpose in this article is to compare different models of eye movement control in reading. The interested reader is invited to consult Rayner (1998) for a more complete review of each of the nine topics discussed in this section.

2.1. Saccades and fixations

Contrary to our subjective impression, the eyes do not move smoothly across the printed page during reading. Instead, the eyes make short and rapid movements, called *saccades* (Erdmann & Dodge 1898; Huey 1908), that typically move them forward about 6–9 character spaces, although there is considerable variability (Rayner 1978; 1998). Since the distribution of saccade sizes, measured in number of character spaces, is largely independent of visual angle when the number of character spaces is held constant (Morrison & Rayner 1981; O'Regan 1983), virtually all studies of reading use number of character spaces as the appropriate metric. Saccades take 20–50 msec to complete, depending upon the length of the movement, and virtually no visual information is extracted during eye movements (Ishida & Ikeda 1989; Wolverton & Zola 1983). Between saccades, the eyes remain stationary for brief periods of time (typically 200–250 msec) called *fixations* (Erdmann & Dodge 1898; Huey 1908). Because visual information is only extracted from the printed page during fixations, reading is similar to a slide show in which short segments of text are displayed for approximately a quarter of a second. It is important to note that there is considerable variability in both saccade length and fixation duration. Some saccades only move the eyes a single character, whereas others are as large as 15–20 characters (although such long saccades typically follow *regressions,* or backward movements to previous parts of the text, and place the eyes beyond the place from which the regression was initiated). Likewise, some fixations are shorter than 100 msec and others are longer than 400 msec (Rayner 1978; 1998). Much of this variability apparently is related to the ease or difficulty involved in processing the currently fixated text.

Erik D. Reichle is Assistant Professor of Psychology at the University of Pittsburgh, and a Research Scientist at the Learning Research and Development Center. He is the author of several papers on eye-movement control in reading, and is on the Editorial Board of the journal *Psychological Review* (beginning 2004). His research interests include word identification, attention, and eye-movement control.

Keith Rayner is Distinguished University Professor of Psychology at the University of Massachusetts, Amherst. He has published extensively in the areas of: eye movements, reading, language processing, scene perception, and visual search. He is the author of *The Psychology of Reading* (with A. Pollatsek) and has edited five others volumes. He is the former Editor of the *Journal of Experimental Psychology: Learning, Memory, and Cognition* (1990–1995) and current Editor of *Psychological Review* (beginning 2004).

Alexander Pollatsek is Emeritus Professor of Psychology at the University of Massachusetts, Amherst. He is co-author of the book *The Psychology of Reading* (Rayner & Pollatsek 1989) and has served on the Editorial Board of numerous journals. His research deals with reading, eye movements, phonological processing, scene perception, and decision making, and he has published widely in these areas.

2.2. Visual acuity

One of the reasons that the eyes are constantly moving in reading is that there are severe limits to how much visual information can be processed during a fixation. Visual acuity is maximal in the center of the retina and rapidly decreases towards the periphery, and fine visual discriminations can only be made within the *fovea,* or central 2° of vision. As a result, the visual features that make up individual letters can be encoded only from a very narrow window of vision. The practical significance of this limitation is that it is necessary to fixate most words so that they can be identified. Indeed, there is considerable evidence that a word becomes increasingly difficult to identify as the angular disparity between the fovea and the retinal image of a word increases (Rayner & Bertera 1979; Rayner & Morrison 1981). Explaining how the reader deals with this limited acuity is one constraint on any model of eye movements.

2.3. Saccade latency

A second kind of constraint on any model of reading stems from the "race" between the processes identifying words and the need to plan a saccade early enough in a fixation so that reading can carry on at about 300 words per minute. On the one hand, experiments in which subjects move their eyes to visual targets indicate that the *saccadic latency,* or the time needed to plan and execute a saccade, is approximately 180–250 msec (Becker & Jürgens 1979; Rayner et al. 1983b). This suggests that the decision to make a saccade is often made within the first 100 msec of a fixation. However, this is seemingly at odds with the intuitively appealing idea that word recognition is a major contributor to driving eye movements during reading because most estimates indicate that lexical access requires 100–300 msec to complete (Rayner & Pollatsek 1989; Schilling et al. 1998; Sereno et al. 1998;). It is thus not immediately obvious how the identification of one word can be the signal to begin planning a saccade to the next. Indeed, early theories of eye movements in reading (Bouma & de Voogd 1974; Kolers 1976) posited that word identification was too slow to be the engine driving eye movements.

2.4. The acquisition of information during reading

During saccades, vision is suppressed so that the information needed for reading is acquired only during fixations (Ishida & Ikeda 1989; Wolverton & Zola 1983). Furthermore, reading proceeds quite smoothly if text is available for processing for only the first 50–60 msec of a fixation prior to the onset of a masking pattern (Ishida & Ikeda 1989; Rayner et al. 1981). This does not mean that words are identified within 50 msec, but rather, that the information that is needed for reading gets into the processing system within 50–60 msec.

2.5. Perceptual span

One solution to the quandary of how word identification can be a signal to move the eyes is that words can be partially processed in the *parafovea,* or region of the retina that extends five degrees on either side of the fovea. McConkie and Rayner (1975) demonstrated the importance of parafoveal processing using an *eye-contingent display change technique,* called the *moving-window paradigm,* which is illustrated in Figure 1. In this paradigm, the letters outside of a "window" spanning a given number of character spaces is distorted in some way (e.g., replaced with Xs). By varying the size of the window and making its location contingent upon where the reader is looking, it is possible to determine the *perceptual span,* or region from which useful visual information can be encoded. With alphabetic text (like English), readers can progress at a more-or-less normal rate when the window extends 14–15 character spaces to the right (McConkie & Rayner 1975; Rayner 1986; Rayner & Bertera 1979; Rayner et al. 1982; Den Buurman et al. 1981) and 3–4 character spaces to the left of the fixation point (McConkie & Rayner 1976; Rayner et al. 1980). However, word encoding probably does not extend more than 7–8 characters to the right of fixation (McConkie & Zola 1984; Rayner et al. 1982; Underwood & McConkie 1985); beyond this distance, only low-spatial frequency information about letter shape (e.g., descenders vs. ascenders) and word length is extracted from the page. The left-right asymmetry reflects covert attention and is language specific; with Hebrew text (which is read from right to left), the perceptual span extends asymmetrically to the left of fixation (Pollatsek et al. 1981).

Four other points about the perceptual span are relevant. First, the perceptual span does not extend below the line that is currently being read (Inhoff & Briihl 1991; Inhoff & Topolski 1992; Pollatsek et al. 1993); readers focus their attention on the line that they are currently reading. Second, studies using various eye-contingent display change techniques have revealed that the size of the span is fairly constant for readers of similar alphabetic orthographies (such as English, French, and Dutch; see Rayner [1998] for further details). Third, characteristics of the writing system influence not only the asymmetry of the span, but also the overall size of the perceptual span. Thus, the span is smaller for Hebrew than for English (Pollatsek et al. 1981) because Hebrew is a more densely packed language than English. And it is much smaller for writing systems like Japanese (Ikeda & Saida 1978; Osaka 1992) and Chinese (Inhoff & Liu 1998) that have ideographic components and hence are even more densely packed than Hebrew. Fourth, the perceptual span is not hardwired, but rather seems to be attention-based. The fact that there is an asymmetry due to the direction of the writing system is consistent with the span being attention-based. In fact, Pollatsek et al. (1982) found that the perceptual span of Israeli readers who were bilingual in Hebrew and English had opposite asymmetries when reading the two languages. Furthermore, Rayner (1986) found that the span was smaller for beginning readers than skilled readers and that the span got smaller when children with four years of reading experience were given text that was too difficult for them. Analogous to this finding, Henderson and Ferreira (1990; see also Inhoff et al. 1989; Kennison & Clifton 1995; Schroyens et al. 1999) found that the span got smaller when the fixated word was difficult to process. Finally, Balota et al. (1985) found that readers obtained more information to the right of fixation when the upcoming word was highly predictable from the preceding text.

```
A. Normal Text

the link between eye movements and language
      *
the link between eye movements and language
            *
the link between eye movements and language
                   *

B. Moving Window: 2 Words

xxx link between xxx xxxxxxxxx xxx xxxxxxxx
      *
xxx xxxx between eye xxxxxxxxx xxx xxxxxxxx
            *
xxx xxxx xxxxxxx eye movements xxx xxxxxxxx
                   *

C. Moving Window: 4 Spaces Left & 14 Spaces Right

xxe link between eye xxxxxxxxx xxx xxxxxxxx
      *
xxx xxxx between eye movemexxx xxx xxxxxxxx
            *
xxx xxxx xxxxxxn eye movements and xxxxxxxx
                   *
```

Figure 1. The moving-window paradigm. Panel A shows the positions of three successive fixations (indicated by the asterisks) in a normal line of text. Panels B and C illustrate how a "window" of normal text is displayed contingent upon where the eyes are currently looking. Panel B shows a two-word moving window; that is, both the fixated word and the word to the right of fixation are displayed normally, and all of the letters in the remaining words are replaced by Xs. In Panel C, the window extends four character spaces to the left of fixation and 14 character spaces to the right of fixation.

2.6. Parafoveal preview effects

Consistent with the findings of the last section, it has been demonstrated that orthographic (Balota et al. 1985; Binder et al. 1999; Rayner 1975) and phonological (Pollatsek et al. 1992) processing of a word can begin prior to the word being fixated. These results indicate that, during normal reading, the *parafoveal preview* of a word can reduce the duration of the subsequent fixation on the word, which is one measure of the time needed for identification (Schilling et al. 1998). Surprisingly, neither semantic (Altarriba et al. 2001; Rayner et al. 1986) nor morphological (Kambe, in press; Lima 1987; Lima & Inhoff 1985) information extracted from the parafovea appears to be of any benefit when the word is later fixated.[3] Furthermore, parafoveal preview benefit is not due to retention of visual featural information, as the case of all the letters can change from fixation to fixation with virtually no disruption to the reading process (McConkie & Zola 1979; Rayner et al. 1980). Instead, the source of the preview benefit seems to be due to abstract letter codes and phonological codes (see Rayner [1998] for a review). However, parafoveal information can produce word skipping (i.e., the word is not fixated) because words that can be identified in the parafovea do not have to be fixated and can therefore be skipped. Many experiments (Balota et al. 1985; Binder et al. 1999; Ehrlich & Rayner 1981; Rayner et al. 2001; Rayner & Well 1996; Schustack et al. 1987) have demonstrated that predictable words are skipped more than unpredictable words and that short function words (like "the") are skipped more than content words (O'Regan 1979; 1980; Gautier et al. 2000). When words are skipped, there is some evidence suggesting that the durations of the fixations preceding and following the skip are inflated (Pollatsek et al. 1986; Reichle et al. 1998).[4]

2.7. Regressions

One indicator of the inherent difficulty of reading (even for skilled readers) is that 10–15 percent of the saccades move the eyes back to previous parts of the text. These backward movements, called *regressions,* are thought to result both from problems with linguistic processing and from oculomotor error. The hypothesis that regressions can be caused by difficulties in linguistic processing is perhaps most clearly supported by the finding that regressions can be induced with structurally difficult "garden path" sentences; because such sentences often lead to incorrect syntactic analyses, readers often make regressions back to the point of difficulty and then re-interpret the sentence (Frazier & Rayner 1982). The idea that regressions are sometimes due to simple motor error is supported by the finding that, when the eyes fixate near the end of a word, they often move back a few character spaces (O'Regan 1990). This presumably happens because the eyes overshot their intended target (near the middle of the word) and a second fixation location affords a better place from which to see the word. This interpretation is consistent with the finding that identification is most rapid if a word is fixated just to the left of its center, on the *optimal viewing position* (Clark & O'Regan 1999; O'Regan 1990; 1992b; O'Regan et al. 1984).

2.8. Eye-movement control

Numerous studies have attempted to determine the characteristics of the mechanisms that control eye movements during reading. There are two different activities that must be explained: (1) What determines where the reader decides to look next? and (2) What determines when the reader moves his/her eyes (either forward or backward in the text)? Although there is not total consensus on these issues, there is some evidence to suggest that decisions about where to fixate next and when to move the eyes are made somewhat independently (Rayner & McConkie 1976; Rayner & Pollatsek 1981). The earliest unambiguous demonstration that the duration of the current fixation and the length of the next saccade are computed online was provided by Rayner and Pollatsek (1981). They varied physical aspects of the text randomly from fixation to fixation and found that the behavior of the eyes mirrored what was seen on a fixation. In their first experiment, they used the moving window paradigm described above and varied the size of the window randomly from fixation to fixation, and found that saccade length varied accordingly. Thus, if the window on the current fixation was small, the eyes only moved a few characters, while if it was large, the eyes moved further. In their second experiment, they delayed the onset of text in the fovea via a mask that appeared at the beginning of a fixation (with the time the mask was on varying randomly from fixation to fixation) and found that fixation durations were adjusted accordingly. In addition, the manipulations affected saccade length and fixation duration independently; in the first experiment, saccade length was affected, but fixation duration was not, whereas in the second experiment, fixation duration was affected, but saccade length was not. Thus, while the decisions about where to fixate next and when to move the eyes may sometimes overlap (see Rayner et al. 2000), there is reason to believe the two decisions are made somewhat independently.

2.8.1. Where to fixate next. Decisions about where to fixate next seem to be determined largely by low-level visual cues in the text, such as word length and the spaces between words. Five types of results are consistent with this claim. First, saccade length is influenced by the length of the fixated word and the word to the right of fixation (Blanchard et al. 1989; O'Regan 1979; 1980; Rayner 1979; Rayner & Morris 1992). Second, when readers do not have information about where the spaces are between upcoming words, saccade length decreases and reading is slowed considerably (McConkie & Rayner 1975; Morris et al. 1990; Pollatsek & Rayner 1982; Rayner et al. 1998a). Third, although there is some variability in where the eyes land on a word, readers tend to make their first fixation about halfway between the beginning and the middle of the word (Rayner 1979; McConkie et al. 1988; 1989; 1991; Vitu 1991b). Recently, Deutsch and Rayner (1999) demonstrated that the typical landing position in Hebrew words is likewise between the beginning (i.e., right-most end) and middle of a word. Rayner (1979) originally labeled this prototypical location the *preferred viewing location.* This position where the eyes typically land in a word is different from the *optimal viewing location,* which is the location in the word at which recognition time is minimized. According to O'Regan and Levy-Schoen (1987), the optimal viewing position is a bit to the right of the preferred viewing location, closer to the center of the word. Fourth, while contextual constraint influences skipping, in that highly predictable words are skipped more than unpredictable words (Balota et al. 1985; Ehrlich & Rayner 1981), contextual constraint has little influence on where the eyes land in a word (Rayner et al. 2001).[5] Finally, the landing position on a word is modulated by the *launch site* (McConkie et al. 1988; Radach & Kempe 1993; Radach & McConkie 1998; Rayner et al. 1996) because the landing position varies as a function of the distance from the prior fixation. As the launch site moves further from the target word, the distribution of landing positions shifts to the left and becomes more variable (see Fig. 2).

2.8.2. When to move the eyes. The ease or difficulty associated with processing a word primarily influences when the eyes move. Although a case can be made that low-level non-linguistic factors can also influence the decision about when to move the eyes, the bulk of the evidence suggests that linguistic properties of words are the major determiner of when to move. A very robust finding is that readers look longer at low-frequency words than at high-frequency words (Altarriba et al. 1996; Henderson & Ferreira 1990; 1993; Hyönä & Olson 1995; Inhoff & Rayner 1986; Just & Carpenter 1980; Kennison & Clifton 1995; Lavigne et al. 2000; Raney & Rayner 1995; Rayner 1977; Rayner & Duffy 1986; Rayner & Fischer 1996; Rayner & Raney 1996; Rayner et al. 1996; 1998a; Sereno 1992; Vitu 1991b; Vitu et al. 2001). There are three additional points with respect to this finding that are relevant. First, there is a *spillover* effect associated with fixating a low-frequency word; that is, fixation time on the next word is inflated (Rayner & Duffy

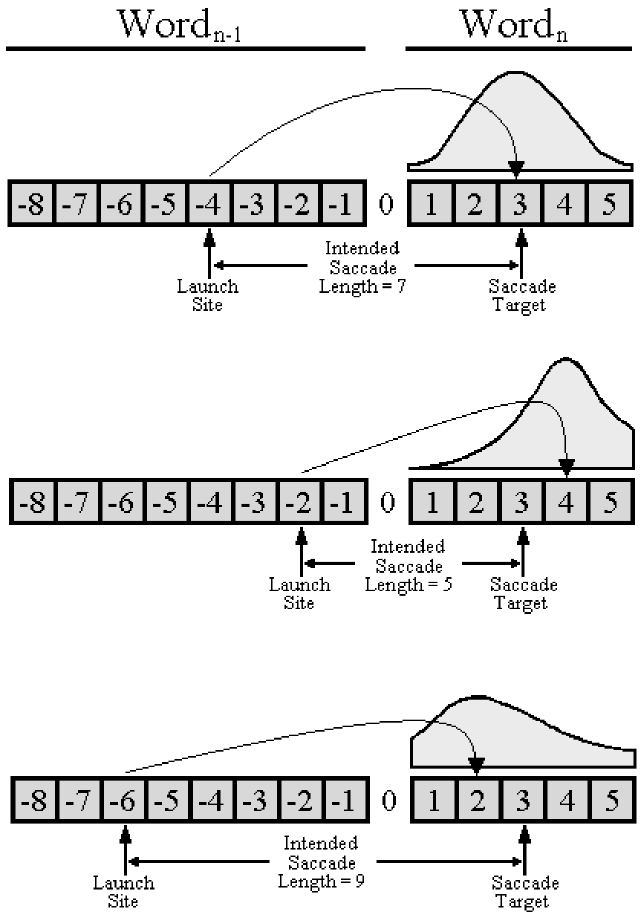

Figure 2. Landing site distribution as a function of the saccade length between the launch site (word$_{n-1}$) and intended saccade target (word$_n$). In all three panels, the launch site and target words are depicted by rectangles, with character spaces represented by numbers (as per convention, the space to the left of word$_n$ is denoted by a zero). The landing site distributions are approximately Gaussian in shape. Although the distributions are centered near the middle of the saccade targets, the oculomotor system is biased towards making saccades approximately seven character spaces in length. This bias results in a systematic range error; that is, the eyes tend to overshoot close targets and undershoot more distant targets. For example, in the middle panel, the intended saccade target is five character spaces from the launch site, so that (on average) the eyes overshoot their intended target, thereby causing the landing site distribution to shift towards the end of word$_n$. In the bottom panel, the opposite happens: The eyes undershoot their target, causing the landing site distribution to shift towards the beginning of word$_n$.

1986). Second, although the duration of the first fixation on a word is influenced by the frequency of that word, the duration of the prior fixation is not (Carpenter & Just 1980; Henderson & Ferreira 1993; Rayner et al. 1998a). Third, high-frequency words are skipped more than low-frequency words, particularly when they are short and the reader is fixated close to the beginning of the word (O'Regan 1979; Rayner et al. 1996).

A second important finding is that there is a predictability effect on fixation time in addition to a frequency effect. Words that are highly predictable from the preceding context are fixated for less time than are words that are not so constrained (Altarriba et al. 1996; Balota et al. 1985; Binder et al. 1999; Ehrlich & Rayner 1981; Inhoff 1984; Lavigne et al. 2000; Rayner & Well 1996; Rayner et al. 2001; Schustack et al. 1987; Zola 1984). Generally, the strongest effects of predictability are not as large as those of the strongest frequency effects. Also, as we noted above, predictability has a strong effect on word skipping: Words that are highly predictable from the prior context are skipped more often than words that are not so constrained.

2.9. Measures of processing time

To investigate the components of reading, researchers typically have subjects read sentences or passages of text while an eye tracker interfaced with a computer records the locations and durations of individual fixations. Because an average college-level reader can read approximately 300 words per minute (Rayner & Pollatsek 1989), this technique produces a staggering amount of data. Accordingly, the data are usually reduced to *word-based measures,* which are across-subject averages that reflect how often and for how long individual words are fixated. A number of word-based measures are standard (Inhoff & Radach 1998; Liversedge & Findlay 2000; Rayner 1998; Rayner et al. 1989; Starr & Rayner 2001). The first is *gaze duration,* which is defined as the sum of all fixations on a word, excluding any fixations after the eyes have left the word (i.e., including only *refixations* before the eyes move on to another word). Gaze duration is usually averaged only over words that are not skipped during the initial encounter (or *first pass*) through that region of text. Two other common measures are *first-fixation duration* and *single-fixation duration.* The former is the duration of the first fixation on a word (again conditional on the word being fixated during the first pass through the text), while the latter is the average fixation duration on words that are fixated exactly once during the first pass. These indices are typically reported along with indices of how often a word was fixated, which reflect the probability of a word being skipped, fixated once, and fixated more than once before moving to another word. Often, the *total time* (the sum of all fixations on the word, including regressions back to the word) is also reported.

The word-based measures provide a complete record of where and when fixations occurred. These two aspects (where vs. when) also provide a useful framework for organizing a discussion of reading models because much of the controversy surrounding reading concerns the determinants of where and how long the eyes remain fixated. The models that have been developed to explain eye-movement control form a continuum, extending from models in which eye movements are determined primarily by oculomotor factors (*oculomotor models*) to those in which eye movements are guided by some form of cognitive control (*processing models*). Prior to comparing different models, we will discuss our model, E-Z Reader (Pollatsek et al. 1999b; Rayner et al. 1998c; 2000; Reichle et al. 1998; 1999; Reichle & Rayner 2001) in some detail. We will also provide an updated version of the model (E-Z Reader 7).

3. E-Z Reader

E-Z Reader is a processing model, and extends the earlier work of Morrison (1984). Morrison drew much of the inspiration for his model from the work of Becker and Jürgens (1979) and McConkie (1979). McConkie (1979) suggested that, during reading, visual attention progressed across a line of text until the limitations of the visual system made it difficult to extract further lexical information; once this point of difficulty has been established, attention shifts and an eye movement is programmed and subsequently initiated, sending the eyes to the problematic location. Although elegantly simple, the model was soon discarded due to problems in defining and explaining what the point of difficulty was, how it might be computed, and whether it could be computed soon enough to be of any use in skilled reading (Rayner & Pollatsek 1989).

The limitations inherent in McConkie's (1979) early model of eye-movement control led Morrison (1984) to propose a model in which the movement of the eyes was a function of successful processing. According to Morrison, the identification of word$_n$ (i.e., the word that is currently being fixated) causes the attention "spotlight" (Posner 1980) to move to word$_{n+1}$, which in turn causes the oculomotor system to begin programming a saccade to word$_{n+1}$. If the program finishes before word$_{n+1}$ is identified, then the saccade will be executed and the eyes will move to word$_{n+1}$. However, if word$_{n+1}$ is identified before the program finishes, the saccade to word$_{n+1}$ may be cancelled. Cancellation can occur some of the time when attention shifts to word$_{n+2}$ while word$_n$ is fixated. In this case, the oculomotor system begins programming a saccade to word$_{n+2}$, which overrides the program to move the eyes to word$_{n+1}$ if the new program interrupts the old program soon enough. Thus, according to Morrison, attention moves serially, from word to word, whereas saccades can be programmed in parallel.

Morrison's (1984) assumption about the parallel programming of saccades followed Becker and Jürgens' (1979) demonstration that saccadic programming is completed in two stages: an initial, labile stage that is subject to cancellation, and an ensuing, non-labile stage in which the program cannot be cancelled. Their results suggested that if the oculomotor system begins programming a saccade while another saccadic program is in its labile stage of development, then the first program is aborted. However, if the second program is initiated while the first saccadic program is in its non-labile stage, then both saccades will be executed, which typically results in a very short fixation between the two saccades.

With these simple assumptions, Morrison (1984) was able to provide an elegant account of both frequency effects and parafoveal preview effects: Because short frequent words are more easily identified in the parafovea than long infrequent words, the former tend to be fixated for less time (and skipped more often) than the latter. Despite its successes, however, Morrison's model cannot explain refixations because the strictly serial attention shifts mean that each word is either fixated exactly once or is skipped.

More fundamentally, however, because Morrison's model posits both that processing of words is strictly serial and that attention shifting is time-locked to word identification, the model is unable to handle some simple and robust phenomena in reading. The first, as we noted above, is that one often gets "spillover" effects due to word frequency (e.g., Rayner & Duffy 1986). That is, lower-frequency words often not only cause longer fixations on that word (word$_n$), but also lengthen either gaze durations and/or first fixations on the succeeding word (word$_{n+1}$). According to Morrison's model, this shouldn't happen because attention doesn't shift until word$_n$ has been processed. Because parafoveal processing on word$_{n+1}$ begins after this attention shift, the amount of information extracted from word$_{n+1}$ before it is fixated will only be a function of how long it takes to program and execute the saccade, and will not vary as a function of the frequency of word$_n$. As a result, Morrison's model predicts no delayed effects of word frequency (or any other delayed effects of word processing difficulty). A related phenomenon (Henderson & Ferreira 1990; Kennison & Clifton 1995) is that the benefit gained through parafoveal preview decreases as foveal processing becomes more difficult (e.g., because the fixated word is lower frequency). By essentially the same argument as above, Morrison's model predicts that this shouldn't happen because parafoveal preview time is only a function of the latency of moving the eyes after covert attention has shifted.

There are at least three ways to circumvent the limitations of Morrison's (1984) model. The first is to add the assumption that if word identification is not completed by a processing deadline, attention does not shift to the next word, but instead remains on the current word, resulting in a refixation (Henderson & Ferreira 1990; Sereno 1992). This leads to the prediction (which has not been supported; Rayner et al. 1996; Schilling et al. 1998) that the first of two fixations should be longer than single fixations because the former reflect cases in which the processing deadline must have been reached. The second solution is to simply assume that difficulties with higher-order linguistic processing somehow cause the eyes to remain on the current word (Pollatsek & Rayner 1990; Rayner & Pollatsek 1989). Unfortunately, how this happens has not been well specified. Finally, a third way to avoid the shortcomings of Morrison's proposal is to assume that word identification is completed in two stages. This last approach is instantiated by E-Z Reader, which is discussed next.

3.1. Overview of the E-Z Reader model

E-Z Reader, like other processing models, makes the basic assumption that ongoing cognitive (i.e., linguistic) processing influences eye movements during reading. Because the model was not intended to be a deep explanation of language processing, it does not account for the many effects of higher-level linguistic processing on eye movements (for reviews, see Rayner 1998; Rayner & Sereno 1994; Rayner et al. 1989). Although this is clearly a limitation, it should also be noted that many of these effects typically occur when the reader is having difficulty understanding the text that is being read, such as when a reader makes a regression to re-interpret a syntactically ambiguous "garden path" sentence (Frazier & Rayner 1982). The model can therefore be viewed as the "default" reading process. That is, we view the process of identifying words to be the forward "driving engine" in reading, as the process of knitting the words into larger units of syntax or meaning would be too slow (whether successful or not) to be a signal to decide how and when to move the eyes forward for skilled readers. Thus, we posit that higher-order processes intervene in eye-movement control only when "something is wrong" and either send a signal to stop moving forward or a signal to execute a regression. Hence, we view E-Z Reader as an explanation of what happens during reading when higher-level linguistic processing is running smoothly and doesn't

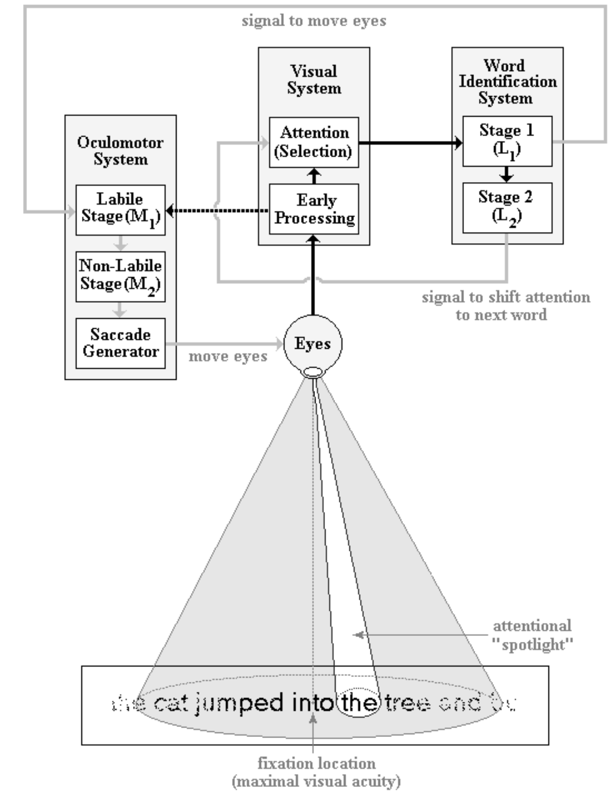

Figure 3. A schematic diagram of *E-Z Reader 7*. Visual features on the printed page are projected from the retina to an early stage of visual processing, which then proceeds at a rate that is modulated by visual acuity limitations. The low-spatial frequency information (e.g., word boundaries) is used by the oculomotor system to select the targets of upcoming saccades. High-spatial information is passed on to the word identification system, which, though attentional selection, allows individual words to be identified by the word identification system. The first stage of lexical processing (L_1) signals the oculomotor systems to begin programming a saccade to the next word. The completion of the second stage of word identification (L_2) causes attention to shift to the next word. Saccadic programming is thus decoupled from the shifts of attention. Saccadic programming is completed in two stages: The first, labile stage (M_1) can be cancelled by the initiation of subsequent programs; the second, non-labile stage (M_2) is not subject to cancellation. Saccades are executed immediately after the non-labile stage of saccadic programming has been completed. Black lines represent the flow of visual information, with the dashed line representing the low-spatial frequency information that is used by the oculomotor system to select the target locations of upcoming saccades. The gray lines represent signals that are propagated among the various components of the model (e.g., the signal to shift attention).

intervene. One implication of this is that the model currently does not explain inter-word regressions.

Like its immediate predecessors (see Reichle et al. 1998; 1999), E-Z Reader 7 consists of a small number of perceptual-motor and cognitive processes that determine when and where the eyes move during reading. Figure 3 is a schematic diagram showing the flow of control among these processes. As is evident in the figure, the central assumptions of the model are that: (1) a stage of word identification is the signal to move the eyes; and (2) attention is allocated from one word to the next in a strictly serial fashion. Notice, however, that both visual encoding limitations and oculomotor constraints also play central roles in the moment-by-moment control of eye movements during reading. In the discussion that follows, we will describe the specific assumptions of our model and how they are related to four major cognitive and perceptual-motor systems: visual processing, word identification, attention, and oculomotor control.

3.1.1. (Early) visual processing. Visual features from the printed page are projected from the retina to the visual cortex so that the objects on the page (i.e., the individual words) can be identified. The earliest stages of visual processing are thought to be pre-attentive in that the features that make up individual words are not fully integrated into perceptual wholes (Lamme & Roelfsema 2000; Wolfe & Bennett 1996). This processing is not instantaneous, with neural transmission from retina to brain taking approximately 90 msec to complete.

In our model, the preceding ideas are formalized by including the early processing stage in the visual system, which, though pre-attentive, is subject to visual acuity constraints (see Fig. 3). The duration of this early visual processing stage, $t(V)$, is a free parameter that corresponds to the base time needed for neural transmission to propagate from the retina to those cortical and subcortical areas that mediate early visual processing. To keep this assumption psychologically plausible, the value of $t(V)$ was set equal to 90 msec. However, because the rate of this early stage of processing is modulated by visual acuity, the rate at which a word is encoded is inversely proportional to both its length and its mean distance from the point of fixation. More specifically, during each fixation, the amount of early visual processing (in msec) that is completed on each word in the visual field is determined by:

$$\text{visual processing} = t/(\varepsilon^{\Sigma|i\_letter - i\_fixation|/N}) \qquad (1)$$

In Equation 1, t is the duration of the fixation (in msec), N is the number of letters in a word being processed, and ε (= 1.08) is a free parameter[6] that modulates the effects of the spatial disparity between each word's letters and the fixation location (i.e., the center of the fovea). Thus, the time needed to encode a word increases as the distance between its center and the fovea increases. Moreover, the time needed to encode a word also increases with its length because the individual letters of long words will (on average) be further away from the point of fixation than will the individual letters of short words.[7] One interesting implication of this equation is that the early visual processing of a word will be most rapid if the word is fixated near its center because a fixation on a word's center will minimize the mean spatial deviations between the fixation and each of the word's letters. This property is also consistent with evidence that word identification is most rapid if the word is fixated near its center (or optimal viewing position; O'Regan 1990; 1992b; O'Regan & Lévy-Schoen 1987; Vitu et al. 1990) and provides one explanation for why the eyes are seemingly directed towards this location during reading (see Shillcock et al. 2000). It also allows the model to account for length effects (i.e., the finding that long words take longer to identify than short words; Just & Carpenter 1980).

Early visual processing is important for two other reasons. First, it is necessary to obtain the word-boundary information that is needed to program saccades to upcoming words. This is denoted in Figure 3 by the dashed arrow that extends from early visual processing to the labile stage of saccadic programming. This arrow represents the flow of low-spatial frequency information that is acquired in the visual periphery (e.g., word boundaries, the presence/absence of ascenders and descenders, etc.). The oculomotor system uses this information to program saccades to upcoming words. Second, early visual processing provides the information that is subsequently used by higher-level visual areas to focus the attention "spotlight" and identify individual words. Word identification (which is discussed in the next section) must therefore wait until the early visual encoding of that word has been completed.

3.1.2. Word identification. The process of identifying a word begins as soon as attention is focused on that word. This identification process is then completed in two stages, reflecting early and late stages of lexical processing. The first stage corresponds to being at (or at least close to) the identification of the orthographic form of the word. We assume that this is not full lexical access, as the phonological and semantic forms of the word are not yet fully activated. We labeled this process the "familiarity check" (i.e., f) in earlier versions of the model, but in E-Z Reader 7 it is simply referred to as the first stage of lexical access (i.e., L_1).

The second stage of word identification involves the identification of a word's phonological and/or semantic forms so as to enable additional linguistic processing. This second stage, therefore, more or less corresponds to what is typically thought to be "lexical access." In prior versions of our model, this stage of word identification was called the "completion of lexical access" (i.e., *lc*). To avoid confusion, however, we will simply refer to this process as the second stage of lexical access (i.e., L_2) in E-Z Reader 7.

The distinction between early and late stages of lexical processing has precedent in the literature; indeed, our distinction was partly motivated by the *activation-verification model* of lexical access (Paap et al. 1982). The two models are broadly consistent if one conceptualizes the first stage of lexical access as a "quick and dirty" assessment of whether or not word identification is imminent, and the second stage as being the actual act of identification. As indicated in Figure 3, this distinction is also important because the two stages of lexical processing play unique functional roles: The completion of the first stage of lexical access causes the oculomotor system to begin programming the next saccade, while the completion of the second stage causes the "spotlight" of attention to shift to the next word. Thus, in E-Z Reader, saccadic programming is de-coupled from the shifting of attention.

As with earlier versions of our model, the time (in msec) required to complete the first stage of lexical access on a word, $t(L_1)$, is a linear function of the natural logarithm of the word's normative frequency of occurrence in printed text and its predictability within a given sentence context. The mathematical statement of this relationship is given by Equation 2:

$$t(L_1) = [\beta_1 - \beta_2 \ln(\text{frequency})]\,(1 - \theta\ \text{predictability}) \qquad (2)$$

In Equation 2, β_1 and β_2 (= 228 and 10 msec, respectively) are free parameters that control how a word's normative frequency (number of occurrences per million, as tabulated by Francis & Kučera [1982]) affect lexical processing time. This time is also modulated by the right-hand term, in which the free parameter θ (= 0.5) attenuates the degree to which a word's predictability in a specific sentence context (as estimated using cloze task probabilities) attenuates the lexical processing time.[8] In all of the simulations reported below, the actual times needed to complete the first stage of lexical processing was found by sampling from gamma distributions having means equal to $t(L_1)$ and standard deviations equal to 0.18 of their means.

The completion of the first stage of lexical processing of a word has two immediate consequences in the model: (1) it cues the oculomotor system to begin programming a saccade to the next word (the details of how the oculomotor system does this will be discussed in detail below), and (2) it initiates further processing of the word. Because all (or at least most) of the orthographic coding has been completed in L_1, the time required to complete the second stage of lexical processing, L_2, is more influenced by a word's predictability. This distinction is reflected in Equation 3:

$$t(L_2) = \Delta[\beta_1 - \beta_2 \ln(\text{frequency})]\,(1 - \text{predictability}) \qquad (3)$$

As in Equation 2, the free parameters β_1 and β_2 control the degree to which a word's frequency of occurrence affects the time necessary to process the word, but this quantity is attenuated by the free parameter $\Delta(= 0.5)$. Note that, in contrast to L_1, a word's predictability fully affects L_2; that is, words that can be predicted with complete certainty within a given sentence context will require no time in this second stage (i.e., if predictability = 1, then $t(L_2)$ = 0 msec). Such cases reflect the situation when top-down information has already fully activated the semantic and phonological codes given reasonable corroborating input from orthography. As was the case with the first stage of lexical processing, the actual process durations were sampled from gamma distributions.

Finally, it should be mentioned that – by adding the early visual processing stage to E-Z Reader 7 – the minimal time needed to identify words in the model is very plausible. Given the parameter values reported above, for example, the mean time needed to identify the word "the" (the most frequent word in English text) when it is centrally fixated and in a completely predictable context is 148 msec, while the time needed to identify the lowest frequency words in completely unpredictable contexts is 432 msec. In contrast, E-Z Reader 6 predicted minimal and maximal mean word identification times of 16 and 278 msec, respectively. E-Z Reader 7 thus predicts word identification latencies that are much more in line with the best available estimates: 150–300 msec (Rayner & Pollatsek 1989).

3.1.3. Attention. A central, and perhaps the most contentious, assumption of E-Z Reader is that covert shifts of attention occur serially, from one word to the next, as each word is identified in turn and then integrated into the discourse representation. By "attention," though, we do not mean spatial orientation; rather, we refer to the process of integrating features that allows individual words to be identified. The separation between these two types of attention has considerable precedence in the literature (LaBerge 1990). For example, Treisman (1969) distinguished between *input selection,* or spatial orientation, and *analyzer selection,* or feature integration. This distinction is important because spatial orientation shifts towards the targets of upcoming saccades (Hoffman & Subramaniam 1995; however, see Stelmach et al. 1997), which in E-Z Reader occur whenever the oculomotor system uses the low-spatial frequency information provided by the visual processing stage to program a saccade (see the dashed line in Fig. 3). These shifts in spatial orientation, however, are decoupled from the shifts in attention (i.e., analyzer selection) that precede lexical processing.

Attention is allocated serially during reading because readers need to keep word order straight (Pollatsek & Rayner 1999). By shifting the focus of attention from one word to the next, readers identify and process each word in its correct order. Although the results of several recent experiments (Inhoff et al. 2000b; Kennedy 1998; 2000; Kennedy et al. 2002; Starr & Inhoff, in press) suggest that properties of two words (particularly visual/orthographic properties) can sometimes be encoded in parallel, we suspect that this does not usually occur in normal reading (see Rayner et al. 2003c, for an extended discussion of these issues). The reason for this is that much of the information that is conveyed by language (both written and spoken) is heavily dependent upon word order.

Furthermore, by decoupling eye movements from attention, our model can also explain aspects of eye-movement control that Morrison's (1984) model could not. For example, E-Z Reader can explain why parafoveal preview benefit decreases as foveal processing difficulty increases (Henderson & Ferreira 1990; Kennison & Clifton 1995). If the eyes are on word$_n$, parafoveal processing of word$_{n+1}$ begins, not with completion of the first stage of lexical processing of word$_n$, but after the completion of second stage. Because parafoveal processing of word$_{n+1}$ ends (by definition) with the onset of the saccade to word$_{n+1}$, more time will remain for parafoveal processing of word$_{n+1}$ when word$_n$ is easy to process (e.g., high-frequency). This is depicted in Figure 4: The time required to complete L_1 and L_2 on word$_n$ increases as its normative frequency decreases (see Equations 2 and 3). Because the saccadic latency is not modulated by word frequency, a saccade will (on average) occur 240 msec (i.e., the mean saccadic latency) after the completion of L_1. This means that, with everything else being equal, the amount of time available to process word$_{n+1}$ in the parafovea will increase as the amount of time needed to process word$_n$ decreases.

In the model, the serial-allocation-of-attention assumption is instantiated as follows: The completion of the second stage of lexical processing on word$_n$ causes attention to shift to word$_{n+1}$, at which point the first stage of lexical processing begins on word$_{n+1}$ when pre-processing of word$_{n+1}$ is complete.[9] The identification of one word thus causes the focus of attention to shift so that the word-identification system can begin identifying the next word (see Fig. 3).

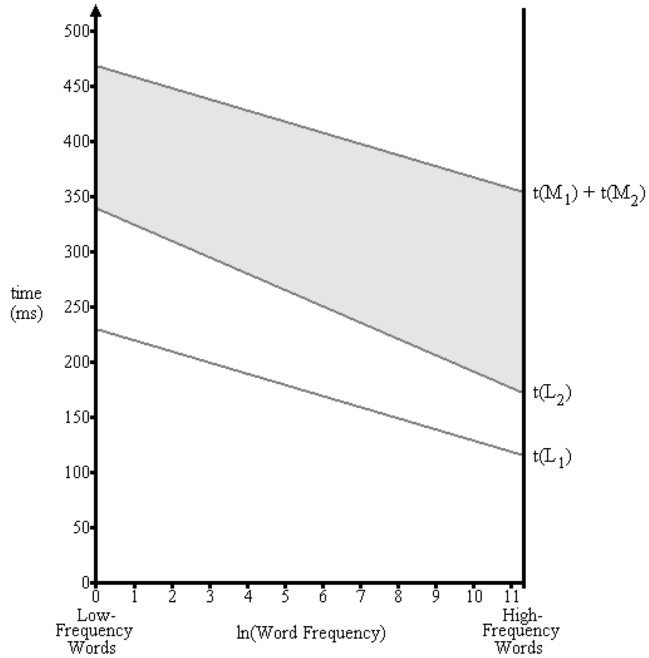

Figure 4. A diagram showing how parafoveal preview benefit is modulated by normative word frequency. The bottom line represents the time required to complete the first stage of lexical processing, $t(L_1)$, as a function of the natural logarithm of word$_n$'s token frequency. The middle line represents the time required to complete the second stage of lexical processing, $t(L_2)$, on word$_n$. Finally, the top line represents the saccadic latency, or time required to initiate a saccade from word$_n$ to word$_{n+1}$. On average, the saccadic latency requires a constant $t(M_1) + t(M_2)$ ms to complete (starting from the point in time when the first stage of lexical processing on word$_n$ has been completed). In E-Z Reader, parafoveal preview begins as soon as word$_n$ has been identified and attention has shifted to word$_{n+1}$. The parafoveal preview is therefore limited to the duration of the interval (depicted by the shaded area in the figure) between $t(L_2)$ and $t(M_1) + t(M_2)$. Notice that, because the relative disparity between $t(L_1)$ and $t(L_2)$ increases as the frequency of word$_n$ decreases, the duration of the parafoveal preview decreases with the frequency of word$_n$.

3.1.4. Oculomotor control. Saccadic programming in E-Z Reader is completed in two stages: an early, labile stage (M_1) that is subject to cancellation by subsequent programs, and a later, non-labile stage (M_2) that is not subject to cancellation. This assumption was motivated by demonstrations that a saccade to a first target can be cancelled by the presentation of a second to-be-fixated target if the second target is presented within approximately 230 msec after the first; after this time, both targets are typically fixated in sequence (Becker & Jürgens 1979). A considerable amount of subsequent research has supported this distinction between labile and non-labile stages of saccadic programming (Leff et al. 2001; McPeek et al. 2000; Molker & Fischer 1999; Vergilino & Beauvillain 2000).

During the first (labile) stage of saccadic programming, the eye-movement system is simply engaged (or made ready) so that it can begin programming an eye movement. The system then computes the distance between the current fixation location and the location of the saccade target (i.e., the *intended saccade length*). Thus, although the target location is represented in terms of spatial coordinates, the saccadic program is represented in terms of a distance metric. This is necessary because the distance that is specified by the saccadic program must ultimately be converted into the appropriate amount of force that has to be exerted (by the extraocular muscles) to execute the actual movement. The labile stage of programming therefore consists of two sub-stages: (1) general system preparation, followed by (2) a location-to-distance transformation, in which the spatial location of the upcoming saccade target is converted into the necessary saccade length. In E-Z Reader, the time needed to complete the labile programming stage is a random deviate that is sampled from a gamma distribution having a mean equal to a free parameter, $t(M_1)$, with each of the two aforementioned sub-stages subsuming half of this time.

An important part of our model is that, when a saccade program is in the labile stage, it is subject to cancellation by a subsequent saccadic program. If the second program is initiated during the system preparation sub-stage of the first program, then whatever amount of preparation has been done to ready the oculomotor system will also be applicable to the second program, so that it will be completed more rapidly than it otherwise would be. If, however, the second program is initiated somewhat later, during the first program's location-to-distance transformation sub-stage, then whatever processing has been done to specify the distance of the first saccade will not apply to the second because the target locations (and hence the distances) of the two saccades are different. This means that the second program will always require a minimal amount of time to finish – the time necessary to convert the spatial location of the saccade target into the intended saccade length.

During the second (non-labile) stage of programming, the command to move the eyes a particular direction and distance is communicated to the motor system. At this point, an intended saccade is obligatory, and cannot be cancelled or modified by subsequent programs. As with the labile stage of programming, the time needed to complete the non-labile stage of programming is sampled from a gamma distribution, with the mean of this distribution being equal to a free parameter, $t(M_2)$. Upon completing the non-labile stage of programming, the saccade is executed immediately.

In E-Z Reader 7, the mean times needed to complete the labile, $t(M_1)$, and non-labile, $t(M_2)$, stages of saccadic programming were set equal to 187 and 53 msec, respectively. To keep the model as simple as possible, the saccade durations were set equal to a fixed value: $t(S) = 25$ msec.[10] Our saccadic-programming parameter values are consistent with estimates from simple saccade latency tasks (Becker & Jürgens 1979; McPeek et al. 2000; Rayner et al. 1983b). It should be noted, however, that these values are in fact estimates of the *minimal* time required to initiate a saccade, often to *pre-specified* targets; in the context of reading text, therefore, the average saccadic latency may be slightly longer in duration than would be suggested by these previous estimates.

Let us examine these assumptions using five key situations in reading. The first situation (shown schematically in Fig. 5A) is the simplest: $Word_n$ is fixated, an eye movement is programmed to $word_{n+1}$, and no subsequent eye-movement command is made while this program is in its labile stage. The program therefore enters its non-labile stage, and an eye movement is made to $word_{n+1}$.

Now consider a second situation (Fig. 5B): $Word_n$ is fixated, a program to fixate $word_{n+1}$ is initiated, but while the oculomotor system is being readied, a second program (to move the eyes to $word_{n+2}$) is initiated. In this case, the program to fixate $word_{n+1}$ is cancelled, and the saccade leaving $word_n$ will move the eyes to $word_{n+2}$ (i.e., $word_{n+1}$ will be skipped). Whatever time elapsed in preparing the oculomotor system to program the first saccade will also be subtracted from the time that would otherwise be necessary to program the second saccade, thereby allowing it to be completed more rapidly than would otherwise be the case. Moreover, because situations like the one just described tend to occur when $word_{n+1}$ is processed rapidly, the model successfully predicts that skipping is more likely to occur whenever $word_{n+1}$ is high frequency, predictable from prior context, and/or short.

Now let's consider a situation (Fig. 5C) that is similar to the one just described: $Word_n$ is fixated, and a program to fixate $word_{n+1}$ is initiated. However, just as the labile stage of this program is about to finish (i.e., the location-to-distance transformation is almost complete), the oculomotor system begins programming a saccade to $word_{n+2}$. As in the previous situation, the program to fixate $word_{n+1}$ will be cancelled, and the eyes will again go directly from $word_n$ to $word_{n+2}$. Because the saccade length specified by the second saccade program is different from the length specified by the first, however, the duration of the second program's labile stage will include the time needed to recompute the distance between the location of the current fixation location and that of the new saccade target. The second program's labile stage will therefore be reduced, but only by the amount of time needed for general system preparation; that is, the second program's labile stage will equal the time needed to complete its location-to-distance transformation.

Finally, let us consider the situations depicted in Panels D and E of Figure 5: In both cases, $word_n$ is fixated, the program to fixate $word_{n+1}$ is initiated, and then (after some time) this program goes into its second, non-labile stage. At this time, a second program (to move the eyes to $word_{n+2}$) is initiated. In both of the situations depicted in Panels D and E, the program to fixate $word_{n+1}$ will run to completion, and the eyes will move from $word_n$ to $word_{n+1}$. However, in Panel D, the second program does not really benefit (i.e., it requires the full amount of time to be completed) because there was no ongoing labile program when the second program was initiated. Because the first saccade is actually executed while the second program is in its early, system-preparation phase, though, the second program's labile stage does not have to be re-started. In contrast, Panel E shows what happens when the first saccade is executed while the second program is in its location-to-distance transformation phase: Because the eyes are now fixated on $word_{n+1}$, the relative distance between the location of the current fixation and that of the saccade target ($word_{n+2}$) must be re-calculated. This means that the location-to-distance transformation has to be re-started, which extends the time needed to complete this part of the second saccade's labile programming.

Our discussion of saccadic programming so far has focused largely on the time needed to program the saccades, and has only addressed the question of where the eyes move at a fairly coarse level (i.e., at the level of individual words). As McConkie and his colleagues demonstrated (1988; 1991), saccades are prone to both systematic and random error. The effects of these sources of error are not negligible, and have been an oft-cited reason for the claim that the control of eye-movements during reading is primarily mediated by fairly low-level visual and oculomotor constraints (e.g., visual acuity limitations, systematic motor error, etc.; see O'Regan 1990; 1992b; O'Regan & Lévy-Schoen 1987; Reilly & O'Regan 1998). It is therefore important to specify how the model handles the effects of saccadic error.

Our assumptions regarding the oculomotor system are based on McConkie et al.'s (1988; 1991) data and analyses. In fact, we more or less directly incorporated their views of saccadic error into our model. In the model, saccades are directed towards the optimal viewing position of the words being targeted. However, these saccades are subject to both systematic and random error, so that, on average, saccades will deviate from their intended targets. More formally, each saccade is the sum of three components:

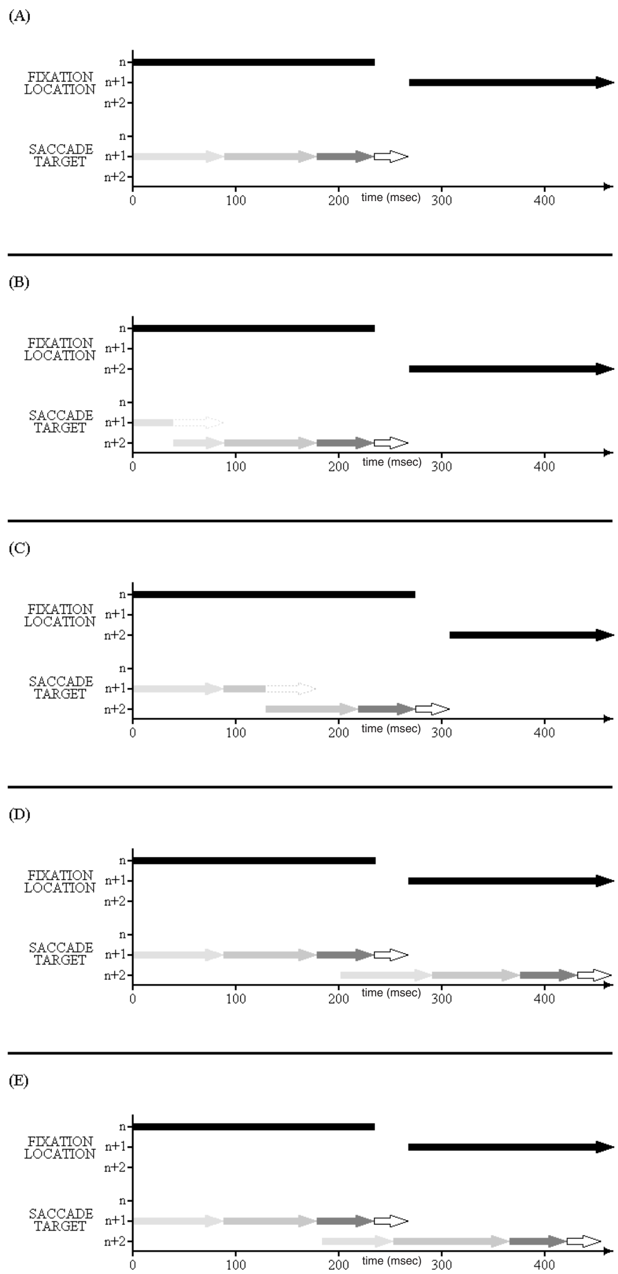

Figure 5. This diagram shows E-Z Reader 7's oculomotor control assumptions and how these assumptions affect saccadic programming in five common situations that occur during reading. In all of the panels, time (in msec) is represented along the horizontal axis, the black horizontal bars indicate the word (*n*, *n*+1, or *n*+2) that is being fixated at each given point in time, and the arrows represent the various stages of saccadic programs that are being directed towards specific word targets (*n*, *n*+1, or *n*+2). The light gray arrows represent the general preparation component of the first, labile programming stage, the medium gray arrows represent the location-to-distance transformation phase of the labile programming stage, and the dark gray arrows represent the second, non-labile stage of programming. The white arrows represent the actual saccades. In Panel A, one program follows another, and the eyes move in sequence from word$_n$ to word$_{n+1}$ to word$_{n+2}$. In Panels B and C, a program is initiated while another, labile program is in progress; in these situations, the first program is cancelled, and the eyes move from word$_n$ to word$_{n+2}$ (skipping word$_{n+1}$). Finally, in Panels D and E, the second program is initated while the first program is in its non-labile stage; in these situations, the first program runs to completion, and the eyes move in sequence from word$_n$ to word$_{n+1}$ to word$_{n+2}$.

$$\text{saccade} = \text{intended saccade length} + \text{SRE} + \text{RE} \qquad (4)$$

In Equation 4, the *intended saccade length* is the distance (in character spaces) between the current fixation (i.e., launch site) and the middle of the word that is the saccade target, and *SRE* and *RE* are the systematic and random error, respectively. The *SRE* emerges from the fact (at least for readers of English) that the oculomotor system "prefers" to make saccades that are seven character spaces in length. Saccades that are intended to be longer than seven character spaces tend to undershoot their targets, whereas saccades that are intended to be shorter than seven character spaces tend to overshoot their targets. The saccades that are executed tend to overshoot (or undershoot) by approximately a half of a character space for each character space that the intended target deviates from the preferred distance. This tendency is modulated by the duration of the launch site fixation, however, with longer fixations (on average) leading to greater saccade accuracy (McConkie et al. 1988; 1991). Both of these tendencies are instantiated in the model using Equation 5:

$$\text{SRE} = (\Psi - \text{intended saccade length})\,[\Omega_1 - \ln(\text{fixation duration})/\Omega_2] \qquad (5)$$

In Equation 5, Ψ is a free parameter representing the preferred saccade length: 7 character spaces. The discrepancy between this preferred distance and the length of the intended saccade is scaled by the right-hand term, which is a linear function of the natural logarithm of the launch site fixation duration. (The values of the free parameters Ω_1 and Ω_2 were fixed at 7.3 and 4, respectively.) Equation 5 thus ensures that the saccades that are executed will tend to overshoot (undershoot) their targets by approximately half of a character space for each character space that the intended saccade is less than (more than) seven character spaces. This systematic error is also modulated by the fixation duration on the launch site, so that there is less error following longer fixations.

The final term in Equation 4, *RE*, is the random error component. Consistent with McConkie et al.'s (1988; 1991) interpretations, this error term is normally distributed, with $\eta = 0$ and σ given by Equation 6. This equation stipulates that the size of the random error component increases proportional to the length of the intended saccade as determined by the values of the two free parameters, η_1 and η_2. (The values of these parameters were fixed at 1.2 and 0.15, respectively.)

$$\sigma = \eta_1 + \eta_2 \text{ intended saccade length} \qquad (6)$$

In closing this discussion of oculomotor control, we must revisit the issue of refixations. A key assumption of earlier versions of our model was that the oculomotor system begins programming a saccade to refixate a given word as soon as it is fixated. This saccade then ensues (resulting in a refixation) unless the first stage of lexical processing on that word finishes before the labile stage of programming, in which case the program is cancelled, and the oculomotor system begins programming a movement to the next word. This "horse race" between the initial stages of saccadic programming and lexical processing allowed the model to predict the correct proportion of refixations, but was problematic because it resulted in a non-monotonic relationship between the first-fixation durations and word frequency (i.e., the first-fixation durations on the low-frequency words were too short). This problem reflected an inherent limitation of the "horse race" assumption. That is, to predict the correct proportion of refixations, the model's parameter values had to set so that the labile programming of automatic refixations completed before the first stage of lexical processing. As a result, the saccades that moved the eyes off the initial landing site (i.e., the refixation saccades) occurred very rapidly, causing the first of several fixations to be too short. Thus, although longer words had a greater probability of being refixated, in the process, they also had a greater number of first fixations that were too short.

In E-Z Reader 7, we modified our assumption about automatic refixations: Rather than being started by default, upon fixating a given word, a program is instead initiated with a probability that is determined by the length of the word that is to be fixated. (The low-spatial frequency information that is used to determine word length is rapidly available from parafoveal vision; see Fig. 3.) Upon fixating a word, the oculomotor system initiates a labile program to refixate the word with probability, p, given by Equation 7. In Equation 7, $\lambda(= 0.07)$ is a free parameter that modulates how word length affects the probability of making a refixation. The model thus correctly predicts that long words are more often the recipients of multiple fixations than are short words. Similarly, the model also correctly predicts more refixations on low-frequency words than on high-frequency words. This is true because the first stage of lexical processing will complete less rapidly on low-frequency words, and as a result be less likely to cancel any labile refixation programs that happen to be pending. Finally, it should be noted that E-Z Reader 7 – like its predecessors – predicts that a substantial proportion of refixations occur because saccades overshoot and undershoot their intended targets.

$$p = \begin{cases} \text{length}\,\lambda & \text{if}(\text{length}\,\lambda) < 1 \\ 1 & \text{if}(\text{length}\,\lambda) \geq 1 \end{cases} \tag{7}$$

3.2. Simulation results

E-Z Reader 7's performance was evaluated using data from an eye-tracking experiment in which 30 college students read 48 sentences containing 8–14 words each (Schilling et al. 1998). We used the norms of Francis and Kučera (1982) to estimate what the token frequencies of the words were for our readers. (For example, the word "torpedo" is used very infrequently in written text, and as a result occurs only once in the corpus, whereas "the," the most frequently used word, occurs 69,974 times.) Before running the simulations, we completed a separate "cloze-task" experiment in which participants had to guess word$_{n+1}$ when given the sentence up through word$_n$ so as to determine each word's mean predictability within its sentence context. Finally, because regressions are outside of the scope of the model, we did not include data from sentences in which readers made inter-word regressions.

The first simulation examined the model's capacity to predict the means and distributions of several commonly used word-based measures of fixation duration and probability. To do this, we first divided the words into five frequency classes. For each of the frequency classes, we computed the means of the following measures: first-fixation duration, single-fixation duration, and gaze duration; and the probability of making a single fixation, the probability of at least one refixation, and the probability of skipping a word. We also constructed first-fixation and gaze duration distributions. Finally, we ran a simulation using 1,000 statistical subjects to determine how well the model could predict the observed means and distributions. The observed and predicted means are presented in Figure 6, and the observed and predicted distributions are presented in Figure 7.

As can be seen in Figure 6, the model does an excellent job predicting the observed means ($r^2 = 0.94$ for fixation durations; $r^2 = 0.98$ for fixation probabilities). In particular, E-Z Reader 7 – in contrast to its predecessors – correctly predicts the negative monotonic relationship between first-fixation durations and word frequency. This pattern was inherently problematic for earlier versions of the model because the relatively slow lexical processing of low-frequency words rarely finished before the "automatic" program to make a refixation, thereby causing the first of several fixations (and the mean first-fixation durations) on low-frequency words to be too short. E-Z Reader 7 avoids this problem by eliminating the assumption that refixations are automatically programmed upon fixating a word.[11]

Figure 7 shows that the model generated first-fixation and gaze duration distributions that are very similar to those that were observed. In fact, this aspect of the model's performance is considerably better than that of its predecessors. Although we have not quantified this improvement, it

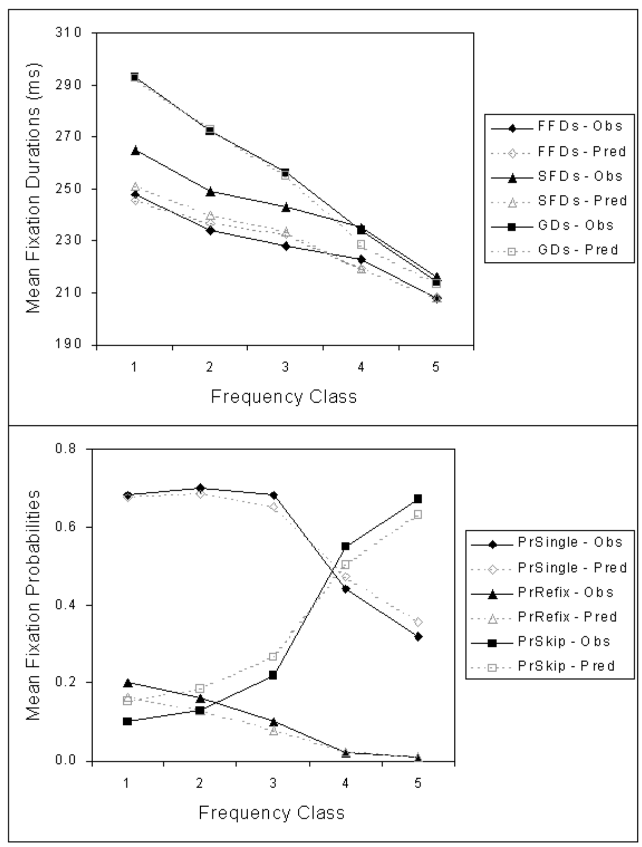

Figure 6. The top panel shows the mean observed and predicted first-fixation (FFD), single-fixation (SFD), and gaze durations (GD) for five frequency classes of words. The bottom panel shows the mean observed and predicted single-fixation (PrSingle), refixation (PrRefix), and skipping probabilities (PrSkip) for five frequency classes of words.

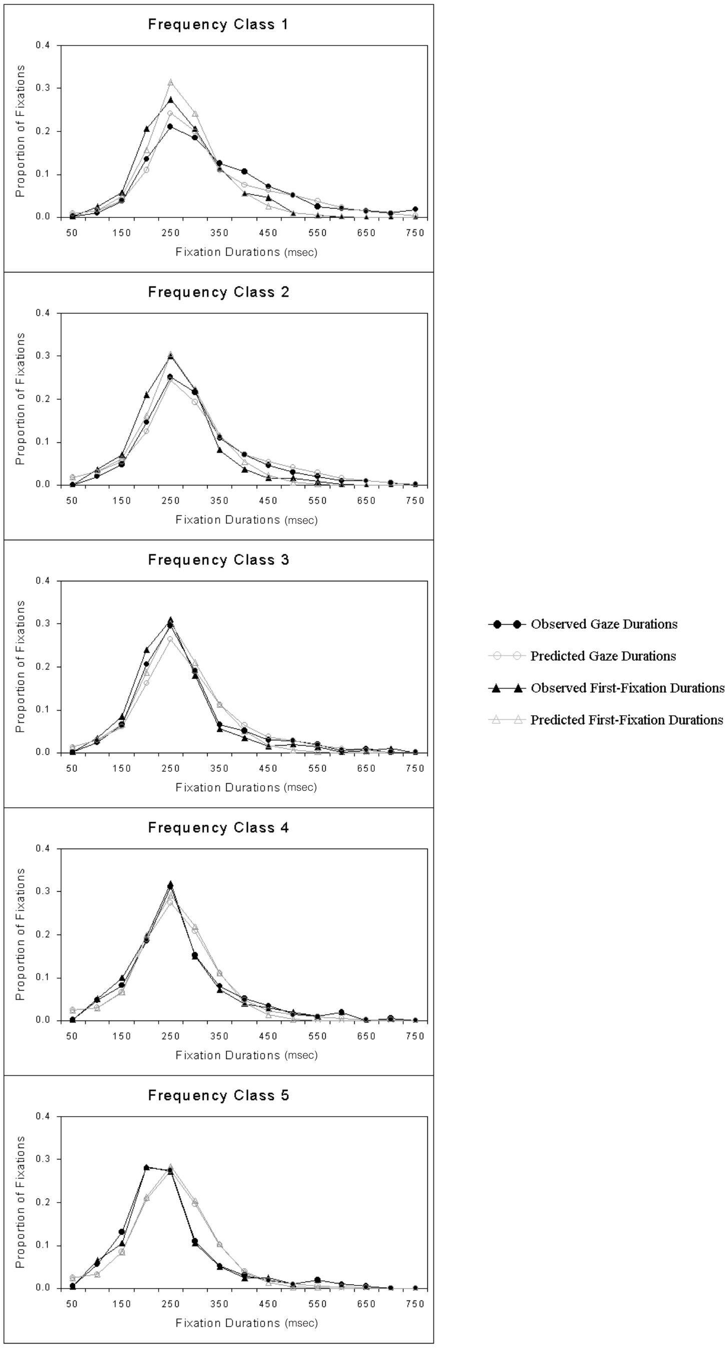

Figure 7. Observed and predicted frequency distributions of first-fixation (FFD) and gaze durations (GD). Each of the five panels shows the distributions for a separate frequency class of words. Each point represents the proportion of fixation durations within a given 50-msec interval (e.g., points above the abscissa labeled "100" represent the proportion of fixation durations between 50 and 100 msec that were observed in the sentence corpus and predicted by E-Z Reader 7).

is clear that the model is no longer over-predicting the amount of variability in the fixation durations (cf. Fig. 7 to Figs. 8 and 9 in Reichle et al. 1998).

Finally, we examined the first-fixation and gaze durations that were predicted for the low- and high-frequency target words that were used by Schilling et al. (1998) to study word-frequency effects during reading. In their experiment, Schilling et al. observed a mean gaze duration difference of 50 msec between the low- and high-frequency target words, as well as a 31-msec frequency effect on the first-fixation durations. E-Z Reader 7 predicted mean gaze and first-fixation duration frequency effects of 54 and 21 msec, respectively. The results of this simulation thus show that the model can handle both the aggregate properties of the Schilling et al. sentences and the frequency effects on specific words. Of course, previous versions of E-Z Reader could also account for a number of other "benchmark" phenomena; in the interest of evaluating the model further, therefore, we completed several additional simulations (each based on 1,000 statistical subjects).

In the first of these simulations, we first replaced the frequency values of all of the Schilling et al. (1998) target words with the mean frequency of the high-frequency targets (141 per million). We then repeated this procedure using the mean frequency of the low-frequency targets (2 per million). In both cases, the mean frequency values were inserted into the same within-sentence word positions as the original targets. The reason for inserting the mean frequency values into the sentence "frames" is that any resulting between-target differences can be attributed entirely to those items. As expected, the model predicted 84- and 44-msec frequency effects on the gaze and first-fixation durations, respectively. More importantly, the model also predicted 30- and 24-msec spillover frequency effects (for gaze and first-fixation durations, respectively) on the words immediately following the targets. These results are consistent with demonstrations that such spillover effects are typically one-third to one-half of the size of frequency effects (Rayner & Duffy 1986; Rayner et al. 1989; Schilling et al. 1998).

The second simulation examined the effects of parafoveal preview. To do this, we calculated the gaze durations on the Schilling et al. (1998) targets both with and without parafoveal processing of these words. The former condition was simulated using the standard (normal) model; to simulate the latter condition, we "lesioned" the model so that the first stage of lexical processing, L_1, on the targets could begin only after the words had been fixated. (Visual pre-processing was allowed.) Typically, the gaze durations on words increase 40–60 msec in the absence of parafoveal preview. Our simulation indicated that, with no parafoveal processing, the model predicted a 26-msec increase in the gaze durations on the target words. Although this prediction is a little smaller than what is typically observed, it is not entirely unreasonable, especially if one considers that the model predicts an additional increase in gaze durations (90 msec) in the complete absence of early visual processing.

The third simulation examined the processing "costs" that are incurred on word$_n$ that are due to: (1) skipping word$_{n-1}$; or (2) skipping word$_{n+1}$. Typically, the gaze duration on word$_n$ will be longer if word$_{n+1}$ is skipped than if word$_{n+1}$ is fixated (Pollatsek et al. 1993; Reichle et al. 1998). Likewise, there is some evidence that the gaze durations on word$_n$ are longer if word$_{n-1}$ is skipped than if word$_{n-1}$ is fixed. To examine these effects, we first calculated the mean

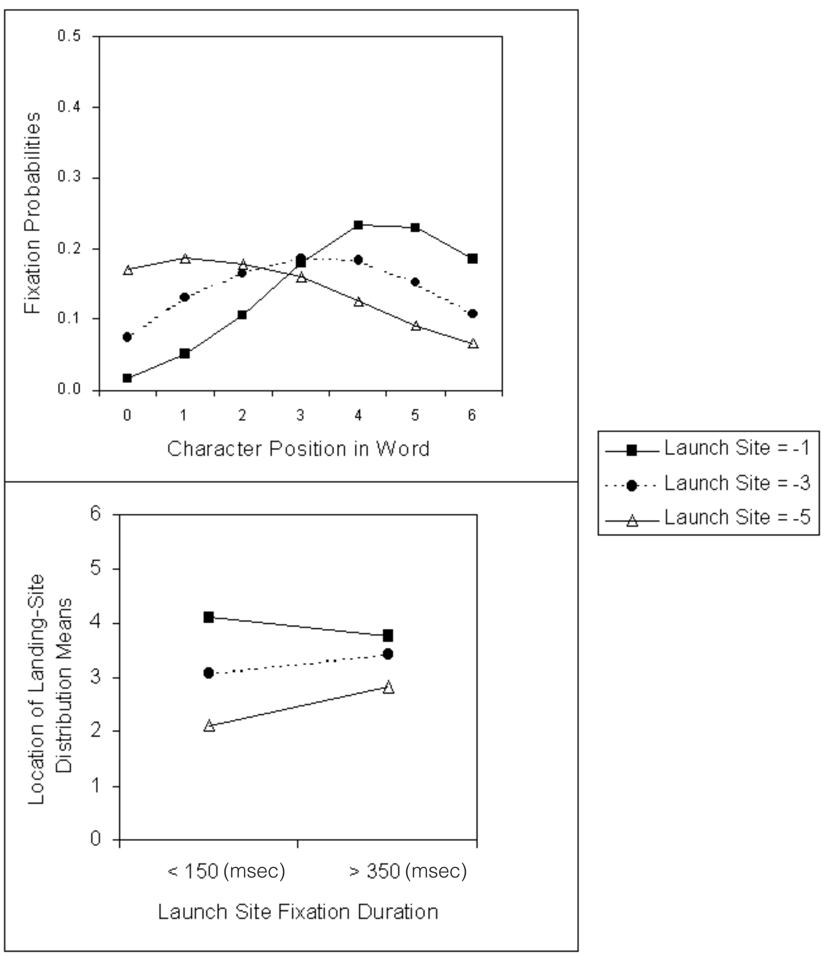

Figure 8. Simulation results showing the details of where the eyes move during reading. The top panel shows the landing site distributions on 6-letter words as a function of saccade length (i.e., the distance between the launch site and the middle of the saccade target). The locations of the launch sites and landing sites are indicated by numbers (in the legend and along the x-axis, respectively) representing ordinal position, from left to right, with the blank space between the two words being zero. The predicted landing sites are similar to those that have been reported elsewhere (e.g., McConkie et al. 1988; cf. Figs. 2 and 8A); that is, the distributions are approximately Gaussian in shape, with means that shift from near the word centers to near their beginnings with increasing saccade length. The bottom panel shows how the predicted systematic range error depicted in the top panel is modulated by the launch site fixation durations. As is evident, the systematic range error is attenuated following longer (above 350 msec) fixations on the launch site words.

gaze duration difference on the Schilling et al. (1998) target words when the following word was skipped versus fixated. The model should predict such an effect because, if $word_{n+1}$ is skipped, then the oculomotor system must modify the program to move the eyes to $word_{n+1}$ so that it instead moves the eyes to $word_{n+2}$, and such modifications take additional time. Indeed, the model predicted a 58-msec effect, which is similar is size to the 38-msec effect observed in the Schilling et al. (1998) corpus. Next, we calculated the mean gaze duration difference on the Schilling et al. targets when the immediately preceding word was skipped as opposed to fixated. Again, the model should handle this effect because, in cases where $word_{n-1}$ is skipped, any parafoveal processing that is done on $word_n$ will be completed from a more distant location than if $word_{n-1}$ is fixated. The model confirmed our predictions; it predicted a 66-msec effect, which again corresponds fairly closely to the 50-msec effect that was observed with the Schilling et al. materials. (E-Z Reader 7 handled these results significantly better than earlier versions of the model.)

The final simulation evaluated the model's capacity to account for the fine-grained details of where the eyes move during reading. This was done by examining the landing site distributions that were generated by E-Z Reader 7 on words of various lengths (again using the Schilling et al. [1998] sentences).[12] Figure 8A shows the landing site distributions that were predicted for 7-letter words. The figure indicates that the predicted landing site distributions closely resemble those reported by McConkie et al. (1988; 1991): (1) the landing sites are normally distributed; (2) the distribution means are located near the middle of the words; and (3) the distributions shift towards the beginnings of the words and become more variable as the distance between the launch sites and landing sites increases. Furthermore, as Figure 8B indicates, the magnitude of this systematic range error (i.e., how much the saccades over/undershoot their intended targets) is modulated by the launch-site fixation duration, so that there is less spread among the landing site distribution, means following longer launch site fixations. Together, the results of this final simulation are inconsistent with Reilly and O'Regan's (1998) claim that models like E-Z Reader cannot explain the patterns of landing site distributions that are normally observed during reading.

Before moving to the alternative models of eye-movement control, it is useful to note that Engbert and Kliegl (2001) sought to evaluate the basic assumption in E-Z Reader that lexical processing is the "engine" driving eye movements during reading. That is, they wanted to know if the time course of saccades is always determined by the time course of lexical processing. To answer this question, they implemented a computational model that, like E-Z Reader, accounts for eye-movement control during reading in terms of a few assumptions about lexical access and saccadic programming. There are two versions of the model, a two-state and a three-state version. The former is quite similar to a simpler version of E-Z Reader (Model 2 in Reichle et al. 1998), but there is only one stage of lexical processing, and it makes somewhat different assumptions about the variability of processes. The three-stage model is similar to the version of E-Z Reader that we are discussing except that, functionally, the first stage of lexical processing is replaced by an all-or-none process. That is, the reader is either assumed to wait until lexical access is completed before programming a saccade, or an "autonomous saccade" (i.e., completely independent of lexical processing) is executed. This all-or-none process (i.e., fully process the word before making a saccade or don't pay any attention to lexical processing) contrasts with E-Z Reader, in which the signal to make the saccade is partial lexical processing of the attended word.

Engbert and Kliegl's (2001) three-state model was first fitted to the same sentences (taken from Schilling et al. 1998) that were used to evaluate E-Z Reader. The model successfully predicted the mean fixation durations and skipping rates for the five frequency classes of words, and in so doing demonstrated that the state transitions can in fact be described using different distributional assumptions (i.e., residence-time dependent probabilities). Because these residence-time dependent probabilities can be implemented as an exact algorithm, whereas sampling from gamma distributions cannot, the model advances our understanding of eye-movement control by providing something like a process model of where the variability is coming from. The introduction of autonomous saccades in the three-state model marginally improved the ability of the model to fit frequency effects on both gaze durations and probability of word skipping. It also allows the model to predict (at least qualitatively) other phenomena that E-Z Reader can predict, such as spillover effects and word-frequency effects on preview benefit. However, it is by no means clear that this improvement can be taken as evidence for the existence of autonomous saccades during reading (as Engbert & Kliegl claim) because our model predicts the same phenomena by positing two stages of lexical processing.

4. Alternative models of eye-movement control

Models of eye-movement control during reading can be compared and contrasted along any number of different dimensions. Historically, the models have most often been classified as being either *oculomotor* or *cognitive/processing;* that is, with respect to whether or not language processing plays a prominent role in guiding the eyes during reading (Reilly & O'Regan 1998). Proponents of the oculomotor models claim that properties of the text (e.g., word length) and operating characteristics of the visual (e.g., acuity) and oculomotor systems (e.g., saccade accuracy) largely determine fixation locations. An auxiliary assumption of this view is that fixation durations are determined largely by where in a word the eyes have fixated. In contrast, proponents of the processing models tend to emphasize the role of language processing in guiding eye movements during reading. According to this view, the decision about how long to fixate is determined by ongoing linguistic processing, whereas the decision about where to fixate is jointly decided by linguistic, visual, and oculomotor factors. Although these two views of eye-movement control in reading have often been treated as completely distinct theoretical "camps," the distinction is one of degree because the actual models vary considerably with respect to how central a role linguistic processing plays in determining the moment-to-moment movements of the eyes through the text.

This fact has been acknowledged in more recent papers. Engbert et al. (2002), for example, have also categorized the existing oculomotor models with respect to their assumptions regarding attention. According to this taxonomy, the models near the cognitive end of the continuum can be further divided into those that assume the serial allocation of attention (i.e., *sequential attention shift*), and those that posit an attention gradient (i.e., *guidance by attentional gradient*). In the sequential-attention-shift models, attention is allocated serially, from one word to the next, whereas in the guidance-by-attentional-gradient models, attention is a gradient, so that more that one word can be attended to (and processed) in parallel. Because the question of how attention is allocated during reading is quite contentious (see Henderson & Ferreira 1990; Inhoff et al. 2000a; 2000b; Kennedy 2000; Murray 1998; Rayner et al. 1998), the models will undoubtedly play a prominent role in guiding future research in an effort to resolve this issue. (How this issue is resolved will also have important ramifications for the models.) Consequently, in the following review, we shall use both of these dimensions in describing existing models of eye-movement control during reading. We shall also use the oculomotor-cognitive dimension to organize our discussion, starting with those models that assign the least significance to linguistic processing.

4.1. Minimal control

In this model, neither fixation durations nor saccade lengths are affected by linguistic or cognitive factors, but are instead affected only by the physical layout of the text (Suppes 1990; 1994). The model consists of a small number of axioms that describe the fixation-duration distributions and a random-walk process that determines where the eyes will move next.

The axioms describing fixation durations are as follows: First, the duration of each fixation is a function of the number of operations (which are never specified) that must be completed during each fixation. Second, the fixation durations are stochastically determined by sampling from an exponential distribution if a single operation must be completed; in cases requiring two operations, the durations are described by the convolution of two independent exponential distributions. Finally, the fixation times are independent of both earlier processing and the current text content. Thus, the model stipulates that variability in fixation durations is not due to variability in the duration of the underlying cognitive processing, but instead reflects the probabilistic nature of the processing.

Saccades are determined by a similar set of rules. First, if the processing within a "region of regard" (which is defined – in the case of reading – by a given word) completes, then the eyes are moved to the next word; otherwise, they remain in the same location. Second, if processing has not finished and the memory for a prior region of regard has decayed, then the eyes are moved back to that prior region. Third, if perceptual processing of the upcoming word has finished from the current location, then the upcoming word is skipped. Finally, the length of each saccade is independent of both earlier processing and the length of prior saccades. (Thus, cognitive processing is posited to affect the locations of fixations.)

Unfortunately, the minimal-control model has only been used to simulate eye movements during an arithmetic task (Suppes 1990; Suppes et al. 1982; 1983), so that it is difficult to evaluate its adequacy with respect to reading. It is clear, however, that the model only makes predictions on the level of individual words, and therefore cannot account for either landing site distributions (McConkie et al. 1988) or the optimal viewing position effects (O'Regan 1990). The model also fails to account for many other factors that are acknowledged by Suppes (1994) to affect eye movements during reading.

4.2. Strategy tactics

This model originated from two observations: First, words are identified most rapidly if they are fixated slightly to the left of center, on the *optimal viewing position;* and second, words are also less likely to be refixated if they are initially viewed from this position (O'Regan 1990; 1992b; O'Regan & Lévy-Schoen 1987). These results led O'Regan to suggest that readers adopt a "strategy" of directing their eyes from word to word in an attempt to fixate each word's optimal viewing position. This reading strategy is "risky" because the saccades often miss their intended targets, so that the words are sometimes viewed from sub-optimal locations. To compensate for this, the reader can also use a "careful" variant of the strategy that includes the following within-word "tactic": If the eyes do not land near the optimal viewing position, then immediately move them to the other end of the word. Using this tactic ensures that every word will either be viewed from its optimal position (in the case of single fixations) or will be viewed from two different locations (in the case of refixations).

Because the within-word tactics are guided by visual factors (e.g., word length), the model predicts that linguistic variables (e.g., word frequency): (1) should only modulate fixation durations when there is a single long fixation or when the fixation is the second of two, and (2) should not modulate refixation probabilities. Unfortunately for the strategy-tactics model, neither of these predictions has been confirmed. Rayner et al. (1996) found that word frequency effects were evident in the first of two fixations (see also Sereno 1992), and that refixations were more likely on low-frequency words than on high-frequency words (with length controlled). In addition, Rayner et al. found that neither fixation durations nor frequency effects on single-fixations varied as a function of landing position,[13] which suggests that the optimal viewing position may be much less important in normal reading than in the identification of single words when they are presented in isolation (see also Vitu et al. 1990). It is worth noting that our current conjecture about refixations (see Equation 8) is similar to that of the "careful" strategy; both assume that the reason for moving the eyes to a second location within a given word is that it affords the reader a better view from which to identify the word.

4.3. Word targeting

This theory was largely motivated by the seminal work of McConkie and his colleagues (McConkie et al. 1988; 1989; 1991; Radach & McConkie 1998). As mentioned previously, they expanded upon the observation that readers typically fixate the preferred viewing location (Rayner 1979), and found that landing site distributions behaved systematically with respect to both saccade length and the launch site fixation duration. These findings led McConkie and his colleagues to conclude the following: First, the landing site distributions (which resembled truncated Gaussian distributions; see Fig. 2) reflect random noise in the oculomotor system, with the missing tails being due to cases in which the eyes undershot or overshot their intended targets. The oculomotor system is also assumed to be "tuned" to make saccades approximately seven character spaces in length, so that longer saccades tend to undershoot their targets, while shorter saccades tend to overshoot their targets. This systematic range error causes the distributions to shift towards the beginnings of words as the launch site becomes more distant from the intended saccade target. With longer launch site fixations, however, the eye-movement system has more time to plan its saccades, which results in more accurate saccades and a reduction in the systematic range error.

The relationships among saccade length, the duration of the launch site fixation, and saccadic accuracy led to the development of precise mathematical descriptions of how these variables affect the landing site distributions during reading (McConkie et al. 1994). Although there have also been attempts to provide similar mathematical descriptions of fixation durations (McConkie et al. 1994; McConkie & Dyre 2000; see also Brysbaert & Vitu 1998), these accounts are little more than precise descriptions of the data, and do not attempt to explain how linguistic processing affects fixation durations during reading. Also, because these descriptions address the "where?" and "when?" questions of eye-movement control independently, they fail to explain why the durations of fixations are related to their spatial locations.

Recently, however, several word-targeting strategies were implemented as computer simulations (Reilly & O'Regan 1998) so that several theoretical assumptions about eye-movement control could be evaluated with respect to how well they handle the findings related to landing-site distributions (McConkie et al. 1988). These simulations included several alternative strategies, including three that might be classified as oculomotor (e.g., *word-by-word, target long words,* and *skip short words*) and at least one in which language processing is important (e.g., *skip high-frequency words*). The results of these simulations indicated that the target-long-words strategy fit the landing-site distributions better than the other strategies, while the language-based strategies fared rather poorly overall. On this basis, Reilly and O'Regan suggested that language-processing models do not provide an adequate account of eye-movement control during reading. As we demonstrated earlier, however, processing models (e.g., E-Z Reader) can generate reasonable-looking landing-site distributions (see Reichle et al. 1999). Our model's successes here are largely due to the fact that McConkie et al. (1988; 1989; 1991) provided such a precise explanation of how visual and oculomotor variables affect eye movements, and that incorporating such an eye-guidance mechanism into our model is fully compatible with our model's other language processing assumptions.

4.4. Push-Pull

Yang and McConkie (2001) recently applied the core assumptions of the Push-Pull theory of saccade generation (Findlay & Walker 1999) to the domain of reading. The name of this model originates from the hypothesis that the timing of saccades is determined by the outcome of competitive ("push-pull") operations that occur among various components of the oculomotor system. These operations are necessary to resolve the ever-present conflict of whether to keep the eyes stationary (i.e., to fixate) or move the eyes to a new location (i.e., to make a saccade). Thus, the key assumption of this model is that the timing of saccades is largely independent of lexical processing (with the exception that processing difficulty can inhibit the oculomotor system from initiating a program). At present, however, the model has not been implemented within a computational framework, so it is difficult to evaluate how well it accounts for the various reading phenomena that have been described in this paper.

4.5. SWIFT

Many of the ideas of the Push-Pull model have been instantiated in the SWIFT (Saccade-generation With Inhibition by Foveal Targets) model (Engbert et al. 2002; Kliegl & Engbert 2003). The model's architecture is shown in Figure 9. If one compares Figure 9 to Figure 3, it is evident that SWIFT and E-Z Reader share several key assumptions: In both models, words are identified in two stages and saccadic programming is completed in two stages. In contrast to E-Z Reader, however, SWIFT assumes that lexical processing is distributed over a four-word attentional gradient (i.e., SWIFT is a guidance-by-attentional-gradient model). Another important difference between the two models is that saccadic programs in SWIFT are initiated autonomously, after a variable (random) time interval, unless this interval is extended because the word being fixated is difficult to process. In contrast to E-Z Reader, therefore, lexical processing in SWIFT is not the engine driving eye movements during reading; instead, saccades are initiated so as to maintain a preferred mean rate of eye movements.

During the first stage of lexical processing, the lexical activity of word$_n$ at time t, $a_n(t)$, increases (i.e., $da_n/dt > 0$) until it reaches some maximum value, L_n. During the second stage of lexical processing, $a_n(t)$ decreases (i.e., $da_n/dt < 0$) until it equals zero. L_n is a function of the word's normative frequency of occurrence in text and its predictability in the local sentence context, as given by:

$$L_n = (1 - \text{predictability}_n)\,[\alpha - \beta \log(\text{frequency}_n)] \tag{8}$$

In Equation 8, α and β are free parameters that modulate the effect of word frequency. (Note the similarity between Equation 8 and the equations that determine lexical processing times in E-Z Reader: Equations 2 and 3.) The lexical activity of word$_n$ reaches its maximum at time t_p. The rate at which a_n approaches L_n is given by:

$$\frac{d\,a_n(t)}{d\,t} = \begin{cases} +f\lambda_k t & \text{if } t < t_p \\ -\lambda_k t & \text{if } t \geq t_p \end{cases} \tag{9}$$

In Equation 9, f and λ_k are parameters that control the rate at which a_n approaches L_n. The parameter f increases the rate of the first stage of lexical processing (relative to the second) so that it is completes more rapidly, and the λ_k parameter adjusts the rate of lexical processing as a function of the distance between the word and the fovea (i.e., the point of fixation). The parameter λ_k has four values (as indexed by the k subscript): One for each of the four words in the attentional gradient. Thus, the word being fixated (word$_n$) is processed most rapidly, word$_{n-1}$ and word$_{n+1}$ are processed less rapidly, and word$_{n+2}$ is processed least rapidly (i.e., $\lambda_n > \lambda_{n-1} = \lambda_{n+1} > \lambda_{n+2}$). This asymmetry in the attentional gradient reflects the well-known fact that, for readers of English, the perceptual span extends further to the right of fixation (McConkie & Rayner 1975; Rayner 1986; Rayner & Bertera 1979; Rayner et al. 1982).[14]

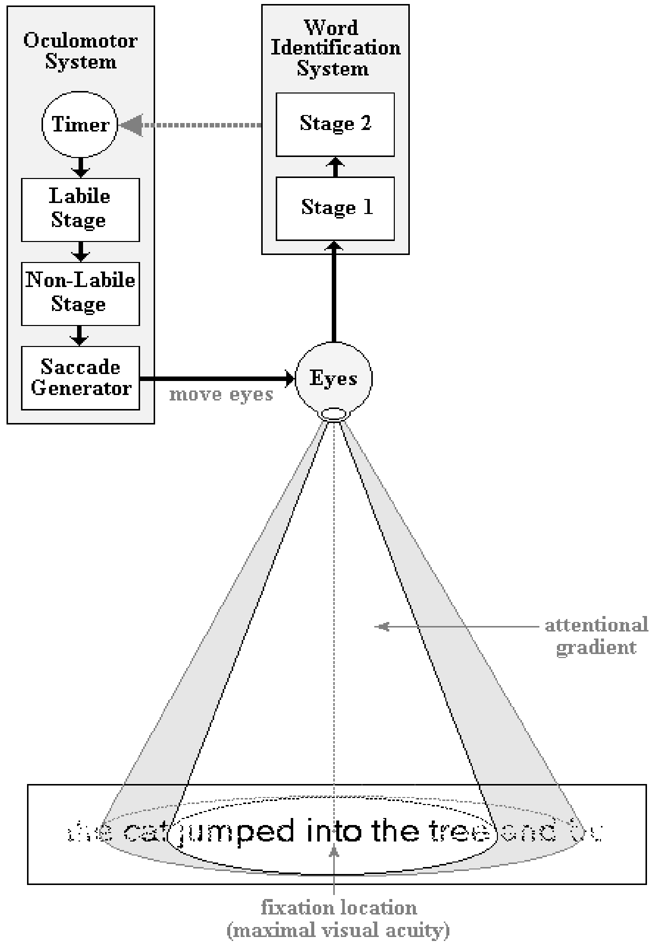

Figure 9. A schematic diagram of the *SWIFT* model (Engbert et al. 2002; Kliegl & Engbert 2002). Lexical processing occurs within a four-word attentional gradient. Saccadic programs are initiated autonomously, by a timing mechanism, so as to maintain a mean rate of eye movements. The dashed gray arrow represents the inhibitory link between the fovea and the oculomotor system. This inhibitory link allows word identification to extend the duration of the current fixation (via increasing the duration of the time interval between saccades) if the word being fixated is difficult to process.

In SWIFT, eye movements are directed towards words that have received intermediate amounts of lexical processing. The conditional probability of a saccade being directed towards word_k at time t if the eyes are currently on word_n is given by Equation 10. In this equation, the subscript m indexes word position within the attentional gradient, which extends two words to the right of the currently fixated word (i.e., word_n). If $\Sigma_m a_m(t) = 0$, then the eyes are directed towards the next word immediately to the right of the attentional gradient that has not been completely processed.

$$\Pr(k, t\_n) = \begin{cases} a_k(t) / \Sigma_m a_m(t) & \text{if } k \le n+2 \\ 0 & \text{if } k > n+2 \end{cases} \tag{10}$$

As already mentioned, saccadic programs are initiated so as to maintain a mean rate of eye movements. Saccadic programs are initiated after a random interval, t, that is given by Equation 11. In Equation 11, t_s is a random time interval (the value of which is determined by sampling from a gamma distribution) and h is a free parameter that lengthens t_s by an amount proportional to the lexical activity of word_n. The intuition behind Equation 11 is that the model's tendency to relentlessly drive the eyes forward will be held in check if the word identification system is experiencing difficulty processing the word that is currently being fixated. Two points about Equation 11 are noteworthy: First, this inhibition by foveal targets is necessary for the model to account for the frequency effects that are typically observed on first-fixation durations. (The model presumably predicts frequency effects on the other word-based measures because, in natural text, word frequency is negatively correlated with word length, so that longer words tend to be fixated more often – purely by chance – than shorter words.) Second, although this inhibition is necessary to produce normal word frequency effects, it is operational only approximately 15% of the time.

$$t = t_s + h\, a_n \tag{11}$$

Finally, the initiation of saccadic programs in SWIFT is separated from the selection of saccade targets. Thus, the target of an upcoming saccade is not selected as soon as the program is initiated; instead, there is a lag, so that there is little "cost" in terms of re-programming time if the labile program has to be cancelled. This assumption provides a means of avoiding the problem associated with earlier versions of our model (e.g., E-Z Readers 5 and 6); namely, that our model predicted costs due to skipping that were too large.

SWIFT was applied to the same corpus used to evaluate

E-Z Reader (i.e., the Schilling et al. 1998 sentences). Like our model, SWIFT successfully predicted the mean values for each of the word-based measures. (Engbert et al. [2002] have not, however, examined the predicted distributions.) Although Engbert et al. did not examine their model's performance on the Schilling et al. high- and low-frequency target words, the model would undoubtedly handle the frequency effects on these specific items, too. Furthermore, in contrast to earlier versions of our model (E-Z Readers 5 and 6) but not to the current version, SWIFT predicts costs for skipping upcoming words that are concordant in size with those that have been reported in the literature. As Engbert et al. indicate, this aspect of the model's performance stems from the fact that the timing of the saccadic programs is decoupled from their target selection. This distinction between the two models has been blurred, however, because of our assumption in the current version of the E-Z Reader model that target selection occurs during the later half of the labile saccadic programming stage.

Kliegl and Engbert (2002) have recently examined SWIFT's capacity to simulate the results of a gaze-contingent display experiment reported by Binder et al. (1999) in which parafoveal preview of specific target words was either allowed or denied. The model successfully captured the pattern of effects observed in this experiment: In the absence of parafoveal preview, the target words tended to be fixated longer, skipped less often, and be the recipients of more regressions.

In the final analysis, we agree with Engbert et al. that SWIFT provides a viable alternative – at least as measured with respect to the model's capacity to handle a wide array of phenomena – to the current sequential-attention-shift models, including E-Z Reader. Although the model has not yet been fitted to the landing site distribution data reported by McConkie and his colleagues (McConkie et al. 1988; 1991), we acknowledge that the model could probably account for these effects if it were augmented with assumptions similar to those used by E-Z Reader (i.e., Equations 4 to 6). Nevertheless, we strongly believe that the remaining differences between the two models are far from being merely cosmetic. To reiterate, in E-Z Reader, attention is allocated serially, from one word to the next, with word identification being the "engine" driving the eyes forward. In stark contrast to this, in SWIFT, attention is allocated in parallel, to several (four) words within an attentional "window," with the tempo of the eye movements being largely independent of the moment-to-moment lexical processing (with the only exception being due to the occasional delays in the initiation of saccadic programs due to foveal inhibition by difficult words). We suspect that, in the future, the relative merits of the two sets of assumptions will be measured with respect to how well they handle the many effects of linguistic variables that have been documented in the reading literature (see Rayner 1998). For reasons that we have discussed elsewhere (Pollatsek & Rayner 1999), we believe that the ability to explain such effects will ultimately support our claim that the intrinsic nature of language processing during reading hinges upon word identification: (1) proceeding in a serial fashion, and (2) being the primary determinant of when the eyes move.

4.6. Glenmore

Yet another model inspired by Findlay and Walker (1999) is the Glenmore model[15] of Reilly and Radach (2003). The model's architecture is depicted in Figure 10. As is evident in the figure, Glenmore is a connectionist model (cf. McClelland & Rumelhart 1986; Rumelhart et al. 1986b) that consists of three major components: (1) a saliency map that selects the saccade targets; (2) an interactive-activation network that identifies words; and (3) a saccade generator that initiates and executes eye movements.

Like both the Push-Pull model (Yang & McConkie 2001) and SWIFT model (Engbert et al. 2002), lexical processing is distributed across a gradient. Letter presence/absence is encoded across a series of 30 letter-sized input units, each of which corresponds to a unique spatial location in the visual array. The activation of these units is scaled so that it decreases for units that are farther away from unit 11 (which, in the model, is the center of the fovea). The scaling is done using a gamma distribution function with a mean centered on unit 11, as described by Equation 12. In this equation, i is the position of the input unit, and μ_G and σ_G are parameters that specify the mean and standard deviation of the distribution, respectively. The scaled input unit activation is then propagated (via direct one-to-one connections) to both the letter units of the word-identification system and units of the saccade target saliency map.

$$\text{activation(i)} = \text{Gamma}(\text{i}, \mu_G, \sigma_G) \tag{12}$$

Each letter unit receives activation from (and sends activation to) the word units, so that a given letter sequence can be mapped onto its corresponding lexical representation. The model thus incorporates many of the basic processing principles of the classic *Interactive-Activation Model* of word-identification (McClelland & Rumelhart 1981), such as top-down modulation of letter activation and a "winner-take-all" competition among word units. Letter units also send activation to the saliency units, which also receive activation from the input units. The saliency units form a map, with each unit corresponding to one of the 30 locations specified by the input units. This saliency map is used to select the targets of upcoming saccades; the unit that is most active will be the target of any saccade that is executed.

Activation is propagated to the letter and saliency units in standard fashion; the input to each unit i at time t is given by Equation 13, in which $o_{j,t}$ is the activation that is being propagated to unit i from unit j, and $w_{i,j}$ is the connection weight between unit i and unit j.

$$\text{input}_{i,t} = \text{input}_{i,t-1} + \Sigma_j \, w_{i,j} \, o_{j,t} \tag{13}$$

The accumulation of activation within these two types of units is described by a Gaussian probability density transfer function; that is, the units accumulate activation over time as described by Equation 14. Here, input_i is the net input to unit i (as given by Equation 13), and μ_N and σ_N are parameters that specify the mean and standard deviation of the distribution, respectively.

$$\text{activation}(\text{input}_i, \mu_N, \sigma_N) = \frac{1}{\sqrt{(2\pi\sigma_N)}} e^{-(\text{input}_i - \mu_N)^2/(2\sigma_N^2)} \tag{14}$$

The activation described by Equation 14 is then propagated to the word units using Equation 15. In Equation 15, $L_{j,t}$ is activation from letter unit j (which is divided by word length, n, to nullify the effect of this variable), $W^R{}_{i,t}$ is the activation from a word unit i to itself (via recurrent connections), and $W^O{}_{k,t}$ is activation from other word units (via inhibitory connections).

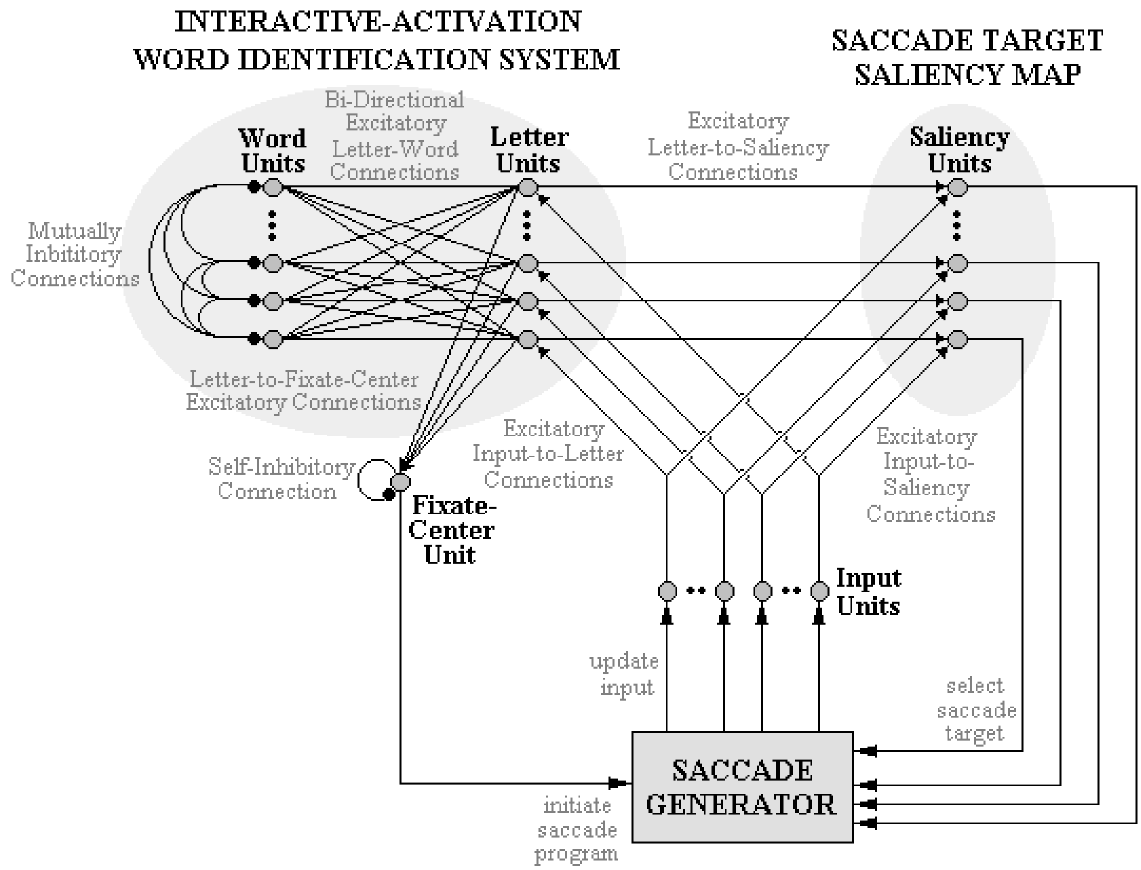

Figure 10. A schematic diagram of the *Glenmore* model (Reilly & Radach 2003). The model has a connectionist architecture and is comprised of three main components: (1) an interactive-activation network that is responsible for identifying words; (2) a saliency map that selects saccade targets; and (3) the saccade generator. Activation of the input units is propagated forward to the letter and saliency units so as to identity and localize the individual letters in the 30-unit input array. Letter activation is then spread to the word units (which provide top-down modulation of the letter units), the saliency units, and a fixate-center unit. A saccade is initiated to the target location that corresponds to the most active saliency unit whenever the activation of the fixate-center unit falls below a certain threshold.

$$\text{input}_{i,t} = \text{input}_{i,t-1} + \frac{\Sigma_j w_{i,j} L_{j,t}}{n} + \Sigma_i w_{i,i} W^R_{i,t} - \Sigma_k w_{i,k} W^O_{k,t} \quad (15)$$

Word unit activation is accumulated using a sigmoid transfer function, so that the activation of unit *i* is given by Equation 16. Activation therefore ranges continuously over the range 0 to 1 and is equal to 0.5 when the net input (given by Equation 15) equals the free parameter α. The other free parameter, β, controls the steepness of the function, or the rate at which activation goes from zero to one as the net input increases. The role of the word units is to support the letters of words that are presented as visual input. This is critical because the letter units also propagate activation to the fixate-center unit, which is responsible for initiating saccades. When the activation of the fixate-center unit falls below a certain threshold, it signals the saccadic generator to move the eyes to the location specified by the saliency map.

$$\text{activation(i)} = 1 / \{1 + e^{-[(\text{input i} - \alpha)/\beta]}\} \quad (16)$$

The saccades that are generated by Glenmore are subject to both systematic and random error. The landing site distribution mean, μ, is centered (i.e., is equal to zero) on the target word and deviates from the target as described by Equation 17. Likewise, the standard deviation of the landing site, σ, also varies as a function of saccade length, as described by Equation 18. In these equations, the slope (b_1 and b_2) and intercept (m_1 and m_2) parameters modulate the effect of saccade length.

$$\mu = b_1 + m_1 \text{ (saccade length)} \quad (17)$$

$$\sigma = b_2 + m_2 \text{ (saccade length}^3) \quad (18)$$

Finally, each landing site, *x*, is a random deviate that is independently sampled from a Gaussian distribution defined by Equation 19, with μ and σ being defined by Equations 17 and 18, respectively.

$$f(x,\mu,\sigma) = \frac{1}{\sqrt{(2\pi\sigma)}} e^{-(x-\mu)^2/(2\sigma^2)} \quad (19)$$

The Glenmore model has been successfully applied to wide range of eye-movement phenomena. However, instead of fitting their model to a sentence corpus (as we and others have done with the Schilling et al. 1998 sentences), Reilly and Radach (2003) have demonstrated their model's competence by running simulations in which they illustrate key properties of its performance. So far, they have shown that Glenmore successfully predicts many of the findings simulated by our model, including word-frequency effects, spillover effects, and preview effects that are modulated by the difficulty of the fixated word. Moreover, although they did not provide evidence that the model reproduces the types of landing site distributions observed by McConkie et al. (1988; 1991), the model has clearly been designed to account for such effects (see Equations 17–19). Likewise, it remains an open question as to whether the model can predict the costs that have been observed for skipping. Based on these results, therefore, we think that the Glenmore model is very promising, and that – again, if one only considers the model's performance – it provides a viable alternative to the cognitive-based, serial attention models (like E-Z Reader). However, we also believe that the model may be inherently limited in that it makes no provisions for explaining how linguistic variables affect eye movements during reading. As it is currently implemented, for example, the Glenmore model cannot handle predictability effects. We suspect that, given the model's core asumptions (e.g., the gradient of lexical processing), many of the well-documented effects of linguistic processing (see Rayner 1998) may prove to be even more challenging for the model.

4.7. Mr. Chips

This model was proposed as a means to evaluate how an *ideal-observer* (i.e., a reader with perfect lexical knowledge and the well-specified goal of maximizing reading speed) would move his/her eyes (Klitz et al. 2000; Legge et al. 1997). Consequently, the model exemplifies a very different approach to understanding the interrelationships among visual processing, word recognition, and eye-movement control during reading. The model does this using three pieces of information: (1) input from a "retina" that encodes a small number of letters in the fovea and indicates whether letters in the parafovea/periphery are present or absent; (2) knowledge about the relative frequencies with which words occur in text; and (3) knowledge of the likelihood of making a saccadic error of a given size for each given saccade length. These three types of knowledge are depicted in Figure 11.

The Mr. Chips model attempts to use all of the above information that is available from a particular fixation location to identify the next word in text using the fewest saccades possible. To do this, the model calculates the expected uncertainty that is associated with being able to identify a word for saccades of each possible length. It then executes a saccade that minimizes this uncertainty. For example, imagine that the model has the following information about a word: It is five letters long and begins with "abo" (see Fig. 11). The model uses this information in conjunction with its lexical knowledge to calculate conditional probabilities of the letter string being each of the words that satisfy these constraints, using Equation 20:

$$p_i = P_i / \Sigma_j P_j \tag{20}$$

In Equation 20, p_i is the conditional probability of the letter string being word$_i$, given the letter information already known ("abo" in the example); P_i is the absolute probability of the letter string being word$_i$; and the P_js are the absolute probabilities of the letter strings in the "candidate set" (in the example, all of the 5-letter word beginning with "abo"). In the Figure 11 example, the conditional probability that "abo—" is "about" is equal to 0.849.

The conditional probabilities are then used to compute the conditional entropy, or degree of uncertainty, H, that would result from a saccade of length, L, under the assumption that the letter string is word$_i$ using Equation 21. For example, from the current fixation, the entropy associated with the letter string is: H(0, abo—) = 0.613. (Smaller entropy values represent less uncertainty about the identity of a word, so that identification occurs with certainty when the entropy value associated with a letter string equals zero.) A saccade of $L = 1$ would reveal one letter, which, given the model's lexical knowledge, must be either "u" or "v." If the letter is "u," then the conditional probability of the word being "about" is $p = 1$, and the conditional en-

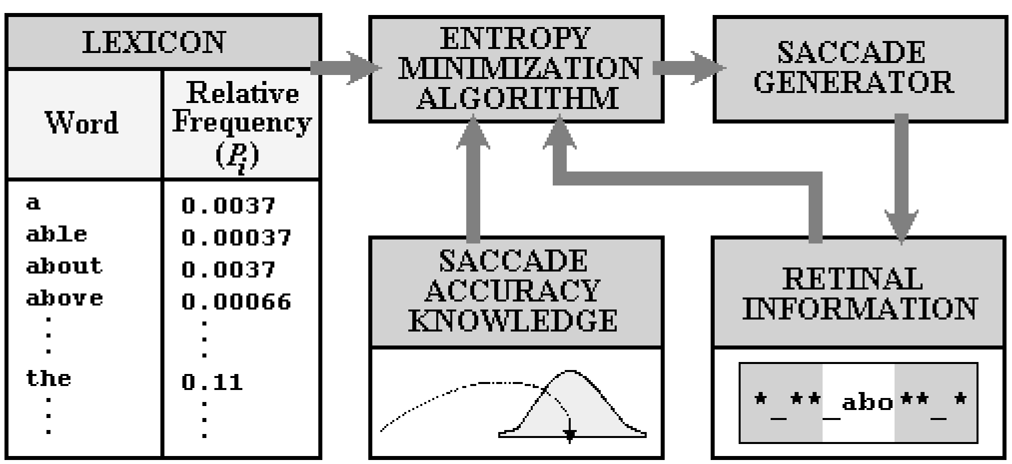

Figure 11. A schematic diagram of the *Mr. Chips* model (Klitz et al. 2000; Legge et al. 1997). The model attempts to compute the saccade length that will minimize the uncertainty about the identity of next unidentified word. It does this using three sources of information: (1) the relative frequencies with which the words in its lexicon occur in text; (2) the accuracy of saccades for each possible saccade length; and (3) visual information from the model's "retina." Visual information is encoded from two regions in the retina: a fovea, in which letters can be identified, and a parafovea, in which letters can be discriminated from blank spaces. (In the Figure, the retina is presented by a rectangle, with the white and gray areas corresponding to the fovea and parafovea, respectively.) The entropy-minimization algorithm computes the saccade length that will minimize the uncertainty of the next unidentified word, and then an error-prone "Saccade Generator" executes the saccade so that the retina can encode additional letter information.

tropy would be reduced to: $H(1, \text{about}) = 0$. Likewise, if the letter is "v," then the conditional entropy is reduced to: $H(1, \text{above}) = 0$.

$$H(L, \text{word}_i) = -\Sigma_i \, p_i \log_2 (p_i) \qquad (21)$$

After the conditional entropies are calculated for each possible saccade length, Mr. Chips computes a probability-weighted average to determine the expected entropy associated with a saccade of each given length. This is done using Equation 22. In the example, $H(L) = 0$ for saccades of lengths 1 to 5. Because of saccadic error, however, each saccade of intended length, L, has an associated landing-site distribution, $P_L(x)$, which determines the probability of making a saccade of actual length, x. The model uses this knowledge to calculate the entropy associated with each saccade length, L, averaged across all of the possible landing sites. Equation 23 gives the expected uncertainty, H_L, associated with making a saccade of intended length L. Finally, the model makes the saccade that minimizes H_L, and thereby maximizes the probability of identifying the word. In cases where more than one possible saccade yields the same expected entropy, Mr. Chips executes the longest saccade possible so as to maximize reading speed.

$$H(L) = \Sigma_i \, p_i \, H(L, \text{word}_i) \qquad (22)$$

$$H_L = \Sigma_i \, P_L \, (x) \, H(x) \qquad (23)$$

Because Mr. Chips was developed with the intent of examining the way lexical knowledge and restrictions on visual encoding affect saccade lengths and fixation locations, the model does not address the "when?" question of eye-movement control. Several of the model's emergent properties, however, are consistent with research findings about where the eyes move. For example, the model predicts that the mean saccade length will be around seven character spaces (McConkie et al. 1988) and that saccades will tend to be directed towards the optimal viewing position (O'Regan 1990). The model also predicts parafoveal preview effects because the left-most letters of upcoming words are often identified before the words are actually fixated.

Unfortunately, it does not seem plausible that human readers compute the expected amount of information to be gained from each possible saccade length so as to make the saccade that maximizes this gain. Klitz et al. (2000) acknowledge this fact, and say that their model "is not intended as an exact model of how humans perform a task, but rather establishes an upper bound (i.e., a level of competence) for human performance." Furthermore, the Mr. Chips algorithm is well approximated by the simple heuristic of left-justifying the target word in the high-resolution part of vision, so that, on some level, the model is psychologically plausible.

Moreover, it is important to point out that Mr. Chips, unlike the other models discussed in this article, was developed to investigate how visual impairment might affect eye movements during reading. In this capacity, the model has been successful (Klitz et al. 2000). A comparison of the model's performance to that of a human in a reading task[16] with a simulated *scotoma* (i.e., a blind spot in the visual field) indicated that, in contrast to the model, the human had difficulty integrating information across central scotomas more than a single character-space in size. The human reader appeared to primarily use visual information from one side of the scotoma and to use the visual strategy of moving the eyes in order to place the region of normal vision over all of the character spaces in turn, rather than using lexical knowledge to winnow down the possible identities of letter strings from a single fixation. Although the human reader's natural strategy produced shorter saccades, it markedly increased reading speed over when they tried to execute the Mr. Chips strategy. These analyses, therefore, suggest that, while the seemingly erratic eye movements of readers with scotomas do not allow the maximal amount of information to be extracted from the page during each fixation, they are nevertheless adaptive in that they allow a maximal overall rate of information extraction.

4.8. Attention shift

In the attention-shift model (or ASM), linguistic processing and eye-movement control are loosely coupled (Reilly 1993). As Figure 12 indicates, the model's architecture consists of pair of interacting connectionist networks that are trained using the back-propagation learning algorithm (Rumelhart et al. 1986a). One of these networks is responsible for word identification; the other is responsible for programming saccades. As each word is identified, the lexical-encoding network signals attention to shift to the next word, so that it can be processed. The movement of attention, in turn, causes the saccadic-programming network to begin programming a saccade to the next word. In contrast to E-Z Reader, the ASM does not allocate attention serially, from one word to the next. The attention "spotlight" is instead fixed in size, so that whatever falls within the spotlight will be the focus of attention. This means that, in cases where two or more short words follow in immediate succession, they both may be in the spotlight and can be encoded on a given fixation. The ASM is therefore a guidance-by-attentional-gradient model.

In the ASM, the times needed to complete both lexical access and saccadic programming are determined by the number of cycles that the two networks require to settle into stable activation patterns. As in E-Z Reader, the visual input to the word identification system is affected by retinal acuity limitations. Thus, the activation patterns that represent letter features become more "degraded" (i.e., the activation values of the units representing the letters decrease and are more prone to noise) as they are encoded further from the fovea, especially for letters that share many features with other letters. This degradation allows the model to account for the finding that word identification becomes more difficult as the distance between the word and the fovea increases (Morrison & Rayner 1981).

Although Reilly (1993) does not provide a detailed account of his model's performance, the ASM does simulate a few of the basic phenomena related to eye-movement control in reading. For instance, the model generates mean fixation durations and saccade lengths that are in close agreement to values that have been reported in the literature. In contrast to E-Z Reader, however, the ASM has not fitted to the various word-based measures, nor has it been shown to generate means and distributions for the different frequency classes of words. Nonetheless, because the amount of training that the word-recognition module receives on each word is proportional to each word's frequency of occurrence, the model does predict that low-frequency words are fixated longer than high-frequency words. Moreover, because two successive short words are

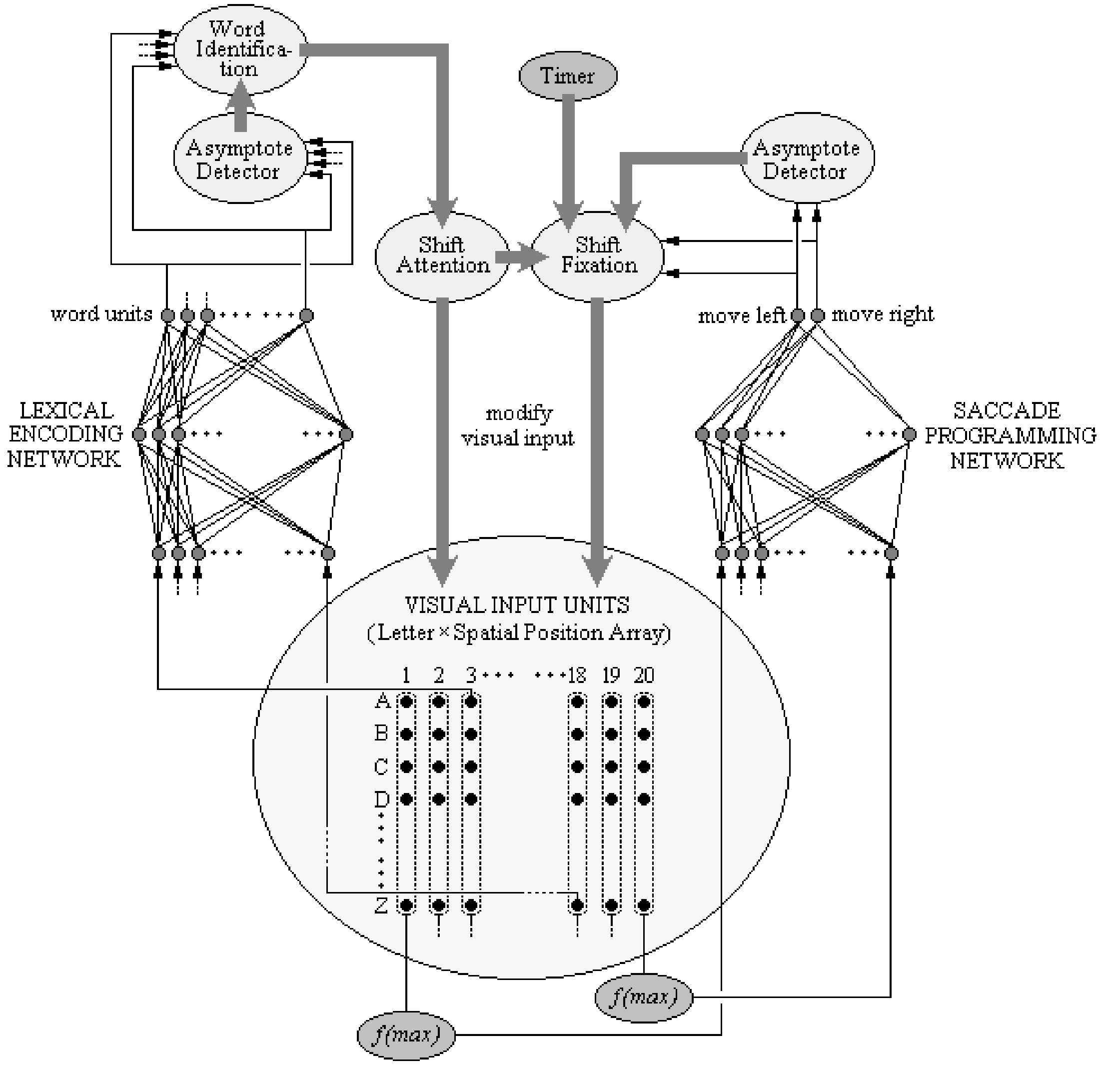

Figure 12. A schematic diagram of the *Attention-Shift* model (Reilly 1993). In the model, visual input is represented by an array of 26 letters that can be in any of 20 different spatial locations (position 8 is the center of the fovea). The core of the model consists of two connectionists networks that work in tandem to identify words and move the eyes. The first network, labeled "Lexical Encoding" in the-Figure, has as its input the activation values of each letter from the 16 central spatial positions. This information is used to identify individual words, which are represented by the word units as unique 8-bit patterns. The input to the second network, labeled "Saccadic Programming" in the Figure, are the maximal values from each spatial position, which is used to compute the direction and amplitude of the saccades. The "Asymptote Detectors" determine when the networks have settled into stable activation patterns, and thus provide an index of processing time. Word identification causes attention to shift, which modifies the visual input by reducing the activation values of unattended spatial input units (this is represented by the thick arrows in the Figure). Attention shift also enable saccades, which are executed after the "Saccadic Programming" network has settled into a stable pattern or after a certain time interval (which is determined by the "Timer"). Saccades also modify the visual input by boosting the activation values of the letters in the next word.

sometimes encoded in parallel, the model is able to account for the skipping of short words, as well as parafoveal preview benefit. It is of interest, though, that the ASM does not account for either of these phenomena in the same way that E-Z Reader does. In our model, skipping occurs whenever the word being fixated is identified, attention shifts to the next word, and it too is identified (in the parafovea). Thus, the models provide quite different accounts of the same phenomena: Whereas the ASM (a guidance-by-attentional-gradient model) allows some degree of parallel processing of upcoming words, E-Z Reader (a sequential-attention-shift model) allows for parafoveal processing via covert shifts of attention. There is one noteworthy difference between the two models with respect to parafoveal processing, however: In contrast to our model, the ASM does not explain why predictable words are skipped more often than less predictable words.

Finally, like E-Z Reader, saccadic programming in ASM is prone to noise, so that individual words can be refixated and/or skipped due to simple oculomotor error. Reilly (1993) has not, however, demonstrated that the model can reproduce the complex dependencies between the locations and durations of launch sites and the landing site distributions. We therefore contend that, unlike E-Z Reader, the ASM has not – at present – provided a complete account of the visual, oculomotor, and language-processing determinants of eye-movement control in reading.

4.9. EMMA

Salvucci (2000; 2001) has recently extended many of the core principles in E-Z Reader to provide a general theory of the interrelationships among cognition, attention, and eye movements. This model, EMMA (Eye Movements and Movements of Attention), has been implemented within the *ACT-R/PM* production-system architecture (Anderson & Lebiere 1998; Byrne & Anderson 1998). *Productions* are procedural condition-action pairs (i.e., "if-then" statements) that perform operations on units of declarative knowledge. For example, the production:

If (letter$_1$ = "c" & letter$_2$ = "a" & letter$_3$ = "t"),
then (word = "cat")

encodes the percept "cat," so that the meaning of the word can then be retrieved from semantic memory.

In EMMA, the encoding time for both words and objects, T_{enc}, is given by Equation 24. In Equation 24, the frequency of occurrence is scaled within the range (0, 1), ε is the eccentricity of the word or object (as measured by the angular distance between it and the fovea), and K and k are free parameters which scale the encoding time and eccentricity parameter, respectively. Like E-Z Reader, EMMA is a sequential-attention-shift model. EMMA also shares the following assumptions with E-Z Reader 7. First, encoding times are a function of both normative frequency and foveal eccentricity. Second, the actual amount of time that is required to encode a given object or word is determined stochastically by sampling random values from gamma distributions having fixed means (cf. Equations 1, 2, and 3, in E-Z Reader 7, and Equation 24, in EMMA) and standard deviations. Third, saccadic programming is completed in two sequential stages (the first being subject to cancellation by subsequent programs, the second not), the durations of which are also sampled from gamma distributions having fixed means and standard deviations. Finally, although saccades are directed towards the centers of their intended targets, they often deviate from their targets because of Gaussian motor error.

$$T_{enc} = \mathrm{K}\,[-\log(\text{frequency})]\,e^{k\varepsilon} \tag{24}$$

Although EMMA and E-Z Reader share many common assumptions, there are a few notable differences. First, in contrast to our model, encoding time in EMMA is not modulated by predictability, so that the model cannot account for predictability effects (Balota et al. 1985; Ehrlich & Rayner 1981; Rayner & Well 1986). Second, the distinction between the first and second stages of lexical processing in E-Z Reader corresponds to the encoding and cognitive-processing stages in EMMA, respectively. As cognitive processing completes, it directs the visual system to encode additional information. However, because only the rate of encoding (and not cognitive processing) is modulated by normative frequency, EMMA cannot account for the interaction between parafoveal preview benefit and foveal processing difficulty (Henderson & Ferreira 1990; Kennison & Clifton 1995; Schroyens et al. 1999). Finally, in EMMA, foveal eccentricity is measured in terms of angular disparity rather than character spaces. Although this last difference between the two models is largely cosmetic, it allows EMMA to simulate tasks other than reading.

So far, EMMA has successfully predicted the patterns of fixation durations and locations in equation solving (i.e., mental arithmetic) and visual search tasks (i.e., subjects scan visual arrays of alphanumeric characters and indicate the presence of pre-defined targets). EMMA has also been fitted to the same six word-based measures used to evaluate E-Z Reader (i.e., the mean fixation duration and fixation probability values observed in the Schilling et al. 1998 sentence corpus). In each of these tasks, the core principles governing attention and eye movements were the same in the model, and only the productions mediating the central, or cognitive, components of the tasks were changed. We view the successes of EMMA as being very encouraging because they suggest that the core principles of the model (which are shared by E-Z Reader) are general enough to describe the link between cognitive processing and eye movements in a variety of task domains. These successes also provide converging evidence supporting the validity of the basic principles shared by E-Z Reader and EMMA. However, the link between cognitive processes and eye movements might not be as tight in tasks where there are no externally composed task demands (such as scene perception).

4.10. Reader

In contrast to all of the models discussed thus far (including E-Z Reader), this model attempts to explain reading in its entirety, including the encoding of visual features, lexical processing, semantic and syntactic analysis, and the schema-guided comprehension and abstraction of key ideas that normally occur during reading (Carpenter & Just 1983; Just & Carpenter 1980; 1987; Thibadeau et al. 1982). In this model, eye movements are tightly linked to cognitive processing. This coupling is based on two assumptions. The first is the *immediacy hypothesis,* which stipulates that each word is processed to the farthest extent possible when it is fixated. The second is the *eye-mind hypothesis,* which stipulates that the eyes remain fixated on a word until the processing on that word has been completed. Both the durations and locations of individual fixations are thus determined by the immediate processing of the word that is being fixated. Thus, Reader (like our model) is clearly a sequential-attention-shift model in that attention (and in the case of Reader, all cognitive processing) is sequentially shifted from one word to the next.

Reader was implemented as a computer simulation with a production-system cognitive architecture (Anderson 1983; Anderson & Libiere 1998; Newell 1990). In Reader, the productions are activation-based; that is, they direct activation towards units of declarative knowledge. These units of declarative knowledge, in turn, have thresholds that must be exceeded if the information is to be "active" in working memory (and thereby satisfy the conditions of other productions). The values of these thresholds are adjusted to modulate the cost associated with using each production. For example, the thresholds of those productions that mediate lexical access are adjusted to reflect each word's normative frequency of occurrence, so that low-frequency words take longer to identify (and are consequently fixated longer) than high-frequency words. Also, in the most recent version of the model (Just & Carpenter 1992), the amount of activation that is available to support processing is limited (and is a free parameter) so that individual differences in working memory capacity can be used to simulate individual differences in reading ability.

The major strength of the Reader model is its comprehensiveness. As mentioned above, the model attempts to explain the entire reading process and therefore does reasonably well simulating a number of language-related reading phenomena, such as word-frequency effects, increased reading times on lexically ambiguous words, and the processing difficulties which are found with syntactically ambiguous sentences. Unfortunately, the model is extremely complex (it consists of 225 productions; Just & Carpenter 1987), and thus lacks the conciseness and controllability of other computational models (e.g., the inner workings of the model are not transparent, and can only be described verbally). It is also difficult to evaluate the model's performance because it depends upon the complex interplay of the productions, many free parameters, and the regression weights on several independent variables (e.g., whether or not a word is the first in a sentence) that are necessary to convert production cycles (arbitrary units of time) into processing time. Furthermore, the model only makes predictions about the locations of fixations at the level of individual words, using a composite measure (*gazes*) that counts skipping as 0-msec fixation durations in the average. This means that the model does not really make precise predictions about which word is fixated. In addition, apart from word-length effects, the model fails to account for any of the phenomena that are explained by the oculomotor models (e.g., landing site distributions).

In addition to the above shortcomings, Reader has been criticized because of the immediacy and eye-mind assumptions. With respect to the former, there is considerable evidence that the lexical processing of a word is often initiated before the word has been directly fixated (i.e., parafoveal preview: Balota et al. 1985; McConkie & Rayner 1975; Pollatsek et al. 1992; Rayner 1975). Furthermore, the depth of linguistic processing assumed before the eyes are allowed to move seems somewhat implausible. With respect to the eye-mind hypothesis, as we have noted a couple of times, there is evidence that the normative frequency of $word_n$ can affect how long the eyes remain on $word_{n+1}$ (Rayner & Duffy 1986; Rayner et al. 1989). These spillover effects indicate that the eyes often leave a word before the processing of that word is complete, contrary to the eye-mind assumption. Moreover, it seems quite implausible that each word can be encoded to the linguistic depth assumed in the model before an eye movement is programmed. This would produce fixation durations (and gaze durations) much longer than those usually encountered in normal reading. Thus, even if eye movements during reading are partially guided by language processing, the Reader model greatly over-simplifies how this occurs.

4.11. Comparison of the models

The processing models extend the theoretical coverage of the oculomotor models by attempting to specify how *the* key component of reading – word identification – affects (and is affected by) both the visual and oculomotor systems. This is important because a large number of linguistic variables have well-documented effects on eye movements during reading (for reviews, see Rayner 1998; Rayner & Duffy 1988; Rayner & Sereno 1994). Indeed, much of the interest surrounding the use of the eye-tracking methodology is that it affords a relatively non-intrusive, on-line way to study language processing. Of course, the processing models are not equally successful in handling the phenomena addressed by the oculomotor models. Table 1 lists the various eye-movement phenomena that have been observed during reading (as we discussed earlier in this article), and which E-Z Reader can explain. In Table 1, we have also presented for comparison a summary of the performance of the other eye-movement control models with respect to each of these phenomena. Thus, we have indicated whether or not (or the extent to which) each of the models can account for particular phenomena. A "Yes" indicates that the model can explain a result; a "No" indicates that (as the model is currently instantiated) it does not; finally, in some cases, we have indicated that the model provides a limited (labelled "Ltd") account in that the account is incomplete.

Table 1. *A comparison of the Reading Models[a,b] with respect to reading-related phenomena[c] that are explained by the E-Z Reader Model (See Table 1 Notes located at end of main Notes section.)*

	Minimal-Control	Strategy-Tactics	Word-Targeting	Push-Pull	SWIFT	Glenmore	Mr. Chips	Attention-Shift	E-Z Reader	EMMA	Reader
	Oculomotor ←			Oculomotor-Cognitive Dimension				→	Cognitive		
Reading Phenomena	POC	POC	POC	POC	GAG	GAG	GAG	GAG	SAS	SAS	SAS
Landing Site Distributions	No	Yes	Yes	Ltd	No	Yes	No	Yes	Yes	No	No
Systematic Range Error	No	Yes	Yes	No	No	Ltd	No	No	Yes	No	No
Word-Based Measures	Ltd	No	No	Ltd	Ltd	Yes	No	Ltd	Yes	Yes	Ltd
Frequency Effects	No	No	No	No	Yes	Yes	No	Ltd	Yes	Yes	Ltd
Parafoveal Preview	Ltd	No	No	No	Ltd	Yes	Ltd	Ltd	Yes	Ltd	No
Spillover Effects	No	No	No	No	No	Yes	No	No	Yes	No	No
Costs for Skipping	No	No	No	No	Yes	Yes	No	Ltd	Yes	No	No
Predictability Effects	No	No	No	No	Yes	No	No	No	Yes	No	Yes

Table 1 indicates that E-Z Reader handles the phenomena discussed in this article. Of course, one might argue that the inventory of phenomena in Table 1 is incomplete, and that there are also other ways by which to evaluate a computational model. Let us examine each of these objections in turn. First, we acknowledge that Table 1 is incomplete. For example, it does not include neighborhood effects (Perea & Pollatsek 1998; Pollatsek et al. 1999a) or lack of case change effects across fixations (McConkie & Zola 1979; Rayner et al. 1980). For E-Z Reader to be able to account for these effects, it would be necessary to extend the model to account for how letter processing maps onto word identification (which is something that we intend to do in future research). Nevertheless, the phenomena contained in Table 1 represent a substantial body of research and are not trivial to explain (as indicated by the fact that many of models have difficulty explaining a majority of them). Moreover, there is obviously some consensus that these phenomena are important "benchmarks" in that so much effort has been spent developing models to explain these phenomena. Thus, although we agree that Table 1 is not exhaustive, it does represent the basic results that any viable model of eye-movement control in reading must be able to explain.

A second criticism – that there are other ways to evaluate computational models – is more difficult to address because what constitutes a "good" model is somewhat subjective (see Hintzman 1991, for a discussion of some of the issues related to the evaluation of computational models). Rather than arguing that our model is better than another, we believe that it may be more productive simply to discuss why we think our model is a "good" model. To begin with, E-Z Reader describes and summarizes a large body of data (those in Table 1). Moreover, it does so in a relatively simple fashion. Although successive versions of the model have included additional free parameters, we have always maintained our "minimalist" approach to modelling; that is, we have added new parameters only when it was absolutely necessary (e.g., to explain some aspect of the data that could not otherwise be explained) or when it made the model more psychological or physiologically plausible.[17] Our reason for doing this is that we wanted the model to be transparent. That is, we wanted the model to be simple enough for us to understand why it worked and why – in some cases – it failed. (We believe that one of the major shortcomings of other modelling approaches, e.g., connectionism and production systems, is that the models are often too complicated to be summarized in a concise and precise manner.)

One final criterion that we use for evaluating our model is its utility as a heuristic device. That is, one measure of a model's usefulness is the degree to which it makes clear predictions that don't depend on specific settings of parameter values, but instead flow from the basic assumptions of the model. For example, prior to any attempts to fit the model, it was clear than an earlier version of our model (E-Z Reader 5; Reichle et al. 1998) predicted inflated fixation durations on word_n in cases where word_{n+1} is skipped. This prediction was subsequently confirmed (Pollatsek et al. 1986; Reichle et al. 1998; but see Note 4). Similarly, the model is currently being used as an analytical tool to evaluate the basic assumptions of other theories of language processing, as will be discussed in the next section of this paper. Finally, we believe that – with everything else being equal – it is better to have a model that at least has the potential to map the behavioral phenomena that are being explained onto their underlying neural processes. As the last section of this paper will indicate, we are currently striving to link the cognitive processes of E-Z Reader onto known brain structures.

On the basis of the preceding analysis, therefore, we conclude that the E-Z Reader model provides the most comprehensive and complete theory of eye-movement control in reading, while still being transparent enough that many of its qualitative properties flow from basic assumptions rather than from specific parameter values. In the final section of this article, we will briefly discuss the possible roles that E-Z Reader may play in future reading research.

5. Future research

In this section, we will focus on a few of the ways in which the E-Z Reader model may be used to guide future reading research, and, conversely, how this research may guide the development of future reading models. This discussion will focus on two main issues. First, we will briefly discuss how the model has been used as an analytical tool to examine some key assumptions about eye movements and language processing. More specifically, our discussion will focus on the ways in which the model might be used to better understand higher-level linguistic processing in the context of natural reading. Second, we will consider how recent advances in cognitive neuroscience have influenced our understanding of eye-movement control in reading, and then speculate on how our model might be viewed in light of this new information.

5.1. Language processing

The core principles of E-Z Reader have been adapted to several different task domains, which suggests that it is capturing the basic "engine" that drives eye movements in tasks like reading. However, as we have indicated above, it is incomplete, as it only takes into account certain relatively "low-level" aspects of the reading process (i.e., up to the level of lexical access). However, we are optimistic that as better quantitative descriptions of higher-order language processing are developed, additional processing modules could be interfaced with our model to expand the domain of the model. This would undoubtedly be beneficial for two reasons. First, our model could be used to help guide what to look for in the eye movement record to test theories of language processing. Second, because a large number of higher-level language processing phenomena are known to affect eye movements during reading (see Rayner 1998, Table 2), the capacity to simulate these results using language models could provide additional benchmarks for evaluating future models of eye-movement control. Two examples of this "bootstrapping" approach to understanding reading and language are discussed below.

5.1.1. Lexical ambiguity. There are now a large number of eye-movement studies (Binder & Rayner 1998; Dopkins et al. 1992; Duffy et al. 1988; Kambe et al. 2001; Rayner & Duffy 1986; Rayner & Frazier 1989; Sereno 1995; Sereno et al. 1992; Wiley & Rayner 2000) that have examined how lexically ambiguous words are processed during reading. The basic findings from this research suggest that both *meaning dominance* (i.e., the relative frequency of the various meanings of the ambiguous word) and contextual information influence the processing of such words. For ambiguous words with two equally likely meanings (e.g., "straw"), readers' gaze durations are longer on such words in neutral, contexts than on a control word matched in length and word frequency. However, when the prior context disambiguates the meaning that should be instantiated, gaze durations are no longer on the ambiguous word than on the control word. Thus, the contextual information helps guide the reader's choice of the appropriate meaning. For ambiguous words where one meaning is much more dominant than the other (e.g., "bank"), when the prior context was neutral, readers look no longer at the ambiguous word than the control word. However, when the subsequent text in the sentence makes it clear that the subordinate meaning should be instantiated, fixation times on the disambiguating information are quite long and regressions back to the target word are frequent (suggesting that the reader incorrectly selected the dominant meaning and now has to recompute the subordinate meaning). Conversely, when the disambiguating information that precedes the biased ambiguous word indicates that the subordinate meaning is instantiated, readers' gaze durations on the ambiguous word are lengthened. Apparently, the contextual information increases the level of activation for the subordinate meaning so that the two meanings are in competition (just as the two meanings of a balanced ambiguous word like "straw" are in competition in a neutral context). This general pattern of results has been interpreted in the context of the *Reordered Access Model* (Duffy et al. 1988) and the data have been simulated using a constraint-satisfaction framework (Duffy et al. 2001).

Using the basic principles of E-Z Reader, we were able to simulate the pattern of data present in these eye-movement studies. This was done by: (1) treating the subordinate meaning of ambiguous words as if readers were dealing with a low-frequency word; and (2) allowing disambiguating context to decrease the time required to complete lexical processing of ambiguous words. Although our early efforts indicated that the model can predict the gaze duration on the ambiguous target words, we were unable to simulate an important finding; namely, that spillover fixations are much longer for ambiguous words than for words matched to the frequency of the subordinate meaning (Sereno et al. 1992). However, the important point for this discussion is that we suspect that, by implementing aspects of the Reordered Access Model into the architecture of our model, progress can be made in understanding lexical ambiguity resolution in reading.

5.1.2. Morphology. A recent survey of prominent reading researchers indicated that one of the major areas of residual ignorance in the domain of reading research concerns the role of morphology in visual word identification (Kennedy et al. 2000). In the last few years, researchers have had some success investigating the role of morphology in word identification by examining how eye movements are affected by the morphemic variables during natural reading (Andrews et al., in press; Hyönä & Pollatsek 1998; 2000; Juhasz et al. 2003; Pollatsek et al. 2000). For example, Hyönä and Pollatsek (1998) examined the eye movements of Finnish readers while reading long compound words embedded in single sentences. The data indicated, among other things, that although the whole-word frequency influenced fixation durations on the word, the frequency of the constituent words of the compounds influenced fixation durations as well. Interestingly, the effect of the frequency of the second constituent was first seen a bit later in processing than the effects of either the frequency of the first constituent or the frequency of the whole word (i.e., on the duration of the second fixation on the word instead of the duration of the first fixation of the word). These findings suggest that access of the compounds is a "race" between a direct lexical look-up process and a compositional process in which the components are assembled (a similar conclusion comes from a study of English suffixed words; Niswander et al. 2000). E-Z Reader 7, which already includes races between various components, is a natural framework to be expanded upon to explain such phenomena. However, expanding the model in this direction is not trivial, as it entails positing that units smaller than "the set of letters between the spaces" can influence the decision of when to move the eyes. Thus, among other things, one has to think carefully about which letter subsets of a word can play an active role in this decision. We are currently working on an expanded version of the model that simulates the major trends that were observed in these data (Pollatsek et al. 2003).[18]

5.1.3. Conclusion. Our discussions of lexical ambiguity and the role of morphology in word identification were meant to illustrate how our model of eye-movement control might be used to advance our understanding of language-related phenomena. These two examples were selected because researchers in both of these areas have made extensive use of data from eye-movement experiments and because explaining these phenomena clearly involved relatively small increments in the development of our model. Of course, this is not to say that eye movements have not already been used in a productive manner to address other language-related questions; on the contrary, eye movements have been used to study a wide array of linguistic phenomena, including (but not limited to) other types of ambiguity resolution (e.g., syntactic and phonological ambiguity), semantic and repetition priming, anaphor and co-reference, and discourse processing (for reviews, see Rayner 1998; Rayner & Pollatsek 1989; Rayner & Sereno 1994). We think that E-Z Reader will also prove to be a useful platform from which to model these other psycholinguistic phenomena.[19]

5.2. Cognitive neuroscience

As mentioned at the beginning of this section, the last decade has witnessed unprecedented advances in our general understanding of the mind-brain relationship. New methodologies, such as brain-imaging (e.g., PET, fMRI), electrophysiological recording (e.g., EEG), and single- and multiple-cellular recording techniques, have provided invaluable tools for examining the relationship between cognitive processes and their neural substrates. Likewise, new theoretical advances, such as those offered by biologically plausible connectionist models (Churchland & Sejnowski 1992; McClelland & Rumelhart 1986; Rumelhart et al. 1986b), promise to bridge the chasm that has until recently separated cognitive psychology from neuroscience (Churchland 1986). It therefore seems appropriate to consider how these recent advances will further our understanding of eye-movement control in reading, and, conversely, how cognitive models of reading might be used to guide neuroscience research.

5.2.1. The neural basis of reading. There is a growing consensus that most high-level and/or complex cognitive processes (e.g., language processing) are supported by large-scale networks that are themselves composed of several cortical and subcortical regions (Mesulam 1990; 1998; Posner & Raichle 1997). Consequently, it is not surprising that reading (which subsumes a large number of complex cognitive operations) is mediated by several of these large-scale networks. In the specific case of reading, these include (minimally) the networks that support vision, attention, eye-movement control, and language. In this section, we will provide a brief overview of these systems, and then speculate about how the language-processing system might interface with the systems that are responsible for programming and executing saccades.

The most natural place to begin an analysis of the neural systems underlying reading is the printed page. Visual processing of the text begins in the retina and progresses by way of the optic nerve to the optic chiasm and then the optic tract. From there, the visual "stream" splits into two pathways: The first projects to the lateral geniculate nucleus, and then the occipital cortex; the second innervates several subcortical structures, including one that is known to play a key role in eye movements – the superior colliculus (Leigh & Zee 1999; Sparks & Mays 1990). On the basis of results from numerous electrophysiological recording experiments with non-human primates, it has been estimated that there are 30 or more distinct cortical areas that are involved in vision (Felleman & Van Essen 1991; Maunsell & Newsome 1987; Van Essen & DeYoe 1995), although many of these areas perform functions that are less central to reading (e.g., motion perception). However, the low-level visual features (which comprise graphemes) are extracted and represented within the primary visual and extrastriate cortices (Grill-Spector et al. 1998).

The visual-processing stream continues on past this first analysis via two anatomically and functionally distinct pathways (Maunsell & Newsome 1987; Sagi & Julesz 1985; Ungerleider & Mishkin 1982; Van Essen & DeYoe 1995). The ventral, or "what," pathway extends along the inferior temporal cortices, and is thought to play an important role in feature integration and object recognition (Ishai et al. 1999; Tanaka 1996). Because words can be considered to be visual objects, the ventral system has also been implicated in the integration of those visual features which are necessary to represent visual word forms (Cohen et al. 2000; Poldrack et al. 1998). However, the location of the word-form area(s) remains controversial (see Posner et al. 1999a; 1999b; and Price 1997), and there is some evidence suggesting that the left medial extrastriate cortex is also intrinsically involved in the recognition of word forms (Peterson et al. 1989; 1990; Pugh et al. 2000).

The dorsal, or "where," pathway is thought to represent spatial information, such as the relative positions and orientations of objects (Ungerleider & Haxby 1994; Ungerleider et al. 1998). (For this reason, the dorsal system may also provide an interface between perception and action; Goodale & Milner 1992.) The dorsal pathway has also been implicated in visuospatial attention. In particular, the regions around the intraparietal sulci (i.e., the *parietal eye fields*) are thought to be central components of the visuospatial attention network. The other components include the superior colliculus (part of the mid-brain), the pulvinar nucleus of the thalamus, and a region that includes the precentral sulci/gyri and the posterior tips of the superior frontal sulci (i.e., the *frontal eye fields*) (Corbetta et al. 1993; Goldberg 1994; Kim et al. 1999; Leigh & Zee 1999; Luna et al. 1998; Rafal & Robertson 1995; Sweeney et al. 1996). Recent neuroimaging and electrophysiological recording research suggests that this network is involved in both covert and overt shifts of visuospatial attention, and that covert attention is probably represented in motor (more specifically, eye movement) coordinates (Corbetta 1998; Kim et al. 1999). This attention network also modulates both the analysis of objects in the ventral visual-processing pathway (Corbetta 1998) and perceptual processing in the striate and extrastriate cortices (Somers et al. 1999).

Although much less is known about language than the other components of reading, a long history of neuropsychological evidence (Caplan 1992) and a large number of more recent neuroimaging experiments indicate that the left inferior frontal gyrus (*Broca's area*) and the posterior part of the left superior and middle temporal gyri (*Wernicke's area*) are the two major language-processing areas. Both areas are engaged by a variety of receptive and expressive language tasks, including: (1) reading (Bavelier et al. 1997; Binder et al. 1997); (2) speech comprehension (Binder et al. 1997; Caplan et al. 1999; Schlosser et al. 1998; Stromswold et al. 1996); and (3) speech production (Bookheimer et al. 1997; Müller et al. 1997). The exact functional roles of these two language-processing areas are not known, but it has been suggested that Broca's area is involved in articulatory and syntactic processing, and that Wernicke's area supports lexical and semantic processing (Mesulam 1990). This hypothesis is (in part) based on the close proximity between Broca's area and the primary motor cortex. Wernicke's area, which receives input from the primary auditory cortex, may play a large role in lexical processing, such as binding the phonological word forms to their semantic representations (which are distributed elsewhere in the associative cortex; Mesulam 1998).

Because a single language network is presumably used to understand both written and spoken language, one of the central questions in reading research has been: How are the graphemes on a printed page converted into linguistic-based codes? The results of several recent neuroimaging experiments suggest that the left angular gyrus (which is located in the posterior part of the inferior parietal lobule) plays a critical role in computing grapheme-to-phoneme correspondences (Horwitz et al. 1998; Pugh et al. 2000). Because the left angular gyrus lies at the juncture of the extrastriate cortex and Wernicke's area, it is ideally situated to convert the orthographic word forms into their phonological counterparts. From the angular gyrus, the phonological word forms could then be used to gain access to semantic representations via Wernicke's area.

With respect to the time course of orthographic, phonological, and semantic processing, a recent meta-analysis (Posner et al. 1999a; 1999b) provides compelling evidence that key components of word-form processing can be completed within the time window that is necessary for it to function as a signal to initiate saccadic programming. The results of a recent ERP experiment, for example, indicate that certain aspects of lexical processing (e.g., word frequency) can be discerned within 120–150 msec of word onset (Sereno et al. 1998). This would leave plenty of time (up to 130–180 msec) for the oculomotor system to program a saccade if one assumes a 250–300 msec fixation. This is an ample amount of time to initiate and complete the labile stage of saccadic program. (In E-Z Reader 7, the time needed to do this, $t(M_1)$, is equal to 187 msec.) Of course, additional programming time is available to the extent that pre-attentive visual processing (which, in our model, subsumes the first 90 msec of processing) allows early processing of parafoveal words. Nonetheless, the Sereno et al. results only show that it is plausible that word identification drives eye movements; they do not demonstrate that word identification drives eye movements, nor do their data suggest how the linkage is made. One possibility is discussed in the next section of this paper.

5.2.2. Specifying a neural implementation. E-Z Reader provides a functionalist account of eye-movement control in reading. As we have stated on previous occasions (Reichle et al. 1998; 1999), the model is neither a deep model of linguistic processing, nor a deep model of oculomotor control; instead, the model is simply our attempt to specify the functional relationships among a few key parameters (i.e., word frequency, predictability, retinal acuity, saccadic accuracy) to explain the time course of word identification and eye-movement control during reading. Consequently, up to now, we have remained completely agnostic about how the cognitive operations in our model might be implemented in the brain. Given the current state of cognitive neuroscience, however, it seems appropriate that this question should at least be considered.

Our answer – which at this time is obviously very speculative – is depicted schematically in Figures 13 and 14. Figure 13 depicts the eye movements that might occur as word$_n$ and word$_{n+1}$ are in turn fixated, the cognitive processes (as specified in our model) which give rise to this pattern of eye movements, and the cortical and subcortical systems in which these cognitive processes occur. Figure 14 shows both where in the brain these neural systems are localized (indicated by the numbers in the text below), and how processing is coordinated among these systems.

The sequence of events depicted in Figures 13 and 14 begins when the visual image of word$_n$ hits the retina. After approximately 90 msec, the features that make up the word's orthographic form are being processed within the primary visual cortex (1). The individual letter features are then integrated at successively higher levels of the visual

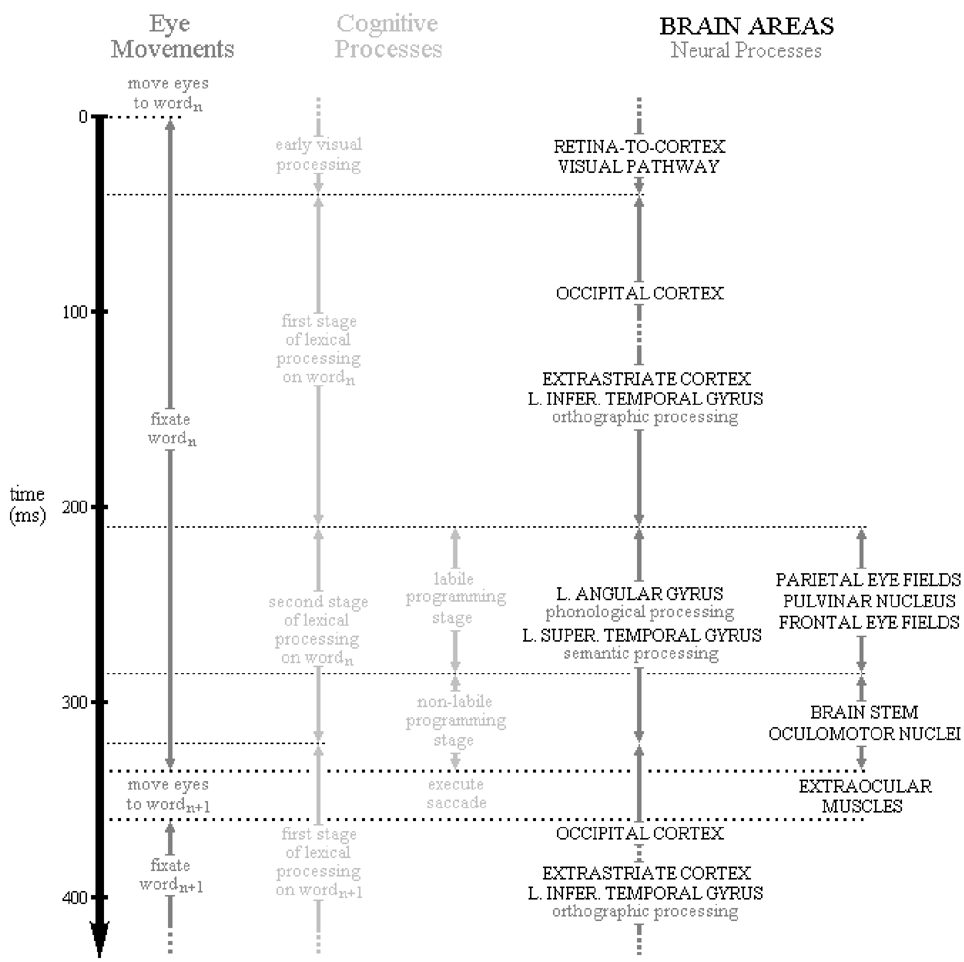

Figure 13. The time course of cognitive and neural processing during reading. The left side of the figure shows the pattern of fixations and saccades as the eyes move from word$_n$ to word$_{n+1}$. The center of the figure shows the cognitive processes specified by the E-Z Reader model. The right side of the figure shows the neural processes (and their locations within the brain) that may mediate these cognitive processes.

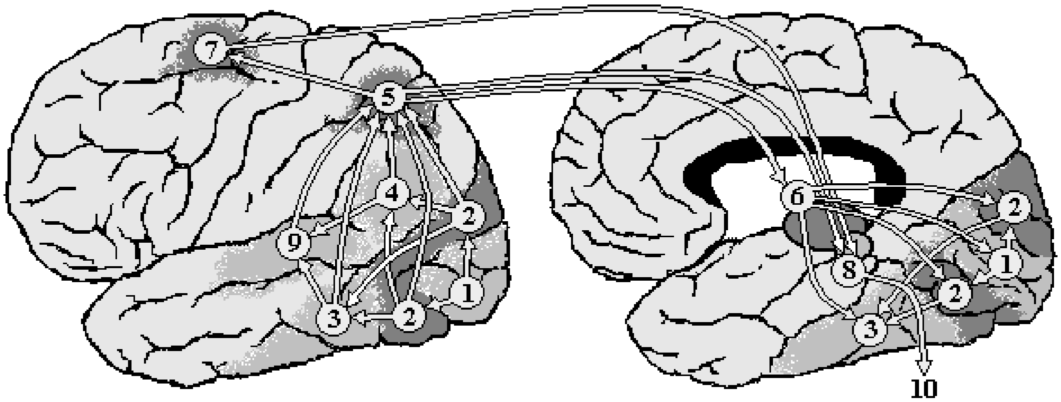

Figure 14. Sagittal views of the left lateral (left side of figure) and medial (right side of figure) cortical, thalamic (i.e., pulvinar nucleus), and mid-brain (i.e., superior colliculus) structures that may mediate the control of eye movements during reading. The number in the figure correspond to the following brain structures: (1) primary visual cortex (Brodmann's Area [*BA*] 17); (2) extrastriate cortex (*BA*s 18 & 19); (3) inferior temporal gyrus (*BA*s 20 & 37); (4) posterior inferior parietal lobule (i.e., angular gyrus; *BA* 39); (5) intraparietal sulci (i.e., parietal eye fields; *BA*s 7 & 40); (6) pulvinar nucleus of the thalamus; (7) superior prefrontal and posterior superior frontal gyri (i.e., frontal eye fields; *BA*s 6 & 8); (8) superior colliculus; (9) posterior middle and superior temporal gyri (i.e., Wernicke's area; *BA*s 21 & 22); and (10) the motor circuits of the brainstem which control the extraocular muscles and actually move the eyes. Although the figure only shows the left hemisphere, the right-hemisphere homologues of structures 1, 2, 5, 6, and 7 are also components of the visuospatial, attention, and oculomotor networks. Finally, the processing pathways among the areas depicted in the figure are not the only pathways that are known to exist; rather, the figure shows a few of the major pathways that have been shown to exist and which have a pattern of connectivity that is sufficient to support those cognitive processes that are important components of reading.

system as processing cascades from the striate to the extrastriate cortex (2). After approximately 150–250 msec, word_n's orthographic form has been assembled in the left extrastriate cortex (2) and/or left inferior temporal gyrus (3), and this orthographic word form has been used to either access or assemble its phonological representation within the left angular gyrus (4).

Up to this point in time, both the eyes and attention have been focused on word_n. With the partial (i.e., orthographic and/or phonological) identification of word_n, however, the parietal eye fields (5) disengage visuospatial attention. The pulvinar nucleus of the thalamus (6) then moves the attentional "spotlight" forward, so that the frontal eye fields (7) and superior colliculus (8) can start using the low-spatial frequency information (e.g., word length) from the primary visual cortex to begin programming a saccade to word_{n+1}. This saccadic program takes (on average) approximately 240 msec to complete. During this time, the processing of word_n continues; its orthographic (2 & 3) and/or phonological form(s) (4) are used to access the word's meaning by way of connections through Wernicke's area (9) to various parts of the associative cortex. If the meaning is accessed before the saccadic program has been completed, then the pulvinar (6) enhances the processing of word_{n+1} (by shifting the internal attentional "spotlight" to the next word) and a preview benefit ensues. Otherwise, a saccade is executed by neural circuitry in the brainstem (10; see Leigh & Zee 1999) and the extraocular muscles, thereby moving the eyes move forward to word_{n+1}.

Again, it is important to note that saccadic programming in our model is initiated after the first stage of lexical processing on an attended word has been completed, whereas attention shifts only occur after an attended word has been identified. Attention is thus allocated serially, from one word to the next as each new word is identified. The serial allocation of attention is necessary because it preserves the temporal order of the words, along with any syntactic information that may be dependent upon word order (Pollatsek & Rayner 1999). This is, of course, not to say that some properties of an upcoming word might not occasionally be encoded in parallel to those of the word that is currently the focus of attention; as reviewed earlier, there is some evidence that (under certain conditions) properties of two words can indeed be encoded in parallel (Inhoff et al. 2000a; Kennedy 1998; 2000; Kennedy et al. 2002; Starr & Inhoff, in press). However, we believe that the default process during normal reading is one in which attention is allocated serially, so that the meaning of each new word that is identified can be integrated into a larger sentence representation, which is at least partially dependent upon word-order information. Furthermore, the version of our model presented in this paper (E-Z Reader 7) includes an early, pre-attentive visual processing stage that surveys the "terrain" of the upcoming text. Orthographic irregularities in the parafoveal might therefore register through this pre-attentive visual processing. This would allow the model to account for parafoveal-on-foveal effects stemming from unusual word beginnings in a manner that does not depend upon the serial shifts of attention that are normally associated with lexical processing.

6. Conclusion

Our contention throughout this paper has been that, although E-Z Reader does not provide a deep explanation of language processing, vision, attention, or oculomotor control, it does provide a viable framework for thinking about how these different cognitive processes interact during the course of normal reading. Like the oculomotor models that were discussed earlier in this paper, E-Z Reader can account for the effects of several basic visual and oculomotor variables on eye movements. In contrast to these models, however, E-Z Reader also accounts for many of the important linguistic variables that are known to affect eye movements during reading. The model thus reflects our belief that, in order to account for the complex relationship between language processing and eye movements during reading, any adequate model of eye-movement control during reading will (almost by definition) have to include an account of language processing. Although our sketch of how the cognitive processes in E-Z Reader might map onto the neural systems responsible for guiding the eyes during reading is undoubtedly a gross over-simplification of what will undoubtedly turn out to be a much more complicated story, we would still argue that the mapping is precise enough to guide future cognitive neuroscience research.

Finally, it is worth emphasizing that E-Z Reader, like all of the other models reviewed in this paper, was developed primarily to explain the results of eye-tracking experiments. This should not be surprising because eye-tracking technology has proven to be an invaluable tool for studying reading. It is only natural that, as our understanding of eye movements and their determinants improve, this knowledge should be used to make inferences about the cognitive processes that occur during reading, and that these inferences should in turn be used to guide our modeling efforts. Because the last decade has witnessed unprecedented theoretical and methodological advances in the study of cognitive neuroscience, however, it is almost certain that these advances, too, will guide the development of the next generation of reading models. Like eye-movement data in the past, the discoveries of tomorrow will provide important guideposts for developing and evaluating future models.

ACKNOWLEDGMENTS
We thank our colleagues Charles Clifton, Susan Duffy, Julie Fiez, and Charles Perfetti for their many helpful discussions about various issues discussed in this article. Preparation of the article was supported by Grants HD17246 and HD27650 from the National Institutes of Health. The second author was also supported by a Research Scientist Award (MH02550) from the National Institute of Mental Health and by the Leverhulme Trust as a Leverhulme Professor at the University of Durham.

Requests for reprints should be sent to: Erik Reichle, Department of Psychology, Learning Research and Development Center, University of Pittsburgh, 3939 O'Hara Street, Pittsburgh, PA 15260; or via e-mail to: *reichle@pitt.edu.*

NOTES

1. Many models of word-identification have been proposed (Brown 1991; Bullinaria 1997; McClelland & Rumelhart 1981; Paap et al. 1982; Plaut et al. 1996; Seidenberg 1989; Seidenberg & McClelland 1989) to explain how orthography maps onto phonology and/or meaning, and how this process is affected by lexical variables (e.g., normative frequency, grapheme-phoneme regularity, etc.). Unfortunately, these models are generally limited in two ways: First, the entry point into these models is usually some highly abstract orthographic representation that bears little resemblance to the features that one might expect to be encoded by the visual system (e.g., homogenous retina acuity). Second, the models are generally fit to data from paradigms other than natural reading (e.g., lexical decision latencies). The models therefore say very little about the relationships among vision, eye movements, and word identification. Two interesting exceptions to this are McClelland's (1986) *programmable blackboard* model of reading and Shillcock et al.'s (2000) *split processing* model. The former model was designed to examine how fixation locations and visual acuity restrictions affect the model's word recognition performance; similarly, the split processing model was designed to examine how bisection of the visual field (and hence words) by the two cerebral hemispheres might explain why words are identified most rapidly when they are fixated near their centers.

2. We did not have a deep reason for choosing the name of our model. "E-Z Reader" was the name of a fictional character in a children's educational program *The Electric Company* in the U.S. and was clearly a spoof on the title of the movie *Easy Rider.*

3. Our discussion of parafoveal preview effects pertains to the processing of English. Indeed, there is some recent evidence (Deutsch et al. 2000; 2003) that indicates that, in Hebrew, morphological previews (in the form of the root morpheme, which is distributed throughout the word) provide preview benefit effects.

4. There is currently some disagreement regarding the extent to which the duration of a fixation prior to a skip is inflated. While there are reports of such an effect (Pollatsek et al. 1986; Reichle et al. 1998), others have reported null effects (Engbert et al. 2002; Radach & Heller 2000). In a very recent examination, we found effects on the order of 23 msec prior to a skip.

5. There is some dispute concerning the influence of "higher order" variables on where readers fixate. For example, Lavigne et al. (2000) reported that the eyes moved further into a word when that word was both high-frequency and predictable from the prior context. However, Rayner et al. (2001) and Vonk et al. (2000) found no such effect. In addition, Underwood et al. (1990; see also Hyönä et al. 1989) reported that the eyes moved further into words when the informative part of the word was at the end of the word. But Rayner and Morris (1992) and Hyönä (1995b) were unable to replicate this finding. On the other hand, there appears to be general agreement that an orthographically irregular letter cluster at the beginning of a word results in the eyes' initial landing position deviating toward the beginning of the word (Beauvillain & Doré 1998; Beauvillain et al. 1996; Hyönä 1995b).

6. A single set of parameter values were used in all of the simulations reported in this paper. These values were estimated by completing multiple grid-searches of the parameter's space so as to find the set that yielded the best overall fit to the Schilling et al. (1998) sentence corpus. For a complete description of our grid-search procedure, see the Appendix of Reichle et al. (1998).

7. Strictly speaking, Equation 1 produces word length effects (holding the eccentricity of the center of the word constant) only if the word straddles the fixation point. We used the arithmetic mean of the absolute distances in these formulas because of computational simplicity. However, if this were changed to some other combination rule (e.g., the geometric mean), then the equation would predict word length effects in all cases.

8. Frequency and predictability are not the only (nor necessarily the best) predictors of the time needed to identify a word in text. One problem with using frequency is that, even if the number of times a reader sees a given word in print was a perfect predictor of the time to identify the word, the Francis and Kučera (1982) norms (and other norms) are derived from corpuses that are unlikely to be representative of the texts that most readers encounter. (Another limitation of the Francis & Kučera norms is that they are derived from a fairly small corpus – only one million words.) Likewise, the predictability norms are also very crude estimates of how sentence context affects "on-line" lexical processing; in contrast to what actually happens during natural reading, the readers in these close-task studies have no visual information about the target words, but unlimited time to use all of the words in the sentence prior to the targets to guess their identities. Finally, the time needed to identify a word is likely to be a function of many other variables, including its part of speech, its concreteness, and the frequency with which it is encountered in spoken language. In summary, then, our decision to use frequency and predictability was not based on any a priori belief that these variables provide a complete explanation of lexical processing during reading. Instead, we are using them because they are known to produce significant effects in reading, and because they are clearly important determinants of word identification speed (i.e., how often a reader has seen the word before and how much top-down influence there is on the word).

9. In the current version of the model, for simplicity, attentional processing of $word_{n+1}$ (or words in general) is assumed to begin only when early visual processing of the entire word is completed. We are currently exploring versions of the model in which this assumption is relaxed, and attentional processing can begin when the early visual processing of parts of words is complete.

10. In our model, both the early pre-attentive visual processing and the non-labile stage of saccadic programming were halted during actual saccades. The former assumption was made because there is evidence that virtually no visual information is extracted during eye movements (Ishida & Ikeda 1989; Wolverton & Zola 1983). The latter assumption was necessary to ensure that a saccade could not be initiated while the eyes were already in motion. It should be noted that lexical processing does continue during saccades (Irwin 1998).

11. Figure 6 indicates that the model is underestimating the durations of single fixations. This problem stems from our increased estimate of the time needed to complete the labile stage of saccadic programming (i.e., $t(M_1) = 187$ msec). Because this "competitor" takes longer completing the "race" that determines whether or not a word will be refixated (i.e., the race between L_1 and M_1), the predicted durations of the first of two or more fixations is slightly too long, as indicated by the fact that the first-fixation durations are similar in length to the single-fixation durations. This also causes the single fixation durations for lower frequency words to be a bit too short. We don't think this is a major conceptual problem, as the primary goal in our simulations was to fit first-fixation durations and gaze durations rather than single-fixation durations. The problem seems fixable, however, by reducing $t(M_1)$ a bit and increasing the effect of frequency on the first stage of lexical access a bit. These changes shouldn't produce any catastrophic effects on other aspects of the fit, although perhaps the gaze durations may not fit quite well as in the current simulation.

12. We did not actually examine the landing site distributions in the Schilling et al. (1998) data because there were too few observations and because the properties of the distributions that we wanted to simulate are quite robust and have been reported in several places (e.g., McConkie et al. 1988; 1991; Rayner et al. 1996).

13. Interestingly, Vitu et al. (2001) recently reported an inverted optimal viewing position effect in reading in which readers' fixations were longer when they fixated near the center of a word than when they fixated away from the center of the word (when only one fixation was made on the word). Like Rayner et al. (1996), Vitu et al. also found frequency effects such that low-frequency words were fixated longer than high-frequency words.

14. In its current version, the model predicts that people will read about as effectively in a moving window condition in which the word to the left of fixation ($word_{n-1}$) and the fixated word ($word_n$) are visible as when the word to the right of fixation ($word_{n+1}$) and the fixated word ($word_n$) are visible (assuming word-boundary information is preserved to guide eye movements). This conflicts markedly with the findings in moving window studies (McConkie & Rayner 1975) where information to the right of the fixated word facilitates reading far more than information to the left of the fixated word. Perhaps the model does not depend critically on this attentional assumption and good predictions can be obtained with better attentional assumptions.

15. The model derives its name from Glenmore, Ireland – the place where much of the model was first developed (cf. Reilly & Radach 2003).

16. These results are open to alternative interpretations because the task was not natural reading, and thus did not actually require eye movements. Instead, the subject was required to read text on a computer monitor that was displayed through a stationary nine-character "window." The text was manually advanced via pressing keys that moved the text forward (1–9 character spaces) or backwards (1–3 character spaces), and a mask (covering 1, 3, or 5 character spaces) was placed over the center of the viewing window to occlude letters in the scotoma conditions.

17. For example, we previously argued that the last version of the model discussed in Reichle et al. (1998), E-Z Reader 5, is superior to an earlier version, E-Z Reader 3, even though the latter model provided a slightly better aggregate fit to the Schilling et al. (1998) data. This claim was based primarily on a qualitative argument: In E-Z Reader 5 (but not E-Z Reader 3), the rate of lexical processing decreases as the disparity between the word being processed and the fovea increases. Although this feature of E-Z Reader 5 makes the model more psychologically plausible, the counter-argument could be made that the lack of an improvement of the model's overall performance does not warrant the additional of two parameters. However, Salvucci and Anderson (1998; 2001) recently found additional evidence supporting our claim. Briefly, Salvucci and Anderson first replicated the Schilling et al. experiment with a different subject population, and then used several different eye-movement protocol algorithms to determine how well E-Z Readers 3 and 5 could account for the eye-movement data of individual subjects. They also examined how well the models could account for two sequential measures: (1) the proportions of saccades of each given length; and (2) the proportions of saccades of each given length following saccades of various lengths. The results of these analyses indicated that E-Z Reader 5 fit all three measures better than did E-Z Reader 3, and that E-Z Reader 5 in fact provided a better account of the finer-grained, sequential aspects of the observed eye-movement data. Moreover, these results suggest that E-Z Reader 7 (which also includes the visual acuity assumption) may also provide better quantitative fits than earlier, simpler, versions of the model.

18. Furthermore, our simulations to date (Pollatsek et al. 2003) indicate that a simple race model (i.e., a race between two independent processes, a direct look-up process and a constructive process) is unlikely to account for the observed pattern of data in Hyönä and Pollatsek (1998) and Pollatsek et al. (2000). This is an illustration of how modeling can help sharpen one's thinking about such issues.

19. Because the effects of higher-order language processing are often delayed and/or apparent over a wider temporal window than are the effects of lower-order language processing, the former may actually be less difficult to simulate than the latter. Paradoxically, it may be more difficult to evaluate a model's capacity to simulate higher-order linguistic effects for these same reasons.

TABLE 1 NOTES:

a. The primary references for the reading models are: (1) *Minimal-Control* (Suppes 1990; 1994); (2) *Strategy-Tactics* (O'Regan 1990; 1992b); (3) *Word-Targeting* (McConkie et al. 1988; Reilly & O'Regan 1998); (4) *Push-Pull* (Yang & McConkie 2001); (5) SWIFT (Engbert et al. 2002); (6) Glenmore (Reilly & Radach 2003); (7) *Mr. Chips* (Klitz et al. 2000; Legge et al. 1997); (8) *Attention-Shift* (Reilly 1993); (9) *E-Z Reader* (Reichle et al. 1998; 1999); (10) *EMMA* (Salvucci 2000a; 2000b); and (11) *Reader* (Just & Carpenter 1980; 1987; 1992; Thibadeau et al. 1982).

b. *GAG* indicates that the model assumes that attention is distributed as a gradient during reading (i.e., "guidance by attentional gradient"); *SAS* indicates that the model assumes the serial allocation of attention from one word to the next during reading (i.e., "sequential attention shift"); *POC* indicates that the model is primarily an oculomotor model and thus makes no specific assumptions about how attention is allocated during reading.

c. *Yes* indicates that a model can explain a result; *No* indicates that the model (as it is currently instantiated) does not explain a result; *Ltd* indicates that the model's account of a phenomenon is incomplete or limited (e.g., the model predicts parafoveal preview benefit, but the benefit is not modulated by foveal processing difficulty).

Open Peer Commentary

Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.

E-Z Reader's assumptions about lexical processing: Not so easy to define the two stages of word identification?

Sally Andrews
School of Psychology, University of Sydney, Sydney, 2006, Australia.
sallya@psych.usyd.edu.au

Abstract: E-Z Reader's account of the interaction between oculomotor and cognitive processes depends critically on distinguishing between early and late stages of lexical processing, because this distinction allows saccadic programming to be decoupled from shifts of attention. Precisely specifying the nature of this distinction has important implications both for current models of lexical retrieval and for the development of E-Z Reader 8.

The distinction between early and late stages of lexical processing is crucial to the E-Z Reader framework because the two stages play "unique functional roles": Completion of the first stage initiates programming of the next saccade, while completion of the second triggers a shift of the "spotlight of attention" to the next word, whether the programmed saccade has been executed or not. The target article provides little elaboration of the rationale or basis for this distinction, because the focus is on comparing E-Z Reader with other models of eye movement control. However, specifying the basis of the distinction is fundamental to the decoupling of saccadic programming from attention shifts that underpin the model's account of the interactions between oculomotor and cognitive processes during reading.

Reichle et al. (1998) labeled the first stage of lexical processing as the *familiarity check* and the second as *lexical access,* but emphasized that these processes "could be the product of a single word recognition module" (p. 133). An early estimate of the word's "familiarity" provides "the signal to the eye movement system that lexical access is imminent and that a saccade should be planned" (p. 133). It is not clear whether the target article's more neutral terms, "first and second stage of lexical access" (*L1* and *L2*), indicates a modification to this conceptualization of the two stages. If so, it is important to specify this change of view, because defining the two stages of lexical processing is important to current models of lexical retrieval and to evaluation and development of E-Z Reader.

Attributing a critical role to a "familiarity" estimate in determining when the eyes shift between words – a fundamental component of "normal reading" – has implications for current debates in the lexical processing literature. Many models of lexical retrieval are based on McClelland and Rumelhart's (1981) interactive activation framework, which uses lateral inhibition to discriminate the correct lexical candidate from similar words. Such frameworks predict that words that are similar to many other words (i.e., words with many "neighbors") should take longer to identify than words with few neighbors (but see Pollatsek et al. 1999a). This prediction has been refuted: The dominant finding for English words in both lexical decision tasks and naming tasks is faster identification of words with many neighbors (Andrews 1997).

Two major models of lexical retrieval that use an interactive activation network to simulate lexical memory – Coltheart et al.'s (2001) Dual Route Cascade model and Grainger and Jacobs' (1996) Multiple Read Out model – have attempted to explain facilitatory effects of neighbor similarity by adding a familiarity process to their computational implementation to allow fast lexical classification of words with many neighbors. However, both models assume that familiarity only contributes to the decision component of the lexical decision task – not to "true lexical retrieval" (Grainger & Jacobs 1996).

Conceptually, E-Z Reader provides a functional role for a familiarity index as an intrinsic component of normal reading while also allowing for additional effects of the second stage of lexical processing. Therefore, the model makes clear that "familiarity" and "complete lexical access" (Reichle et al. 1998) might both contribute to "normal reading" (Andrews & Heathcote 2001). However, it is not clear that this distinction is effectively operationalized in the model's current implementation.

E-Z Reader implements the two "stages" as different weightings of exactly the same parameters of log frequency and predictability. Although consistent with Reichle et al.'s (1998) claim that the stages reflect different computations derived from the same lexical system, this implementation fails to capture potentially important aspects of the conceptual description. Using word frequency to summarize the word-level attributes that influence lexical retrieval identifies processing with item-specific attributes. This ignores Reichle et al.'s (1998) suggestion that the distinction between *familiarity* and *lexical access* may be likened to that between *matching on the basis of global similarity* and *retrieval through reintegration,* as instantiated in many mathematical models of memory (e.g., Humphreys et al. 1989; Raaijmakers & Shiffrin 1981). If *L1* reflects average similarity of the perceptual input to items in memory, it is not solely a function of the item-specific attributes summarized by word frequency. A measure of average similarity to words, like the N metric of neighborhood density (Andrews 1997) seems a more plausible operationalization of the construct.

Reichle et al. acknowledge that using frequency and predictability as parameters of the equation predicting *L1* and *L2* is a simplification because "the time needed to identify a word is likely to be a function of many other variables" (target article, Note 8). In particular, Reichle et al. recognize the failure to incorporate neighborhood effects on word identification as a limitation of E-Z Reader. Reichle et al. recognize that extending the model to such phenomena is unlikely to be a simple matter of including measures such as N, as model parameters. There is, as yet, relatively little evidence about how neighborhood structure influences eye movements, but what is available suggests the story is complex.

Pollatsek et al. (1999a) found that words that produced facilitatory effects of neighborhood size in the lexical decision task yielded inhibitory effects on gaze durations when presented in sentence contexts, consistent with Grainger and Jacobs' (1996) claim that facilitatory effects are specific to the lexical decision task. However, a subsequent experiment that manipulated only lower frequency neighbors showed facilitatory effects of neighborhood size on measures of skipping, but no significant effects on fixation measures – consistent with the view that overall familiarity can influence the likelihood of making a saccade. The different effects of measures of neighborhood structure on skipping and fixation duration support the previous suggestion that different parameters may determine *L1* and *L2.*

The variable effects of N manipulations on eye movement data may also indicate that N is not the most valid measure of lexical similarity, particularly when applied across words of different lengths (Andrews 1997). Similarity between longer, multisyllabic words may be a function of syllabic or morphemic units that are not effectively captured by letter-based measures like N. As Reichle et al. recognize, extending the model to afford sensitivity to lexical similarity and within-word structure will require an account of "how letter processing maps onto word identification" (target article, sect. 4.11, para. 2) and may require changes to the current serial operation of "early visual" and lexical processing. Refinements will also be necessary to accommodate evidence that morphological subcomponents of words influence the timing of fixations (Andrews et al., in press; Pollatsek et al. 2000) which imply that "units smaller than 'the set of letters between the spaces' can influence the decision of when to move the eyes" (target article, sect. 5.1.2).

Reichle et al. are clearly sensitive to most of the issues I have raised. They acknowledge that their implementation of the two stages of lexical processing is a simplification adopted more for empirical convenience than theoretical commitment and emphasize that the model is not intended to provide a "deep explanation of language processing" but simply "a viable framework" for thinking about how cognitive processes interact during reading. However, it is precisely because of the remarkable success of the relatively simple framework offered by E-Z Reader in accounting for a number of complex features of reading behavior that it seems worthwhile to implement a more sophisticated set of assumptions about the factors influencing lexical processing to provide further support for Reichle et al.'s claim that "the intrinsic nature of the language processing during reading hinges upon word identification" (sect. 4.5, last para.). Validating and refining the distinction between the two stages of lexical access will be crucial to such developments.

ACKNOWLEDGMENT
This commentary was prepared with support from Australian Research Council grant A10007174.

How can selection-for-perception be decoupled from selection-for-action?

Cécile Beauvillain and Pierre Pouget
Laboratoire de Psychologie Expérimentale, UMR 8581 CNRS, Université René Descartes, Paris 5 92771, Boulogne Billancourt Cedex, France.
beauvi@idf.ext.jussieu.fr pouget@psycho.univ-paris5.fr

Abstract: Evidence is presented for the notion that selection-for-perception and selection-for-action progress in parallel to become tightly coupled at the saccade target before the execution of the movement. Such a conception might be incorporated in the E-Z Reader model of eye-movement control in reading.

A key assumption of the Reichle et al. model of eye-movement control in reading is that covert attention and eye movements are decoupled. Although the saccadic system relies on a spatial selection mechanism to calculate the saccade endpoint in the next word, shifting of attention refers to the process of disengaging the mechanisms responsible for identification so that they can be shifted to the next word. This conception is based on classic theories of attention – as applied to linear stage models – that assume that it functions to filter certain objects in the environment (Broadbent 1971; Treisman 1969). Capacity limitations at the perceptual stage was Reichle et al.'s central concept, and they considered selection its functional consequence. As with classical linear stage theories, Reichle et al. emphasize dissociation between attention and intention, as perception and action are viewed as relatively distinct. This leads to the hypothesis that the selection for processing is decoupled from the selection of the target for the next saccade. Instead, we think that selection does not require an attentional control system separate from action-perception cycles. Selection-for-perception and selection-for-action accompany each other within a spatial selection window. The preparation of a motor action, however, binds the perceptual system to the movement target. Then, both selection processes are strictly spatially and temporally coupled. This claim may be incorporated in recent modelizations of overt eye movements (Findlay & Walker 1999; Henderson et al. 1999) and covert shift of attention (Itti & Koch 2000), which use the concept of a saliency map, a topographic map that codes the position and relative salience of potential saccade targets.

A growing body of behavioral evidence given in tasks other than reading shows that the preparation of a movement induces a concurrent shift of attention toward the same location (Deubel & Schneider 1996; Hoffman & Subramaniam 1995). The preparation of an eye movement toward an obligatory spatial location enhances the perceptual processing of the stimulus at locations toward which the motor program is directed. We recently described a similar coupling of saccade programming and shift of attention at the intended landing position into spatially extended letter strings (Doré-Mazars et al., in press). Fine-grained information about the time and the location of the attentional shift was obtained by measuring the detection of a letter change that occurred tachistoscopically at different SOAs after the presentation of the letter string and before the saccade execution. The letter change could occur at different positions within the letter string. Given the variability of the actual landing positions within the letter string, it was possible to examine whether actual landing positions are coupled with corresponding attentional shifts. A strict and selective spatial coupling of attention at the saccadic endpoint was found in the last 50 msec preceding the saccade execution. Then, detection capability was shown to be at its best when the letter change location coincided with the actual landing position to decrease at other landing positions. This shows that attention shifted at the saccade target at the expense of the others when the saccade was prepared to be executed. However, before the selection of one precise saccade target, a salient, orthographically illegal cluster at the beginning of the letter string attracted attention automatically, independent of the intended landing position. At this stage, different targets seem to be selected for action and perception. However, when the command to move to a precise target location is sent to the motor effectors, the target that occupies this location receives prioritized visual processing. At this stage, both selection-for-action and selection-for-perception processes are strictly spatially and temporally coupled.

Recent brain-imaging studies have shown that covert spatial attention and overt oculomotor shifts are tightly integrated at the neural level. A large overlapping network of regions including FEF (frontal eye field), IPS (intraparietal sulcus), and STS (superior temporal sulcus) was active during both tasks (Corbetta et al. 1998; Nobre 2000). These studies, however, do not reveal the temporal relationship between the robust, attention-related signals obtained in the frontoparietal network during tasks requiring sustained covert attention to spatial locations in the visual field and the extrastriate activations. ER-fMRI studies, however, suggest that the engagement of a top-down attentional control circuitry that is centered on the dorsal posterior parietal and frontal cortex is the source of the facilitation of attended input in extrastriate visual cortex (Corbetta & Shulman 2002; Corbetta et al. 2000). Recent studies using ERP measurement revealed that directing covert attention to the location of a stimulus results in an enhancement of the P1 and multiple N1 components evoked by that stimulus, compared with the nonattended condition (Hillyard & Anllo-Vento 1998). Spatial attention exerts a selective amplification of sensory information flow in the visual pathways, the early signal enhancement around 80 msec after stimulus onset being localized in extrastriate cortex (Heinze et al. 1994; Martinez et al. 2001; Woldorff et al. 1997) and the later one (100–130 msec) in ventral areas specialized for pattern and object recognition (Mangun et al. 1997). Whether this enhancement in the extrastriate areas is coupled in time with the programming of an eye movement in the frontoparietal network has not yet been studied in humans. There is however, electrophysiological evidence of an activation enhancement around 50 msec prior to the monkey making a saccade to a target in the frontal eye field neurons (Schall & Hanes 1993) and in infero-temporal cortex (Chelazzi et al. 1993).

In short, the Reichle et al. model will likely need modification to integrate the above evidence. The spatial selection window in reading may be defined as a region including several words (Inhoff et al. 2000; Kennedy et al. 2002) that can be processed in parallel with a bias towards higher salience for the fixated word. During a given fixation, the point of maximum salience dynamically changes to be highest at the saccade word target before the saccade execution. In interactive activation models (McClelland & Rumelhart 1981), the processing systems (as lexical access) are controlled by the connections among different interconnected units (features, letters, and words) and are not capacity limited.

Please stop using word frequency data that are likely to be word length effects in disguise

Marc Brysbaert[a] and Denis Drieghe[b]

[a]*Department of Psychology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom;* [b]*Department of Experimental Psychology, Ghent University, B-9000 Ghent, Belgium.*
marc.brysbaert@rhul.ac.uk **denis.drieghe@rug.ac.be**
http://psyserver.pc.rhbnc.ac.uk/staff/mbrysbaert.html
http://allserv.rug.ac.be/~ddrieghe/

Abstract: Reichle et al. claim to successfully simulate a frequency effect of 60% on skipping rate in human data, whereas the original article reports an effect of only 4%. We suspect that the deviation is attributable to the length of the words in the different conditions, which implies that E-Z Reader is wrong in its conception of eye guidance between words.

A computational model is as good as the data it simulates. This is why Reichle et al. rightly pride themselves about the good fit of the model's outcome with human data. The human data predominantly come from a reading study, reported by Schilling et al. (1998), in which 30 college students read 48 sentences. According to Figure 6 in the target article, the observed frequency effects in this study were roughly 70 msec for gaze duration, 30 msec for first fixation duration, and a 60% for word skipping rate. What Reichle et al. did not mention is that the Schilling et al. study was originally designed to look at the word frequency effect under very controlled circumstances (i.e., with words that were matched on all other variables except for word frequency, and with sentence context that constrained the target words equally). Each participant saw a number of sentences with low frequency words (2 per million) and a number of sentences with high frequency words (141 per million). These frequencies probably coincide with the frequency classes 1 and 5 of Figure 6 in the target article. If we look at the data reported by Schilling et al. for these particular stimuli, we obtain a frequency effect of 67 msec for gaze duration and 35 msec for first fixation duration, but only 4% for skipping rate ("Subjects fixated on HF words 89% of the time and on LF words 93% of the time" – Schilling et al., p. 1,272). That is, for this particular subset of well-controlled stimulus words, in Schilling et al., the effects for gaze duration and first fixation duration agree well with the overall data used by Reichle et al., but this is not true for the skipping rate. How come E-Z Reader "correctly" simulates a 60% difference in skipping rate between low-frequency and high-frequency words, whereas in the human data there was only a 4% difference attributable to word frequency?

After a review of all previously published word skipping data, Brysbaert and Vitu (1998) concluded that the frequency effect on word skipping is 4% on average (i.e., exactly the effect reported by Schilling et al., as well), and that the effect was 9% for contextual predictability (i.e., very predictable words in a sentence are skipped, on average, 9% more often than unpredictable words). In addition, they observed a 60% difference attributable to word length: 2-letter words are skipped more than 60% of the time, whereas 10-letter words are virtually never skipped in first-pass reading. To us, these data strongly suggest that what Reichle et al. simulate in the lower part of Figure 6 is not so much a frequency effect on skipping rate but a word-length effect on skipping rate. The authors themselves are clearly aware of this problem, because in Rayner et al. (1998c, p. 256, footnote 3), they wrote:

> In our modelling, to minimize the number of parameters, we did not distinguish between frequency and word length effects. Thus "frequency effects" in our model are really a combination of frequency and word length effects because the two are highly correlated in our sample of text as in printed English in general.

For this reason, we were very surprised to see that in the present article they still refuse to report the data separately for word length and word frequency, even though the current model is supposed to have a mechanism to deal with the effects of the length of the parafoveal word (see Equation 1 of the target article). What we ask is that Reichle et al. give us a figure in which the word-skipping rates of the Schilling et al. corpus are shown as a function of word length and word frequency, together with the predictions of E-Z Reader. If these provide a good fit, we will rest our case. However, we strongly suspect that the model will largely overestimate the effect of frequency and underestimate the effect of word length. For this reason, until proven wrong, we still believe that E-Z Reader is fundamentally flawed in its conception of interword behaviour in general and word skipping in particular.

ACKNOWLEDGMENT
Denis Drieghe is a research assistant of the Fund for Scientific Research (Flanders, Belgium).

Reading the scene: Application of E-Z Reader to object and scene perception

Peter De Graef and Filip Germeys

Laboratory of Experimental Psychology, University of Leuven, Leuven, B-3000, Belgium. **peter.degraef@psy.kuleuven.ac.be**
filip.germeys@psy.kuleuven.ac.be
www.psy.kuleuven.ac.be/labexppsy/top/peterweb
www.psy.kuleuven.ac.be/labexppsy/top/filipweb

Abstract: We discuss five basic principles of E-Z Reader in terms of their potential for models of eye-movement control in object and scene perception. We identify several obstacles which may hinder the extrapolation of the E-Z Reader principles to nonreading tasks, yet find that sufficient similarities remain to justify using E-Z Reader as a guide for modeling eye-movement control in object and scene perception.

Eye-tracking has provided vision science with a powerful tool to unobtrusively monitor on-line perceptual and cognitive processing. Unfortunately, eye movement records generate a host of overt measures which all may (or may not) reflect some aspect of covert processing, leading to much debate about which measure would be most appropriate (e.g., Inhoff & Radach 1998). The most promising solution to this debate is to consider multiple eye movement measures simultaneously (Henderson et al. 1999). However, to do this, an integrated model is required that specifies the relations between the various overt measures as well as their correspondence to covert processes. This is precisely what Reichle et al. have achieved with E-Z Reader.

As users of eye-tracking methodology in object and scene perception, we can be only envious of this situation, yet at the same time Reichle et al. inspire some optimism with their suggestion that the basic principles of E-Z Reader may apply to other visual information processing tasks (sect. 4.9). We would like to evaluate the grounds for such optimism by examining five basic principles of E-Z Reader to determine whether and how they can be applied to the study of eye-movement control in object and scene perception.

First, according to E-Z Reader, the main engine of eye movements in reading is serial word identification. This makes sense given (a) the importance of individual word order and meaning to understand the whole sentence, and (b) the ease with which individual words can be segregated from a sentence. In scene perception, neither of these conditions is fulfilled. It is quite possible to achieve scene interpretation without identifying any of the objects in the scene, let alone identifying them in a certain order (Oliva & Torralba 2001). Moreover, object-background segregation in scenes can be so computationally demanding that object identification precedes figure-ground organization (Vecera & Farah 1997). Thus, object fixation times could reflect processing of object identity as well as background and figure-background relations. Consequently, in scene perception we may have to adopt a more general principle than object identification as the engine behind eye movements. One possibility is that the visual system simultaneously monitors the rate of activation buildup in an object lexicon and in an object localization module: As soon as both rates drop below a criterion level, the current fixation position is deemed to be suboptimal and an eye movement is planned.

Second, E-Z Reader assigns an important role to pre-attentive processing of the upcoming word. Given the rigid serial structure in which information needs to be acquired in reading, this implies that pre-attentive processing is restricted to the word that is about to become the saccade target. In scene exploration, however, there is no inherent spatial order in which objects need to be processed. Therefore, pre-attentive processing occurs for saccade bystanders as well as for the saccade target (Germeys et al. 2002). It will therefore be necessary to determine the spatial and temporal windows within which pre-attentive processing of a saccade bystander can influence that object's fixation duration or skipping probability once it has finally become the saccade target.

Third, E-Z Reader posits that attention shifts and eye movements are decoupled. We wonder whether the sparse and serial stimulus structure inherent in sentences may not be a necessary prerequisite for such a decoupling. In other words, autonomous attention shifts and eye movements may be possible only because the next relevant stimulus component is always easily discriminated on the basis of rudimentary boundary information. In scene perception, however, the next relevant stimulus component could be anywhere in the visual field; a more sophisticated process is required to mark the location of the next saccade target. As demonstrated by behavioral data (Deubel & Schneider 1996) as well as single-cell recordings in LIP (Colby et al. 1996), spatially selective shifts of visual attention appear to be that process, indicating a strong coupling between attention and eye movements. This implies that in scene perception we must invoke different mechanisms when our eye-movement records indicate refixations, spillover effects, and foveal-on-parafoveal effects.

Fourth, E-Z Reader elegantly limits the number of factors influencing fixation duration to visual acuity, word frequency, and word predictability. In object and scene perception these factors are also likely to play a role, although some may not be easy to estimate (e.g., what would constitute a good estimate of object frequency?). The relative importance of these factors is likely to be different in scene perception than in reading. Specifically, because scene identity is available early on in scene exploration (Biederman 1981) more subsequent fixations may show predictability effects than in reading where context develops more gradually. In addition, the list of factors influencing fixation times probably also needs to be extended. For example, ease of object identification has been argued to be a function of object orientation (Boutsen et al. 1998), object size (Theios & Amrhein 1989), and object camouflage (De Graef et al. 1990), all of which may have effects on eye-movement measures.

Fifth and finally, E-Z Reader capitalizes on the incorporation of very task-specific constraints in the model, such as the preferred saccade length in reading English. One could argue that this limits the generality of the model, but we feel such parameters are justified when they accurately reflect eye-movement behavior in the task under study. Moreover, while the parameter value is obviously task-dependent, the parameter itself may not be. Specifically, that preferred saccade length in reading English is estimated to be seven characters may be linked to the fact that the perceptual span for word encoding in reading English extends about eight characters to the right of fixation (Rayner et al. 1982). In other words, readers prefer to saccade to the edge of their perceptual span, a principle which may also apply to much less constrained tasks such as scene exploration (Shioiri & Ikeda 1989). Other task-specific constraints derived from reading data may be less suitable to extrapolate. For example, E-Z Reader assumes that all fixation times are sampled from a unimodal distribution. However, in other tasks, fixation time distributions may be multimodal, raising the question of whether fixation times in the various component distributions can all be modeled in the same fashion (De Graef 1998).

In summary, it would be unwise to extrapolate E-Z Reader to object and scene perception without careful consideration of task-specific differences in the interplay between visual processing, processing goals, attention, and oculomotor control. However, E-Z Reader does provide a valuable framework for thinking about the best design principles for a model of eye movements in object and scene perception.

ACKNOWLEDGMENTS

This work was supported by Concerted Research Effort Convention GOA 98/01 of the Research Fund K.U. Leuven and the Research Training Network "Perception for Recognition and Action" (RTN-2001-00107), under the direction of the European Commission.

Are there two populations of refixations in the reading of long words?

Karine Doré-Mazars, Dorine Vergilino-Perez, and Thérèse Collins

Laboratoire de Psychologie Expérimentale, UMR 8581 CNRS Université René Descartes Paris 5, France. **dore@psycho.univ-paris5.fr vergilino@psycho.univ-paris5.fr http://d.vergilinoperez.free.fr**

Abstract: This commentary focuses on the limitations of the E-Z Reader model in its attempt to explain refixation saccades in reading. Listing factors that influence probability of refixating leads the model to assume two sorts of refixations. However, taking into account data on the metrics of refixation saccades allows us to propose an alternative explanation for empirical observations reported in the literature.

In reading, the probability of refixating a word – that is, to read it with two fixations – is known to increase with word length, first fixations landing far from the word center and the decrease of word familiarity (McConkie et al. 1989). The E-Z Reader model accounts for these empirical data by making the following assumptions. The rapid parafoveal integration of the length of the to-be-fixated word results in the preparation of a refixation program. Although it remains implicit, this assumption suggests that the *decision* to refixate long words is taken before landing on the word. The computation of the refixation saccade is then initiated once the eyes land in the word. As the ability of the saccadic system to modify or cancel previous motor plans is now well documented in the literature (since the famous Becker & Jürgens [1979] study; see also Vergilino-Perez & Beauvillain, in press), the target model proposed that the refixation saccade program can be canceled during the first fixation on the word. The cancellation of a refixation saccade program would be more likely in a high frequency word than in a low frequency word because the progression of the first stage of the lexical processing is faster on the former than on the latter. Such an assumption is an elegant explanation for the word frequency effect on refixation probability (Inhoff & Rayner 1986; McConkie et al. 1989).

However, such a scenario does not fit with the classical interpretation of the effect of the first landing position on refixation probability. When the first fixation position was imposed at different locations in an isolated word (e.g., O'Regan et al. 1984), refixation probability increased when locations were far from the word center, a location usually called the optimal viewing position

(OVP). This relationship is expressed by a U-shaped curve. As a similar pattern was found in text reading (e.g., McConkie et al.1989; Vitu et al. 1990), it was popularly assumed that a decision to refixate was made because of errors in the execution of saccades which do not land on the intended saccade target. To integrate these empirical observations, the E-Z Reader model admits that a proportion of refixations is planned because of mislocated initial positions.

We would like to address two questions to the authors. First, whether their model assumes two populations of refixations. Second, whether the presupposed factors that affect the decision to refixate also play a role in the computation of the metrics of refixation saccades. Indeed, even if this model addresses the question of the refixation probability, nothing is said about refixation saccade metrics – for example, what is the target for the refixation saccade?

An experiment was conducted in our lab (Doré-Mazars et al. 2003) to examine these questions further during reading of isolated long words. High- and low-frequency words of 8, 10, and 12 letters were displayed in parafoveal vision. With this procedure, the launch site (eccentricity) and the parafoveal preview were held constant. Critical aspects of early work about the refixation decision are replicated here: both length and frequency effects, and also the classical U-shaped curve describing the relation between the refixation probability and the initial landing position on the word. For each initial landing position, we found an effect of the length and the frequency of the word, the amplitude of the first being more important than the second one.

More interestingly, we provide arguments for the view that refixations do not result from saccadic error but are preplanned and sometimes canceled. We observed that the distribution of landing positions in refixation cases is clearly leftward-shifted relative to single fixation cases. In addition, the examination of the refixation saccade amplitude demonstrates that the saccade is planned on the basis of the word length with no effect of the initial landing position on the word. Indeed, the slope of the linear regression between first and second fixation position close to 1 indicates that the refixation saccade is computed as a fixed motor vector applied irrespective of the initial landing position on the word. We replicate here previous findings indicating that the refixation saccade is preplanned in parafovea relative to the word length integrated at this time (Vergilino & Beauvillain 2000). The absence of a target for refixation saccades stands against refixations as corrective saccades. In such a framework, we interpret the difference in initial landing position on the word between single- and refixation cases found in our experiment as the consequence and not as the cause of the planning of refixation saccades. Of course, because of the inherent variability of the text-reading situation (e.g., in launching sites), some refixations could be caused by mislocated landing positions, but their proportion and metrics remain to be assessed. Moreover, while refixation probability was affected by word frequency, no role of this factor in the computation of the refixation metrics was observed in our experiment. Indeed, we found a frequency effect neither on the mean refixation saccade amplitudes nor on the slope of the linear regression. This result is compatible with the notion that the lexical processing that progresses throughout the first fixation is likely to cancel a preplanned refixation saccade. However, since the frequency effect on refixation probability is around 10%, as usually observed in the literature, we assume that only a small proportion of refixations would be canceled by lexical processing. Word processing plays only the secondary role in refixating of long words.

In conclusion, one of the future challenges of the E-Z Reader model is to take into account not only the factors that influence the decision to make a refixation saccade, but also those that determine its metrics, to better explain refixations in reading.

The game of word skipping: Who are the competitors?

Ralf Engbert[a] and Reinhold Kliegl[b]

[a]*Department of Psychology and Center for Dynamics of Complex Systems,* [b]*Department of Psychology, University of Potsdam, D-14415 Potsdam, Germany.* **engbert@rz.uni-potsdam.de** **kliegl@rz.uni-potsdam.de**
www.agnld.uni-potsdam.de/~ralf/
www.psych.uni-potsdam.de/people/kliegl/index-e.html

Abstract: Computational models such as E-Z Reader and SWIFT are ideal theoretical tools to test quantitatively our current understanding of eye-movement control in reading. Here we present a mathematical analysis of word skipping in the E-Z Reader model by semianalytic methods, to highlight the differences in current modeling approaches. In E-Z Reader, the word identification system must outperform the oculomotor system to induce word skipping. In SWIFT, there is competition among words to be selected as a saccade target. We conclude that it is the question of competitors in the "game" of word skipping that must be solved in eye movement research.

In computational models based on the concept of sequential attention shifts (SAS), word skipping is a consequence of a competition between lexical processing and saccade programming (target article; cf. Engbert & Kliegl 2001; 2003; Reichle et al. 1998). This mechanism was proposed first by Morrison (1984). Such an explanation of word skipping is *qualitatively* different from the assumption underlying the SWIFT model (Engbert et al. 2002; 2004; Kliegl & Engbert 2003), that a field of lexical activities builds up during the eyes' random walk over the sentence. It is the relative strength of activity that determines the probability of selecting the next saccade target. The related theoretical framework of competition between targets for action is the dynamic field theory of movement preparation (Erlhagen & Schöner 2002). Consequently, the SWIFT model may be generalized as a model for eye-movement control in situations with many potential saccade targets such as visual search or general scene perception. To compare these differences between SAS models and SWIFT, we investigate the mechanism for word skipping using semianalytical techniques.

In E-Z Reader 7, currently the most advanced SAS model, a new saccade program is initiated at the end of stage 1 of the word identification system (Fig. 3 in the target article). Word skipping occurs if the saccade program is canceled by another saccade command during the labile stage. Such a cancellation will occur if the sum of the durations of L_2 (of the currently fixated word) and L_1 (of the skipped word) is smaller than the average duration of the labile saccade program M_1. To calculate the probability of skipping, we have to consider that saccade program stages are gamma-distributed[1] in E-Z Reader. As a consequence, the probability of skipping is given by an integral over the distribution $q^n_\tau(t)$ of durations of the labile saccade stage M_1,

$$p_{\text{E-Z Reader}} = \int_{L_1+\langle L_2\rangle}^{\infty} q^{n}_{\tau}(t)\,\mathrm{d}t \qquad (1)$$

where the time constant τ is related to the mean of the labile saccade program by $\tau = M_1/9$. It is important to note that there are two oculomotor parameters, n and τ, in the probability. The integral in Equation 1 can be evaluated analytically. The probability for skipping a word, which needs an average processing time L_1 of the first stage of word identification, is given by

$$p_{\text{SAS}} = \left(\sum_{k=0}^{n} \frac{1}{k!} \left(\frac{L_1 + \langle L_2\rangle}{\tau} \right)^{k} \right) \exp\left(-\frac{L_1 + \langle L_2\rangle}{\tau} \right) \qquad (2)$$

Since stage L_1 refers to the skipped word, we have to estimate the average processing time during stage 1 by computing means over the five word-frequency classes for L_1. From low to high word frequency (classes 1 to 5) we computed the values 128.0 msec, 100.7

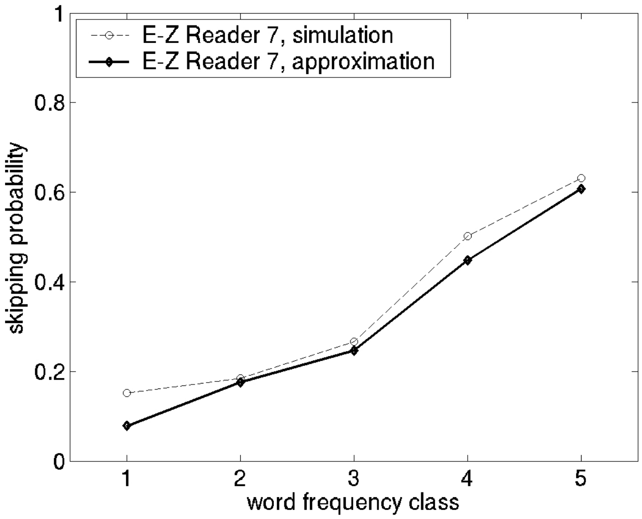

Figure 1 (Engbert & Kliegl). Skipping probability as a function of word frequency class.

msec, 90.8 msec, 60.7 msec, and 44.4 msec for L_1, using Equation 3 and corresponding parameter values given in Reichle et al. The average value of L_2 corresponds to an arbitrary word (the word left of the skipped word). Therefore, we used the ensemble average of L_2 over all words the corpus of sentences,[2] denoted by $\langle L_2 \rangle = 82.3$ msec. For a gamma distribution of order $n=8$ and a mean labile saccade duration $M_1=187$ msec, we obtained $\tau=20.8$ msec. The resulting estimates for the skipping probability $p_{\text{E-Z Reader}}$ are in good agreement with simulated data from the target article (see Fig. 1).

In SWIFT, a field of lexical activities $a_n(t)$ evolves over time. The probability of target selection is given by the relative lexical activity. As a consequence, no additional assumptions must be made to produce forward saccades, refixations, and regressions. The probability of skipping word$_{n+1}$ is given by the probability to select word$_{n+2}$ as the next saccade target, which is computed by the fraction

$$p_{SWIFT} = \left. \frac{a_{n+2}(t)}{\sum_{k=1}^{n+2} a_k(t)} \right|_{t = \text{target selection}} \tag{3}$$

There is no oculomotor contribution to the skipping probability in Eq. (3) – an important difference to Equation (2) for E-Z Reader. Numerical estimates for p_{SWIFT} can be obtained by evaluating the set of lexical activities at the point in time where target selection occurs in SWIFT (for details see Engbert et al. 2002).

Diverging predictions can be derived from SAS and SWIFT models. In E-Z Reader, the probability of word skipping will depend on oculomotor parameters, because of the competition between saccade programming and word identification. In SWIFT, the competition between words for becoming selected as the next saccade target implies a structural stability of word skipping against oculomotor parameters. Therefore, dynamic models generate highly specific predictions, which might be most stimulating for future research: The current controversy on mechanisms of eye-movement control will still be resolvable by experimental results.

ACKNOWLEDGMENTS

We thank Erik D. Reichle for providing us with the simulated data for E-Z Reader 7. This work was supported by Deutsche Forschungsgemeinschaft (grants KL-955/3).

NOTES

1. The gamma distribution for saccade latencies can be written as $q_\tau^n = \frac{1}{\tau - n!}\left(\frac{t}{\tau}\right)^n \exp\left(-\frac{t}{\tau}\right)$, where τ is a time constant and n is the order of the distribution. Mean value and standard deviation are given by $\mu = (n+1)\tau$ and $\sigma = \sqrt{n+1}\tau$. For a relation of standard deviation to mean of one third (Reichle et al. 1998), we have to choose a gamma distribution of order $n = 8$.

2. This procedure may be interpreted as a *mean field approximation*, that is, using the average processing difficulty of the word left to the skipped word. To compute L_1 and $\langle L_2 \rangle$ according to Equation 3 in the target article, we used word frequencies, predictabilities, and the parameters β_1, β_2, and Δ.

Throwing the baby out with the bathwater: Problems in modeling aggregated eye-movement data

Gary Feng

Department of Psychology: Social and Health Sciences, Duke University, Durham, NC 27708. **garyfeng@duke.edu**
http://www.psych-shs.duke.edu/faculty/facultyfeng.html

Abstract: Parameters in E-Z Reader models are estimated on the basis of a simple data set consisting of 30 means. Because of heavy aggregation, the data have a severe problem of multicolinearity and are unable to adequately constrain parameter values. This could give the model more power than the empirical data warrant. Future models should exploit the richness of eye movement data and avoid excessive aggregation.

Because eye movement techniques produce an enormous amount of raw data, it is common practice to perform some sort of data reduction prior to modeling. However, there is a delicate balance between reducing computational complexity and preserving meaningful variance in the data. The E-Z Reader 7 model and its predecessors (Reichle et al. 1998; 1999) posit a comprehensive and elegant set of eye-movement control mechanisms, but the data set used to fit the models is too impoverished to adequately test these models.

The empirical data for E-Z Reader models (from Schilling et al. 1998) are averages of six eye movement variables (single fixation duration, first fixation duration, gaze duration, and the probability of skipping, making single fixations, and making two fixations) over five word frequency levels (Reichle et al. 1998, Table 1; see also Note 6 of the target article). Unfortunately, the structure of the empirical data is ill formed. The six variables are so highly correlated that the data space has far fewer than six independent dimensions, a problem known as *multicolinearity* in linear regression analysis.

Pair-wise correlation coefficients among the six variables range from 0.85 (between skipping rate and first fixation duration) to 0.998 (between first fixation duration and single fixation duration). Furthermore, all eye-movement measures are highly correlated with the logarithm of word frequency. A principal component analysis showed that the first component accounts for 94.6% of the total variance, the first two components account for 98.6%, and the first three components account for 99.999% of total variance. In short, with only 5% loss of information, the six eye-movement variables can be effectively reduced to a single variable, which in turn has an almost perfect linear relationship with log-transformed word frequency.

The consequences of multicolinearity in the dataset are profound. Free parameters (ranging from five in E-Z Reader 1 to at least seven in E-Z Reader 7 models) were effectively estimated on the basis of only five data points, creating a classic identification problem in parameter estimation, where some parameter values may be varied freely without affecting model fit. Moreover, flaws in the data threaten the internal validity of E-Z Reader as an empirical model. If we believe in the principle of parsimony, then the only model that will survive Occam's Razor would be something like "any eye-movement measure is a linear function of the log-transformed word frequency," which is both uninformative and wrong (see Kliegl et al. 1982).

Is the problem of multicolinearity confined to the Schilling et al. (1998) data set? The answer is no. A similar analysis based on our own data (Feng et al. 2003) shows the same pattern of multicolinearity. The problem stems from two sources.

1. Composite eye-movement measures (see Inhoff & Radach 1998), such as gaze duration and probability of skipping, are statistics computed from individual fixations. Because these statistics are calculated on the same sample of fixations, moderate to strong correlations are expected among them. For example, the fixations counted toward single-fixation duration are a subset of first-fixation duration, which is in turn a subset of gaze duration.

2. These correlations are further concentrated as raw data are aggregated to get a "clean" picture suitable for modeling. For example, in our adult reading data, the correlation between first fixation duration and gaze duration is 0.71 when the unit of analysis is per subject per word (N=24,089). It becomes 0.80 when we average across subjects (N=3,599 words), and 0.95 if we only consider five word frequency levels (N=5) and average across both subjects and words.

As long as only a few means of composite eye-movement variables are used for modeling, the problem of multicolinearity will be unavoidable. Therefore, ingenious and intricate theories such as E-Z Reader will remain untestable. The only solution to this problem is to reinstate the richness of the eye-movement data for modeling. There are at least three approaches to this end:

1. Use less aggregated data.
2. Model distribution functions of eye movement variables (e.g., Feng et al. 2001; 2003; McConkie & Dyre 2000).
3. Use raw data instead of composite eye-movement measures (Feng 2001).

In addition to the multicolinearity problem, there are several important flaws in the parameter estimation procedure shared by all E-Z Reader models (see Reichle et al. 1998, p. 157). Instead of normalizing the difference between model predictions and observed values, the authors erroneously squared the difference. Consequently, fixation duration variables, which have a much larger scale than do probability variables, contributed approximately 100 times as much to the index of model fit as did the probability variables (estimation based on Reichle et al. 1998, Table 1). Another error is the use of the standard deviation in the normalization. Because the comparisons were between observed means and simulated means, the sample standard error should be used in the denominator (see Hayes 1988). As a result, the goodness-of-fit index was shrunk by a factor of the square root of *N*. Finally, it is disappointing that there was no attempt to test the fit of each model statistically, or statistically compare successive models. Further analyses on the impact of these factors can be found in Feng (2001).

It may seem paradoxical that even though it has serious problems in parameter estimation, E-Z Reader 7 is successful in simulating many well-known reading eye-movement phenomena. A possible explanation is that precisely because the impoverished empirical data could not provide adequate constraints over parameter values, the authors gained more freedom in assigning parameter values that maximize simulation performance. This would predict that the model's simulation performance would be hampered if the data contain more information, something that could be empirically tested.

In summary, the problems discussed here – multicolinearity in data and issues with parameter estimation and model testing – are fairly low-level. However, a model is ultimately only as good as the data and algorithms on which it is based. There is not enough evidence to conclude that E-Z Reader is empirically validated. Nonetheless, we should not throw out the model with the statistical problems. These issues are not difficult to fix. I look forward to seeing an E-Z Reader 8 that is on a solid statistical footing. Meanwhile, future modeling work should fully exploit the richness of reading eye movements and be wary of the limitations of aggregated data.

ACKNOWLEDGMENTS
I am grateful to Kevin Miller and Ruth Day for helpful comments on early drafts.

Serial programming for saccades: Does it all add up?

John M. Findlay and Sarah J. White
Centre for Vision and Visual Cognition, Department of Psychology, University of Durham, Durham, DH1 3LE, United Kingdom.
j.m.findlay@durham.ac.uk **s.j.white@durham.ac.uk**
http://psychology.dur.ac.uk/staff/details.asp?name=FINDLAY
http://www.dur.ac.uk/s.j.white/

Abstract: This commentary analyses the quantitative parameters of Reichle et al.'s model, using estimates when explicit information is not provided. The analysis highlights certain features that appear to be necessary to make the model work and ends by noting a possible problem concerning the variability associated with oculomotor programming.

Reichle et al.'s target article presents a model of eye control during reading that is impressive in a number of ways. It is fully explicit, quantitative, and economical, and it brings in known features of the visual system (differential magnification) and oculomotor system. It provides a good account of a number of observed phenomena and a quantitative fit to data. Its appearance in *BBS* is particularly welcome because if it proves robust against criticism, this must be regarded as a triumph not only for the model itself but also for the serial stage approach to modeling that underpins it.

The duration of fixations are modeled on the basis of a signal traveling through a number of stages that are strictly serial with the visual, lexical, and oculomotor processes taking place sequentially. These stages are shown in Figures 3 and 13 of the target article; and the latter figure in particular suggests that the time-consuming processes leading to saccades are conceived as the time for signals to traverse brain regions. This represents a different tradition and philosophy to the approach of Findlay and Walker (1999), where the emphasis was on specific time-consuming processes of competitive inhibition, particularly in the late oculomotor stages. Some common ground might be found in the separation of the programming of saccade amplitude from the remainder of the programming. This occurs through the direct (dashed line of Fig. 3) pathway from the early stage of visual processing bypassing the word identification system. Section 3.1.3 indicates that this pathway provides the low spatial frequency information needed to program a saccade. However, it would appear that there also needs to be a modulatory influence from the word segmentation process on this pathway, since the whole basis of the model is that saccades are programmed to words.

The remainder of this commentary works through the model in detail, following the commentators' understanding and looking particularly at the time course of events.

The seriality has the consequence that the duration of a fixation can be expressed as a sum of contributions from the component stages

$$\text{FXDUR} = t(V)+L_1+M_1+M_2-(OV_V+OV_L+OV_M) \qquad (1)$$

where $t(V)$, L_1, M_1, and M_2 are defined as in the model. OV_V, OV_L, and OV_M are introduced to denote the modifications that occur when the model is working dynamically, since overlap (OV) processes can occur. OV_V and OV_L are savings in the visual and lexical stages, respectively, that can come from peripheral preview. OV_M reflects changes in oculomotor preparation time when saccadic programming stages overlap. All the components are described clearly in the target article and, although the detailed magnitudes can be made only with precise knowledge of the text being read, it is possible to make estimations of the distributions. The terms in Equation 1 (above) are all stochastic variables and therefore the grand mean of FXDUR must equate to the combination of the means of the various contributing components. How does this work out? Our analysis below ignores refixations but otherwise tries to follow through Reichle et al.'s model.

Section 3.1.1 discusses the variable $t(V)$, which takes values upwards from 90 msec. L_1 is defined in Equation 2 of the target article as a product of two factors. The first ranges from 110 to 228 msec dependent on word frequency, and the multiplier ranges from 0.5 to 1.0 dependent on word predictability. A plausible overall mean value might be 130 msec. M_1 and M_2 are clearly set out to have mean values of 187 msec and 53 msec respectively.

OV_V, OV_L, and OV_M are not defined explicitly and depend on what happens when the model runs. OV_V and OV_L represent savings on the visual and lexical stages through peripheral preview advantage. OV_M represents modifications when saccadic programming stages overlap. Fixation durations are *shortened* when the planning for a saccade is able to take advantage of preparation already made (as with the second fixations in 5D and 5E of Fig. 5). Fixation durations may also be *lengthened* when saccade skipping necessitates a reprogramming of the location-distance stage, as in the first fixation of 5C.

The sum of the means of the first four terms of Equation 1 above is 460 msec. Therefore, to obtain a plausible overall mean, it seems necessary for the *OV* components to be quite substantial. OV_M can, as far as we can see, only be positive when two conditions are satisfied. First, peripheral preview has allowed completion of the $t(V)$ and L_1 stages of word_{n+1}. Second, the triggering signal falls in the 53 msec non-labile stage of the previous saccade preparation or during the saccadic movement itself (25 msec). Therefore, OV_M cannot exceed 78 msec. Whenever this combination of circumstances occurs, OV_L must equal the full value of L_1 (50 msec–228 msec). This suggests that the OV_M component will usually be smaller than OV_L. Our estimates of plausible parameters are as follows: $t(V)$ 90 msec, L_1 130 msec, M_1 187 msec, M_2 53 msec, OV_V 90 msec, OV_L 60 msec, OV_M 30 msec, summing to a mean FXDUR of 280 msec. Of this figure, 70 msec is "visual-lexical" and 210 msec "oculomotor." This reasoning assigns a very considerable role to peripheral preview, and two predictions seem to follow. If preview is prevented, fixations should be considerably lengthened; consequently, we find the 26 msec preview benefit figure given in section 3.2 surprisingly small. Second, the very first fixation on a text should be substantially longer than subsequent ones.

A similar exercise can be carried out with the variance of FXDUR, which again must be predictable from the variances of the component distributions, taking into account any nonindependence of the terms. How does the variance divide among the various components of the sum, and in particular between the visual-lexical and the oculomotor components? The calculations above suggest that the oculomotor components contribute about 75% to the mean. Unfortunately, the variance of the gamma-distributions from which M_1 and M_2 are drawn are not given in the target article (we very much hope the authors will supply these in their response). However, our rough estimates suggest the oculomotor components must contribute a considerable amount.

If indeed this is the case, it must be reconciled with the fact that in studies of saccades in simple situations, distributions with standard deviations in the 25–30 msec band are often found (Carpenter & Williams 1995; Walker et al. 1995; Wenban-Smith & Findlay 1991). It is, of course, possible that oculomotor variability depends on the circumstances in which the system is used and is higher in reading than in the cases cited. However, it could also be that the serial assumptions of the model are the source of the problem.

Frontal lobe functions in reading: Evidence from dyslexic children performing nonreading saccade tasks

Burkhart Fischer

Brain Research Unit, University of Freiburg, D-79104 Freiburg, Germany.
bfischer@uni-freiburg.de www.brain.uni-freiburg.de/fischer

Abstract: Reichle et al. show that saccades in reading are controlled by linguistic processing. The authors' Figure 13 shows the parietal and frontal eye fields as parts of a neural implementation. This commentary presents data from dyslexics performing nonreading saccade tasks. The dyslexics exhibit deficits in antisaccade control. Improvement of the deficits is achieved in 85% of the cases and results in advantages in learning how to read.

From many different pieces of converging experimental evidence (Fischer 1987) the main components of saccade control have been identified as: (i) fixation, which stabilizes the direction of gaze; (ii) an optomotor reflex, seen under certain conditions as express saccade, when fixation/attention is disengaged; and (iii) a voluntary component, challenged by the instruction to generate antisaccades, that is, saccades in the direction opposite to a visual stimulus (Hallett 1978). Fixation is supported mainly by parietal (Motter & Mountcastle 1980; Robinson et al. 1978) and tectal functions (Munoz & Wurtz 1992), and the reflexes are mediated by the superior colliculus (Schiller et al. 1987; Sommer & Schiller 1992). The voluntary component relies on frontal lobe functions, because successful performance of the antisaccade task is impaired in patients with unilateral frontal lobe lesions (Guitton et al.1985).

Figure 1 shows the basic optomotor cycle consisting of series of periods of fixation (Stop) and saccades (Go). The cycling must not work on its own. It must be controlled by voluntary and/or cognitive processes that make each saccade a meaningful event within the process of active vision. Neurons in the frontal eye fields are activated before purposive saccades – not before any saccade (Bruce & Goldberg 1985).

How can one get more inside, into the relationships between the cognitive processing and the neural systems for saccade control? One possibility is to look at saccade control in nonreading tasks and to compare the corresponding data obtained from subjects who read normally with those of subjects who have reading problems; for example, dyslexics.

Deficits in the acquisition of reading skills may be (and have been) attributed to deficits of a number of different subfunctions within the reading process. One possibility is a deficit in saccade

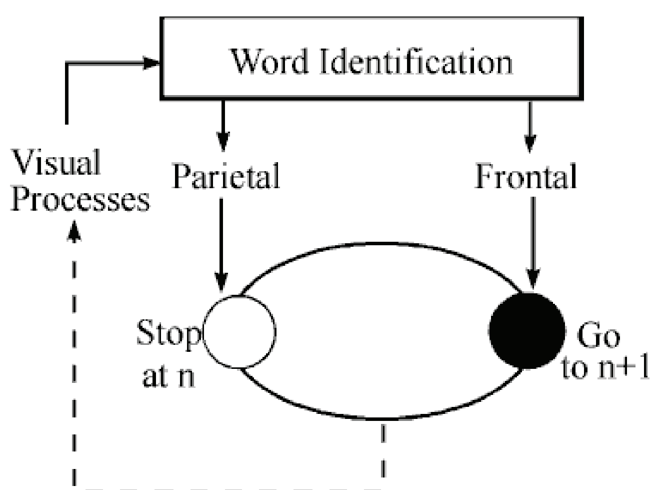

Figure 1 (Fischer). Schematic drawing of the optomotor Stop-and-Go cycle and its control by parietal (Stop) and frontal (Go) functions.

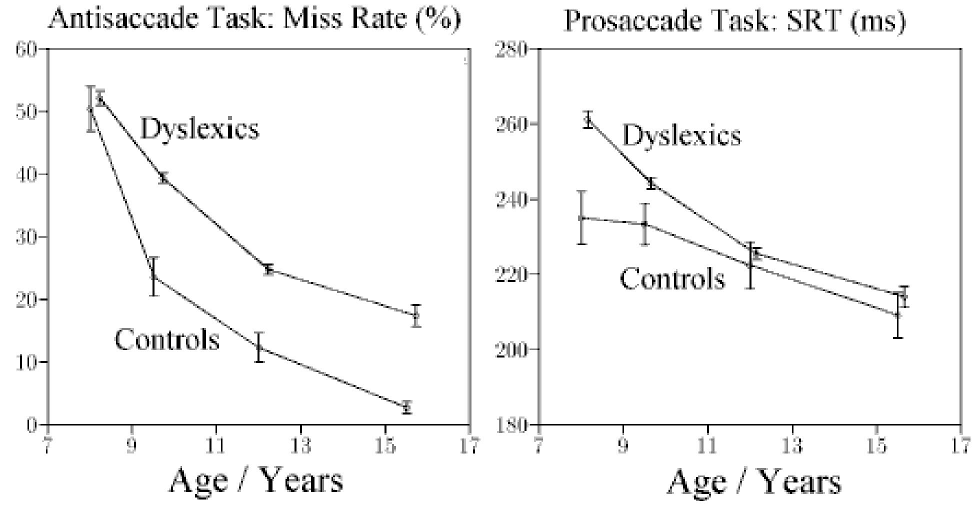

Figure 2 (Fischer). Pairs of age curves obtained from dyslexics (N=1,849) and controls (N=117). Vertical bars represent the confidence intervals.

control. Controversial results have been accomplished with interpretations ranging from "eye movements hold the key of dyslexia" (Pavlidis 1981) to "saccade control is normal in dyslexia" (Olson et al. 1983). Figure 1 reconsiders the problem and makes the following predictions:

1. If only linguistic processing (e.g., word identification) is impaired, saccade control will be affected during reading, but not during nonreading tasks.
2. If the frontal system is impaired, saccade control will be affected during reading and also during those nonreading saccade tasks, which challenge the frontal control component. However, other saccade tasks, for example prosaccade tasks, may be performed normally.
3. If fixation is unstable, reading may become more or less difficult even with intact linguistic processing.

Case number 2 is supported by experimental evidence: Large proportions of dyslexics have a specific problem with the voluntary saccade component, while only a minority exhibit deficits also in prosaccade generation during nonreading tasks (Biscaldi et al. 2000). A preponderance of intrusive saccades (Fischer & Hartnegg 2000b) and/or binocular instability (Stein & Fowler 1993) has been reported in dyslexia as well (case number 3), but will not be discussed here in detail.

Children between the age of 7 and 17 years were categorized as dyslexic or control using the diagnostic tests described earlier (Biscaldi et al. 2000). Two nonreading saccade tasks were administered: a prosaccade task with overlap conditions and an antisaccade task with gap (200 msec) conditions. From the prosaccade task we determined the reaction time, from the antisaccade task we measured the percent number of errors and the percent number of corrective saccades. The methods are described elsewhere (Fischer et al. 1997b).

The normal development of the different components of saccade control from the age of 7 to the age of 85 years has been assessed earlier (Fischer et al. 1997a; Klein et al. 2000). A comparison with the data of dyslexic subjects was described (Biscaldi et al. 2000). Here we present an updated analysis of the data.

Figure 2 shows a pair of age curves of the reaction times of prosaccades (overlap condition) and another pair of age curves of the percentage of those error trials, in which the errors were not corrected. These trials are called *misses.*

The curves show that the initiation of prosaccades is affected only for the youngest group. However, the generation of antisaccades exhibits systematic deficits increasing with age. Counting the percent number of dyslexics, who performed the antisaccade task with miss rates above the mean of the controls plus one standard deviation, reveals that the percentage of affected dyslexics increases with age from about 30% to 55%.

In these dyslexics the reading problem is caused partly by an insufficient frontal control of saccade generation, not by a general impairment of saccade control.

Earlier experiments have shown that daily practice can change saccade control (Fischer & Ramsperger 1986). Three visual tasks were designed for training: One requires fixation, one prosaccades, and another antisaccades (Fischer & Hartnegg 2000a; Fischer et al. 2000). The miss rates of eye movements of 148 dyslexic subjects were measured before and after the training (Fig. 3). About 85% of the subjects improved their antisaccade performance. The training improved only those aspects of saccade con-

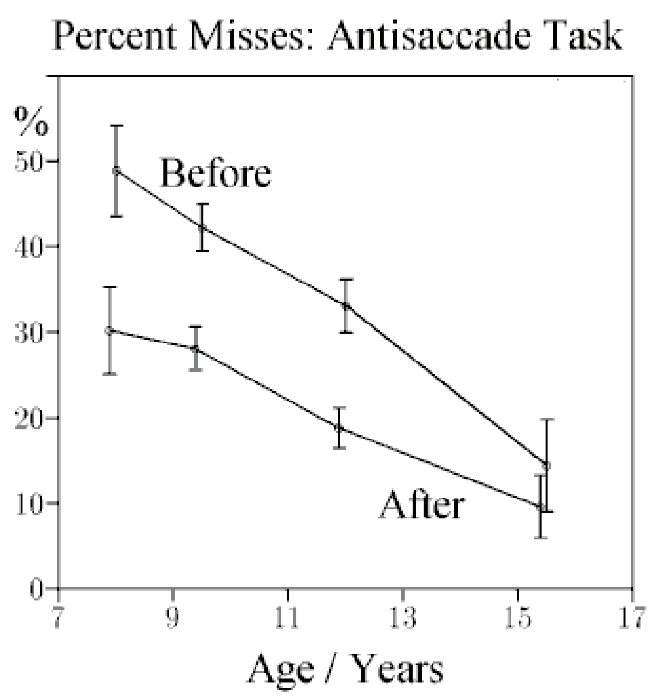

Figure 3 (Fischer). Percent misses in the antisaccade task before and after the training (N=148).

trol that were part of the training program (Fischer & Hartnegg 2000a).

Finally, a group of 20 dyslexics with deficits in saccade control was divided into a test and a control group. Only the test group was given the antisaccade training. Then both groups were recombined and received six weeks of reading instruction. The test group reduced their reading error rate by about 49%, the control group by only 19% (p=0.01). The improvement of saccade control facilitates the learning process but does not replace it.

Among the executive functions of the frontal lobe is the execution of saccadic eye movements during reading. An impairment of this function does not imply that reading is completely impossible, only that the chances of reading errors due to inappropriate saccades are increased. It is suggested that a neural implementation of the E-Z Reader model does indeed include the frontal lobe, and that the model could also serve as a model of dyslexia.

Dimensionality and explanatory power of reading models

Douglas Hanes and Gin McCollum

Neurotology Research, Legacy Clinical Research and Technology Center, Portland, OR 97232. **dough@lhs.org** **mccollum@ohsu.edu**

Abstract: The authors' review of alternative models for reading is of great value in identifying issues and progress in the field. More emphasis should be given to distinguishing between models that offer an explanation for behavior and those that merely simulate experimental data. An analysis of a model's discrete structure can allow for comparisons of models based upon their inherent dimensionality and explanatory power.

The authors are to be congratulated for their structural analysis and comparison of several widely varying models of eye-movement control during reading. The emphasis on the basic structures and assumptions of the models is welcome, as is the authors' recognition that the ability to simulate experimental results is not the only measure of a model.

This commentary presents a further evaluation of reading models in terms of their discrete structures and dimensionalities, which represent the inherent expressive power of the model. It is seldom necessary to consider the output of computer simulations, using specific formulae, to understand the range of phenomena that can be predicted. Moreover, a focus on structure distinguishes the facets of a model that offer explanatory relationships from those that only quantitatively simulate data.

We begin with an example of such a discrete, structural explanation with regard to the spillover effect, then consider two notions of the dimensionality for reading models.

Two lexical stages accommodate the spillover effect. The ability of the E-Z Reader model to correctly predict the "spillover effect," under which the difficulty of one word can lengthen the fixation on the next, has nothing to do with specific formulae or simulations but is inherent in the separation of lexical processing into two stages. In Morrison's earlier model (Morrison 1984), of which E-Z Reader is an elaboration, lexical processing is considered as a single unit, and the signal to generate a saccade originates only after this process is complete. This basic structural assumption implies that the difficulty of processing the current word can have no influence on the following fixation. Therefore, Morrison's model cannot possibly account for the spillover effect.

In contrast, the E-Z Reader model has a second stage of lexical processing following the signal to generate a saccade. This automatically gives the possibility of a spillover effect of various amplitudes, because the duration of this second stage can contribute to the *following* fixation. Therefore, the ability of E-Z Reader and the inability of Morrison's model to account for the spillover effect does not in any way depend upon the specific equations but is implicit in the models' structures.

If the model is fundamentally correct, the requirement of two degrees of freedom in the lexical processing system must be reflected in physiological processes. In this way, the dimensionality of a model of sensorimotor function has implications for physiological organization.

Dimension as the measure of the space determining average fixations or individual fixations. There are various ways of assessing the dimensionality of a model, depending on the facets of interest. Here we give two examples of ways to make dimensional assessments based on a model's ability to predict fixation durations.

One notion of dimensionality is the average number of input variables determining fixation durations in a sequence. In Morrison's model, for example, the sequence of fixation durations is determined by the sequence of lexical processing times for each word. This is an average of one variable per fixation, giving the model a dimension of one. In the E-Z Reader model, the sequence of fixation durations depends on the durations of both stages of lexical processing for each word, as well as the word lengths (which determine early processing rates). This is an average of three variables per fixation, so E-Z Reader has dimension three.

Alternatively, dimension can be determined as the potential number of variables affecting the duration of an individual fixation. In Morrison's model, the length of the fixation is determined by either the duration of lexical processing on the fixated word $L(0)$ or by this duration plus that of lexical processing on the next word $L(+1)$, in case the next word is skipped. A graph of the possible contributions of these variables to the duration of a single fixation is given in Figure 1A. This is a two-dimensional subset of real two-space, giving Morrison's model a dimension of two. A similar analysis shows that with the E-Z Reader model (excluding early processing), individual fixations are determined from the durations of L_2 on the preceding and fixated words and L_1 on the fixated and following words. A graph of possible contributions of $L_1(0)$, $L_2(0)$, and $L_1(+1)$ to a fixation duration is shown in Figure 1B. This three-dimensional graph gives E-Z Reader a dimension of four when the possible contribution of L_2 on the previous word is included.

Both estimates of dimensionality show E-Z Reader to be more complex than Morrison's model, as expected, but they do give different numbers. The reason is that, because of parallel processing of saccades, an individual fixation can involve more cognitive processes and more free variables than does the average fixation. Note that if some fixations have more than average freedom of determination, then others necessarily have less! This is reflected in Figure 1 by the lower-dimensional components of the graphs. The two-dimensional measures are not incompatible but emphasize different aspects of the models.

Conclusion. The essential complexity and expressive power of a model can be represented in a discrete, schematic way. A correctly designed discrete model indicates all of the variables influencing the system, and all the ways in which values of one parameter can constrain those of another. Given the discrete model, simulations can be generated by constructing formulae that provide the best "fit" to the data. However, a focus on simulations can obscure the fundamental properties of the model by presenting results in a form similar to experimental data.

We feel that it is of great importance to distinguish those parts of a model that offer an explanation for behavior and physiology from those parts that merely simulate data. E-Z Reader offers real explanations for how the brain controls saccade timing, while its handling of saccade lengths and refixations is explicitly constructed for purposes of quantitative fit. Other models show their strengths in other areas, as can be seen from the excellent analysis of the target article. Our understanding of reading would be best served by attention to the dimensionalities necessary to explain observed sensorimotor behaviors and their implications for physiological processes.

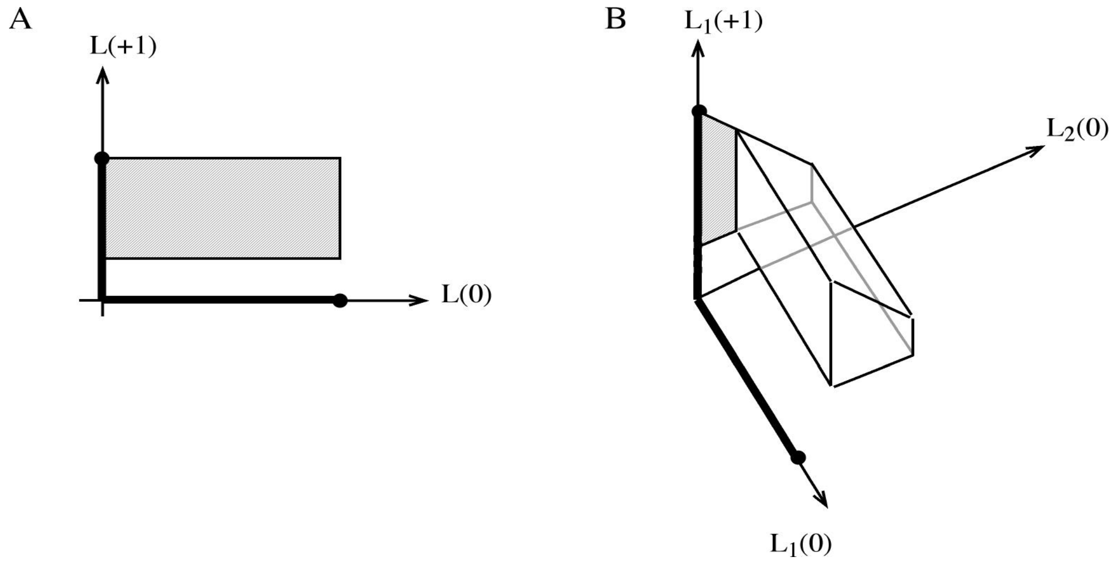

Figure 1 (Hanes & McCollum). A. Values of the pair (L(0), L(+1)) that can determine a fixation duration in Morrison's model. B. Values of the vector ($L_1(0)$, $L_2(0)$, $L_1(+1)$) that can determine a fixation duration in the E-Z Reader model. L(0) and L(+1), lexical processing time on fixated and following words, respectively; $L_1(0)$ and $L_1(+1)$, durations of first stage of lexical processing on fixated and following words, respectively; $L_2(0)$, duration of second stage of lexical processing on fixated word. Enclosed three-dimensional regions, shaded planar regions, and thick line segments are included in the graph.

Visual word recognition and oculomotor control in reading

Lynn Huestegge[a], Jonathan Grainger[b], and Ralph Radach[a]

[a]*Institute of Psychology, Technical University of Aachen, 52056 Aachen, Germany;* [b]*Université d'Aix-Marseille, LPC/CNRS, 13 621 Aix en Provence, France.* **lynn.huestegge@post.rwth-aachen.de grainger@up.univ-mrs.fr ralph@psych.rwth-aachen.de**

Abstract: A central component in the E-Z Reader model is a two-stage word processing mechanism made responsible for both the triggering of eye movements and sequential shifts of attention. We point to problems with both the verbal description of this mechanism and its computational implementation in the simulation. As an alternative, we consider the use of a connectionist processing module in combination with a more indirect form of cognitive eye-movement control.

The E-Z Reader suite of computational models is characterized by the role played by lexical processing in the complex system that guides the eyes through text during reading. The word identification module of E-Z Reader 7 is seen as the "engine" that drives the whole process, determining both the dynamics of saccade generation and the assumed sequential shifting of attention. Although the section on word identification occupies only a modest proportion of the target article, there is no denying that this is a central component of the model.

As argued by Grainger (2003), the time is ripe for a fruitful interaction between research on visual word recognition and research on eye movements in reading. Including a word identification component in a model of eye-movement control in reading is already a significant step in the right direction, and Reichle et al. are to be congratulated for their pioneering work towards this goal. However, there are several ways to go about generating such an interaction. One is to examine how a given model of visual word recognition, motivated by research using isolated word presentation techniques, could be integrated into a more global reading system that includes oculomotor control. Another way is to define a minimalist model of visual word recognition that, when coupled with a model of oculomotor control, optimally fits the data collected using eye-movement paradigms. Reichle et al. have adopted the latter approach, with some unfortunate consequences.

E-Z Reader 7 implements a two-stage approach to word identification. An early stage (L_1) is assumed to play a crucial role in the identification of the orthographic form of the word, whereas a second stage (L_2) is rather related to phonological and semantic processing. Completion of the first stage initiates the preparation of a saccade to the next word, and completion of the second stage initiates an attention shift to the next word. However, an examination of the verbal description of this part of the model and its mathematical implementation reveals a number of problems with this approach.

As the authors note, the major motivation for the two stages in the word identification component of E-Z Reader 7 is the decoupling of eye movements and attention in the model, and the fact that this allows E-Z Reader to capture a wide range of eye-movement data. In the mathematical implementation of E-Z Reader 7, it is apparent that only two empirically observed variables are used to model the various types of processing associated with each stage: word frequency and predictability (see equations for $t(L_1)$ and $t(L_2)$ below). This is because the model was designed to capture the influence of word frequency and predictability on the various measures obtained in eye movement recordings.

$$t(L_1) = [\beta_1 - \beta_2 \ln(\text{frequency})]\ (1\text{-}\theta\ \text{predictability})$$

$$t(L_2) = \Delta[\beta_1 - \beta_2 \ln(\text{frequency})]\ (1\text{-predictability})$$

Two problems come immediately to mind. First, the only difference between these two processing stages concerns the relative weight assigned to frequency compared to predictability. In the first stage, the influence of predictability is reduced and that of word frequency enhanced, compared with the second stage. Second, the relationship between these variables and the orthographic, phonological, and semantic processing described in the verbal model, is left unspecified.

The verbal model is said to be partly motivated by the activation-verification model (AVM) of Paap et al. (1982). It is true that the AVM can be described as a two-stage (activation and verification) model of visual word recognition, but the analogy between the AVM and the word identification component of E-Z Reader stops here. Worse, the precise structure of the AVM was generated to account for experiments showing no frequency effects on early orthographic processing in perceptual identification paradigms. E-Z Reader does just the opposite, by pronouncing frequency effects in the first stage of lexical access. The same is true for predictability. In the AVM, predictability as well as word frequency are thought to influence the second stage of processing exclusively. This is in contradiction to the implementation in E-Z Reader, where effects of predictability on stage L_1 are necessary to account for both word skipping and reduced fixation times for predictable words, since the completion of this early stage triggers eye-movement behavior. While it seems reasonable to assume that predictability may reflect top-down influences from higher-level sentential or text-level representations in the later phases of word identification, the authors need to provide independent evidence for such influences on early orthographic processing.

Furthermore, the option to put orthographic processing in one stage and phonological and semantic processing in the other stage appears to be totally arbitrary. Why not have three separate stages, or put orthography and phonology together (as representations of form), separate from semantics? Indeed, the authors' own presentation of parafoveal preview effects would motivate a regrouping of orthographic and phonological processing. The exclusion of phonology from early word processing is clearly not in harmony with the results obtained by the same group of authors using their fast priming paradigm (see Pollatsek et al. [2000] for a recent review of these issues).

In sum, it appears that there is a considerable gap between the verbal description of the model and its actual mathematical implementation – and no clear theoretical justification for the particular option adopted in either the verbal or the mathematical version of the model.

As mentioned by the authors in their detailed discussion of alternative models, a different approach to the integration of linguistic processing and eye-movement control has been taken by Reilly and Radach (2003). They used a letter- and word-processing module that implements a well established type of interactive activation model as developed in research on single word recognition (Grainger & Jacobs 1996; 1998). The results of word processing are continuously fed back to a spatial salience vector that serves as an arena to integrate visual and linguistic processing in the selection of words as saccade targets. The trigger for the execution of an eye movement comes from an independent fixate center that is codetermined by the level of activity in the word processing module (see Engbert et al. [2002] for a similar approach). In this architecture, there is no distinct processing event that triggers saccade programming, and no shifting of attention. Hence, there is also no need to divide word processing into stages, and some of the problems originating from this division can be avoided. However, although the Glenmore model appears to represent a promising theoretical alternative, it has still to be tested in simulations with a realistic corpus of reading data. It remains to be seen whether it can then match the impressive performance of E-Z Reader in accounting for a wide range of eye-movement phenomena in reading.

Future challenges to E-Z Reader: Effects of OVP and morphology on processing long and short compounds

Jukka Hyönä and Raymond Bertram

Department of Psychology, University of Turku, Turku, FIN-20014, Finland.
hyona@utu.fi rayber@utu.fi http://users.utu.fi/hyona/

Abstract: We argue that although E-Z Reader does a good job in simulating many basic facts related to readers' eye movements, two phenomena appear to pose a challenge to the model. The first has to do with word length mediating the way compound words are identified; the second concerns the effects of initial fixation position in a word on eye behavior.

As Reichle et al.'s target article convincingly demonstrates, the research on eye guidance has reached a stage where the accumulation of empirical data has paved the way for attempts to model and simulate the basic findings related to eye movements in reading. Without any doubt, E-Z Reader has been the most influential model in this respect.

For the information-processing models, it is desirable that they (1) are transparent; (2) are psychologically and neurally plausible; (3) account for the basic empirical facts; and (4) make novel, testable predictions. On the one hand, we think E-Z Reader fares well with respect to the first three requirements (with some limitations, mentioned below). On the other hand, although E-Z Reader may also be capable of producing novel empirical predictions, we are not told what these might potentially be (with one exception concerning a previous version of the model). We are left wondering whether the model is indeed restricted to predicting only the effects that it is designed to simulate, or whether the authors have not yet fully exploited its capacity to generate novel predictions.

In what follows, we take up two empirical phenomena that pose a challenge to E-Z Reader, namely, the role of morphology in the processing of compound words and the effects of fixation location in a word on eye behavior.

Many languages (e.g., German, Dutch, and Finnish) depart significantly from English in having highly frequent compounding. For example, in Finnish, more than 50% of all existing words are compounds. If we are to understand the basic reading processes in these languages, we need to acquire a good insight into how compounds are processed. In a recent study, we (Bertram & Hyönä 2003) demonstrated that word length mediates this process: For relatively long compounds (12–14 characters) the recognition process starts with lexical access of the first constituent and not of the whole word (see also Hyönä & Pollatsek 1998; Pollatsek et al. 2000), whereas the opposite is true for the relatively short compounds (7–8 characters). The potential challenge these findings pose to E-Z Reader is that word length appears to determine whether word or constituent frequency affects the initial fixation on a word.

An initial attempt to model compound word processing with E-Z Reader demonstrates that the job is not trivial (Pollatsek et al. 2003). Pollatsek et al. showed that a version that fitted the data best on reading long compounds was the one in which word identification was assumed to appear serially via the constituents (i.e., accessing first the initial constituent, then the second one, then gluing the two together). However, such a model runs into problems in accommodating our finding (Bertram & Hyönä 2003) that short compounds are recognized via the whole word form and that the first constituent does not seem to get activated. What might be needed are letter-level representations feeding activation to word-level nodes, where letter-level activation varies as a function of eccentricity from the center of fixation point (cf. the Glenmore model of Reilly & Radach [2003]). In the case of short compounds this is not enough, but there should also be a mechanism that gives priority to word-level nodes over compound word constituents. Without such a mechanism, there would be faster and more pronounced activation of the first constituents, because they are typically of higher frequency than the whole words. In morphological processing models (e.g., Schreuder & Baayen 1995) it is assumed that a morphologically complex word is segmented into its component morphemes before they can be mapped onto their corresponding access representations. The priority of word-level nodes over constituent nodes might be accounted for by incorporating this additional segmentation time into the above-mentioned mechanism. We have unpublished data suggesting that segmentation might indeed not be a straightforward operation, especially when there is a lack of clear segmentation cues (Bertram et al. 2003). The apparent reason why the first constituent of long compounds is more strongly activated than the whole-word form (despite the potential segmentation problems) is due to low letter-level activation for the end letters (attributable to eccentricity).

The second empirical phenomenon that appears to pose a challenge to E-Z Reader is, on the one hand, the relationship between within-word fixation location and refixation probability, and, on the other hand, the relationship between fixation locations and fixation durations. Regarding the former relationship, it has been shown that the location of initial fixation has a marked influence on the probability of refixating a word (and subsequently on gaze duration; e.g., Rayner et al. 1996). The farther away the initial fixation lands from the optimal viewing position (OVP), the more probable it is that a refixation occurs. In E-Z Reader, refixation probability varies only as a function of word length (see Equation 7 in the target article). Therefore, we suggest that the posited mechanism for refixations is modified by also including fixation position in the equation.

As regards the second OVP phenomenon, Vitu et al. (2001) found that the closer the fixation is to OVP, the longer its duration. We have unpublished data (Bertram & Hyönä 2002) that replicate the results of Vitu et al.: First fixation is clearly shorter when it lands on the beginning letters in comparison to OVP. We observed this for both short and long compound words; thus, the relationship appears replicable across languages and for different word lengths (however, see Rayner et al. 1996). Although the authors do not acknowledge it, their Equation 1, which quantifies early prelexical visual processing, may be able to simulate the above-mentioned effect. According to our calculations, the time for early visual processing is shortened the further away a fixation is from OVP – a pattern in line with the inverted OVP effect observed in first fixation duration. Therefore, the authors' claim that the equation is "consistent with evidence that word identification is most rapid if the word is fixated near its center" (sect. 3.1.1, para. 3) appears to be at odds with the data and our interpretation of the equation.

ACKNOWLEDGMENT

The authors acknowledge the financial support of Suomen Akatemia (the Academy of Finland).

Selection for fixation and selection for orthographic processing need not coincide

Albrecht W. Inhoff and Kelly Shindler

Department of Psychology, State University of New York at Binghamton, Binghamton, NY 13902-6000. **inhoff@binghamton.edu**

Abstract: The E-Z Reader model assumes that the parafoveal selection for fixation and the subsequent selection for attention allocation encompass the same spatially distinct letter cluster. Recent data suggest, however, that an individual letter sequence is selected for fixation and that more than one letter sequence can be selected for attention allocation (processing).

Reading requires that word recognition and sentence comprehension be coordinated with the programming of saccades, as the eyes need to be near the to-be-identified words. Although an abundance of evidence shows that eye movements are influenced by spatial and linguistic properties of a word at or near fixation, these data often reveal a complex relationship between word recognition and saccade programming. The main goal of the E-Z Reader model, and of other models of oculomotor control during reading, is to explain how local demands of text processing can yield a relatively heterogeneous – but carefully calibrated and well adapted – pattern of eye movements.

The central assumption of the model is that the individual word is the exclusive unit of orthographic and semantic processing at each point in time. Each and every word of text is the target of a programmed (but not necessarily executed) saccade and of a corresponding shift of attention that controls a word's orthographic and semantic processing. The pattern of executed saccades can differ from the strictly sequential pattern of attention shifts, because saccade programming and the shifting of attention are controlled by functionally distinct word recognition processes, and because the programming of an overt saccade is more intricate than the covert shifting of attention.

Earlier data (Inhoff et al. 2000; Starr & Inhoff, in press) challenged a core assumption of the E-Z Reader model, that the programming of a saccade to the next word in the text begins before the word is attended. The current commentary examines another assumption of the E-Z Reader, that the initial selection of a parafoveal word for saccade targeting is followed by its selection for processing so that their spatial areas coincide.

The current version of the E-Z Reader model (i.e., the target article) includes a pre-attentive stage of early (low-spatial frequency) visual processing. It segments visible text into spatially cohesive objects that are separated by interword spaces. A parafoveally visible object that occupies the space of the next word in the text becomes the target of a saccade after the orthographic form of the attended word is specified, and it becomes the subsequent target of a corresponding shift of attention after the meaning of the attended word is accessed. Pre-attentive processing thus discerns a parafoveally visible spatial object, the next word in the text, so that it can become the target of a saccade and of an attention shift.

The results of a recent study (Inhoff et al. 2003), in which we manipulated the spatial segmentation of a parafoveally visible word, are not in harmony with this concept. We created four parafoveal viewing conditions, two with matching and two with mismatching parafoveal word length previews. Mismatches were created by replacing a center or near-center letter of a parafoveally visible five- to eight-letter target word with a blank space that separated it into two spatially separate sequences of letters. For instance, the mismatching preview of the target word *student* consisted of *stu ent.* Matching previews were created by replacing the corresponding center letter with a dissimilar letter, for example, *stusent.*[1] Orthogonal to the manipulation of parafoveal word length, we also manipulated the linguistic informativeness of the preview by creating two conditions, one in which useful orthographic information was available, as shown in the examples, and one in which it was denied (e.g., *vip asp* and *vipsasp*). The intact target (*student*) was visible when the eyes moved to the right of the blank space preceding it, and target viewing durations were measured to determine effects of spatial segmentation and of useful orthographic information. If pre-attentive segmentation was used for the targeting of saccades, then saccades should have been longer in the matched condition than the mismatched condition, as the full target preview could be reached with a larger saccade. Furthermore, if pre-attentive segmentation directed attention, then more useful orthographic information should be obtainable from the matched parafoveal preview, as it reveals more useful orthographic information than the first segment of a mismatched preview.

As can be seen in Table 1, saccades to the target were longer in the matched condition, as readers appeared to direct the eyes to an individual parafoveally visible letter string, the first segment in the mismatched condition and the full letter string in the matched

Table 1. (Inhoff & Shindler) *The size of the saccade to the target (in character spaces) and ensuing target first fixations and gaze durations (in msec) as a function of the length and the orthographic informativeness of the preview*

	Matched length		Mismatched length	
	Letter preview	No preview	Letter preview	No preview
Saccade to word	8.8	9.0	8.4	8.2
First fixation duration	199	217	224	235
Gaze duration	271	307	311	347

condition. First fixations and gaze durations on the target did not yield the expected result, however. Instead, letter previews were equally useful in the matched and mismatched conditions, suggesting that parafoveal selection for processing in the mismatched condition encompassed more than the next parafoveally visible letter sequence.

The spatial segmentation of the parafoveal target preview also did not influence the usefulness of available orthographic information on that subset of trials in which the eyes landed at the previously visible first segment of the mismatched target preview (see Table 2). Readers thus obtained useful orthographic information from both segments of the mismatched parafoveal target preview, irrespective of the selection of a particular segment for subsequent fixation.[2]

Table 2. (Inhoff & Shindler) *The size of the saccade to the target (in character spaces) and ensuing target first fixations and gaze durations (in msec) as a function of the length and the orthographic informativeness of the preview for the subset of 760 trials in which the eyes landed at the first segment of the target preview*

	Matched length		Mismatched length	
	Letter preview	No preview	Letter preview	No preview
Saccade to word	8.3	7.9	8.5	7.5
First fixation duration	193	223	211	233
Gaze duration	291	337	317	369

Other results (Briihl & Inhoff 1995; Inhoff 1989a; 1989b; 1990; Lima & Inhoff 1985), argue against the possibility that acquisition of useful orthographic information from the parafoveal target preview was confined to beginning letters in the matched and mismatched conditions. In Lima and Inhoff (1985), preview of the target's beginning trigram yielded a benefit of approximately 20 msec in the gaze duration data, substantially less than in the current data, and preview of the full word yielded a benefit of approximately 40 msec, quite similar to the current data.

Together, these findings indicate that the spatial areas used for saccade targeting and for the acquisition of useful orthographic information need not coincide. This conception is in harmony with our view (Inhoff et al. 2000; Starr & Inhoff, in press) that more than one spatially distinct linguistic unit can be processed concurrently during a fixation.

NOTES

1. The blank target space in the mismatched length condition was replaced with an uninformative letter in the matched length condition to reveal the identical sequence of useful and non-useful parafoveal letters in both conditions.

2. This does not consider potential inaccuracies due to oculomotor error.

On keeping word order straight

Alan Kennedy

Department of Psychology, University of Dundee, Dundee, DD1 4HN, Scotland, United Kingdom. **a.kennedy@dundee.ac.uk**
www.dundee.ac.uk/psychology/akennedy

Abstract: E-Z Reader is a highly successful model of eye-movement control, employing the notion of a serial-sequential attentional spotlight switched from word to word. Evidence of parallel processing of words in text calls this notion into question. Modifications to the model to accommodate this evidence are possible but will not address the fundamental objection that reading should not be seen as "surrogate listening."

At least until the present version of E-Z Reader, there was a sharp theoretical distinction to be drawn between a model committed to the strict operation of a serial attentional switch and alternatives involving a degree of parallel processing, through the operation of an attentional gradient, for example. The primary testing ground for this theoretical divide was the existence, or otherwise, of parafoveal-on-foveal effects; that is, demonstrations that properties of word$_{n+1}$ (not yet fixated, but possibly attended) influence processing on word$_n$. As noted by Reichle et al., most early evidence on this question was negative (Carpenter & Just 1983; Henderson & Ferreira 1993), but the balance of evidence has now shifted, with a majority of recent studies supporting the proposition (Kennedy 1998; 2000b; Kennedy et al. 2002). The current version of E-Z Reader goes a long way towards accommodating these data by adding an early pre-attentive mechanism, able to survey the terrain of the upcoming text. In particular, orthographic parafoveal-on-foveal effects can now be accounted for without deserting the crucial claim that processing is to be characterised by default as involving a serial attentional switch. The modification to the model will no doubt have the effect of changing the focus of future empirical work from questions about *whether* such effects occur, to questions about their precise locus of operation. For example, sublexical effects exerted from initial letter familiarity or constraint might be characterised as less subversive than apparent lexical or supralexical effects. Nonetheless, it is difficult to avoid the conclusion that further erosion of the "seriality" assumption underlying E-Z Reader must occur. This is, indeed, foreshadowed in the target article by the concession that the SWIFT model of Engbert et al. (2002) "provides a viable alternative" to the E-Z Reader model.

The critical problem with the E-Z Reader model may not, in fact, be its adherence to "contentious" assumptions about the serial allocation of attention. It may rather be the set of tacit assumptions underlying that process and, in particular, the claim that words must be identified in turn because readers "need to keep word order straight" (Pollatsek & Rayner 1999). Although not made completely explicit in the target article, it is clear that this claim relates to the authors' belief that higher-level linguistic processes such as the computation of syntactic relationships can be understood *only* within the framework of a model where word-order can be preserved ("The serial allocation of attention is necessary because it preserves the temporal order of the words" [sect. 5.2.2, para. 5]). Given that reading is essentially a spatial activity, involving the deployment of eye movements over a visual object, the reference to *temporal* order is revealing: It suggests a commitment to the assumption that reading is, in fact, a form of surrogate listening. The clearest statement of this assumption, largely accepted but normally undeclared in the literature, is provided by Ferreira and Clifton (1986): "there is no reason to doubt that the syntactic processing strategies used by listeners and readers would be the same." This assumption has had a profound effect on theories of reading and has played a part in the evolution of models of eye-movement control that assign a significant role to an attentional spotlight, deployed sequentially across "word objects." It is, nonetheless, fallacious. Written text is normally continuously available as a spatially extended object on the page or screen, and, unlike the speech signal, can be inspected at will (Kennedy 2000a). This bestows unique advantages on the reader because, as noted by Reichle et al., re-inspection is typically deployed as a reprocessing option. The cost of incorrect initial structural analysis is far higher when processing speech, where no process equivalent to visual re-inspection is available (Watt & Murray 1996).

It should be noted, however, that the equivalence between *re-inspection* and *reanalysis* is itself in need of careful justification: It can be drawn only because readers compute and retain enough spatially coded information to make saccades to specific locations possible. This point needs to be stressed, because saccades made in the service of reanalysis are often very large and cannot be controlled accurately by physical identification of their correct landing site (Murray & Kennedy 1988). For readers of English, the relationship between initial landing position and launch site for progressive saccades is linear over the whole range of launch positions likely to be found in normal reading, with one character shift in launch position producing a corresponding shift of about a third of a character in landing position. However, the equivalent analysis of inter-word regressive saccades shows a virtually flat relationship between extent and accuracy. Landing position is close to the target word's centre, regardless of launch position, an outcome that strongly implies spatial control over saccades to "already-inspected" sites. The fact that readers know where previously inspected words lie (Radach & McConkie 1998) means that their temporal order of fixation need bear no relationship at all to their spatial succession. Indeed, the fact that the reader, unlike the listener, can afford to adopt a single "preferred" structural analysis rests on the fact that a repair process is readily available, underpinned by this spatially coded information (Kennedy et al. 2003; Kennedy & Murray 1987; Pynte et al. 1988).

The E-Z Reader model takes as a theoretical given the concept of *word object,* an idea which found its place in the literature because it neatly captured the idea of a visually-presented word functioning as a cognitive *event,* bounded by the time spent inspecting it (McConkie 1979). Processing could thus mimic the (temporal) sequence of events that would obtain if the word were heard rather than seen. It is hardly surprising, therefore, that the claim that text presented without inter-word spaces can be read quite well (Epelboim et al. 1993) should have been seen as peculiarly subversive and rejected with vigour (Rayner & Pollatsek 1996; see also Epelboim et al. 1996). What else but inter-word spaces could underpin the notion that the "primitive object level" in reading is the word (McConkie & Zola 1987)? Inter-word spaces are the visual equivalent of auditory segmentation, defining the boundary of attention. They allow the notion of an attentional spotlight, switched to each word in turn, to smuggle into any model of reading that employs it the plausible, but questionable, belief that reading is a form of surrogate listening. From this point of view, properties of a word in the parafovea *cannot* influence current foveal processing, because it is obviously not possible to hear the next word before it is uttered. The irony is, of course, that the serial-sequential nature of speech is precisely the property that a writing system avoids.

ACKNOWLEDGMENT

This paper was prepared with the assistance of Grant No R000223650 from the UK Economic and Social Research Council.

How tight is the link between lexical processing and saccade programs?

Reinhold Kliegl[a] and Ralf Engbert[b]

[a]*Department of Psychology;* [b]*Department of Psychology and Center for Dynamics of Complex Systems, University of Potsdam, Potsdam, 14415, Germany.* **kliegl@rz.uni-potsdam.de**
www.psych.uni-potsdam.de/people/kliegl/
engbert@rz.uni-potsdam.de **www.agnld.uni-potsdam.de/~ralf/**

Abstract: We question the assumption of serial attention shifts and the assumption that saccade programs are initiated or canceled only after stage one of word identification. Evidence: (1) Fixation durations prior to skipped words are not consistently higher compared to those prior to nonskipped words. (2) Attentional modulation of microsaccade rate might occur after early visual processing. Saccades are probably triggered by attentional selection.

Two core assumptions of the E-Z Reader model are: (1) attention is shifted across the text in a strictly serial fashion from word$_n$ to word$_{n+1}$, and (2) eye-movement programs are initiated by completion of the first of two stages of word identification. The SWIFT model does not make these assumptions; its three core assumptions are: (1) eye-movement programs are initiated by an autonomous timer, (2) words within the perceptual span are processed in parallel at rates that depend on the distance from the current fixation, and (3) lexical processing difficulty of the fixated word delays the initiation of a saccade program (Engbert et al. 2002). We comment on implications of E-Z Reader's core assumptions in the light of new experimental results and speculations.

Lexical cancellation of saccade programs. E-Z Reader is strongly committed to the assumption of serial word-to-word shifts of attention. Together with the assumption that saccade programs are triggered by completion of a first stage of lexical processing, this implies longer fixation durations prior to skipped than prior to nonskipped words. The reason is that word skipping requires that a saccade targeting the next word is canceled and reprogrammed to the following word; saccade cancellation necessarily increases the fixation duration. E-Z Reader and our own previous model generate a large effect of this kind (58 msec in E-Z Reader 7; 75 msec in Engbert & Kliegl 2001). The empirical evidence, however, is not consistent with this assumption, with effects ranging from 38 to −26 msec (see Kliegl & Engbert [2003a] and Radach & Heller [2000] for references to six studies/twelve estimates with a median close to zero). We checked this effect in a recent corpus based on 32 young and 33 older adult readers of German sentences (Kliegl et al., in press). In both samples the critical fixations were shorter (−9 to −4 msec) in various analyses (e.g., contrasting saccades from word$_n$ to word$_{n+2}$ with saccades from word$_n$ to word$_{n+1}$; matching words that preceded a skip with the same words [read by different participants] when they preceded a saccade to the next word).[1] Obviously, this experimental variance needs to be explained, but given that E-Z Reader 7 overestimates even the maximum reported increase in fixation duration by 53%, the assumption of lexical-processing triggered saccade-cancellation does not appear to be well supported. Guidance by attentional gradients does not predict a specific increase of fixation durations prior to skipped words.

Lexical initiation of saccade programs. The SWIFT assumption of autonomous saccade generation was motivated by the considerable explanatory power of models assuming primary oculomotor control and by research demonstrating a close link between attentional selection and saccade program initiation (see Deubel et al. [2000] for an excellent review). Recently, covert attention shifts were shown to modulate microsaccade rate and orientation in spatial cueing paradigms (Engbert & Kliegl 2003; see also Hafed & Clark 2002). If attention modulates microsaccade rate and orientation in reading fixations, we can test the E-Z Reader assumption that saccade programs are triggered exclusively by the completion of the first stage of word identification (*L1*), which is

130 to 150 msec after the beginning of a fixation. Alternatively, the completion of early visual processing or attentional selection in the visual system might trigger saccade programs. Indeed, preliminary analyses indicate a linear increase in microsaccade rate from 90 to 200 msec (Kliegl & Engbert 2003), which is after early visual processing but already before the completion of *L1*. We speculate that, aside from effects related to the spatial position, some of these microsaccades represent traces of inhibited saccade programs as postulated in SWIFT, but any reliable link between perceptual or lexical processing and microsaccade rate or orientation will provide important constraints for attentional and ocular control during reading.

Reading as a special case of dynamic attention allocation. Obviously, attention and ocular control did not evolve for reading, but reading is a special application of the attentional/ocular control system. Indeed, the highly constrained spatial nature of the reading process represents an ideal testbed for the further development of theories of attentional/ocular control models. We argue that SWIFT can be ported more readily to nonreading situations (such as visual search) than E-Z Reader, because it does not make any reading-specific assumptions with respect to target selection; indeed a variant of the model was applied to searching for a target in a display of Landoldt rings (Engbert & Trukenbrod 2003). Moreover, the combination of target selection via attentional gradient and parallel processing of words within the perceptual span allows us to generate all types of reading eye movements from the same underlying mechanism. In contrast, in E-Z Reader some reading eye movements require special assumptions (i.e., word skipping or refixations) and others are not even part of the present framework (i.e., interword regressive movements). SWIFT may actually be too flexible, given emerging empirical constraints. For example, it may be necessary to constrain parallel processing within the perceptual span to lexical preprocessing to reduce semantic parafovea-on-fovea effects. Such constraints, however, can be implemented and tested in nested models.

Conclusion. Although E-Z Reader 7 and SWIFT differ in core assumptions, it does not seem insurmountable to introduce flexibility of saccade triggering in E-Z Reader and to constrain parallel processing and possibly autonomous timing in SWIFT. Therefore, E-Z Reader may need to abandon the assumption that all saccades are canceled or triggered by the completion of lexical processing stages; SWIFT may need to restrict parallel processing to visual/lexical preprocessing. Such adjustments, if necessary, will be forced by experimental results. The purpose of a computational model is to provide a coherent perspective on a complex set of empirical results and generate new hypotheses. Computational models of attentional and ocular control of reading already live up to this expectation.

ACKNOWLEDGMENT

This research was supported by Deutsche Forschungsgemeinschaft (KL 955/3).

NOTE

1. We replicated longer fixation durations following a skipped word. Also, skipping saccades started closer to the end of word_n and landed closer to the beginning of word_{n+2} compared with matched movements from word_n to word_{n+1}, as expected from oculomotor control theories.

Psycholinguistic processes affect fixation durations and orthographic information affects fixation locations: Can E-Z Reader cope?

Simon P. Liversedge and Sarah J. White

Centre for Vision and Visual Cognition, Department of Psychology, University of Durham, Durham, DH1 3LE, United Kingdom.

s.p.liversedge@durham.ac.uk s.j.white@durham.ac.uk

http://psychology.dur.ac.uk/staff/details.asp?name=LIVERSEDGE

www.dur.ac.uk/s.j.white/

Abstract: This commentary focuses on two aspects of eye movement behaviour that E-Z Reader 7 currently makes no attempt to explain: the influence of higher order psycholinguistic processes on fixation durations, and orthographic influences on initial and refixation locations on words. From our understanding of the current version of the model, it is not clear how it may be readily modified to account for existing empirical data.

E-Z Reader 7 provides an impressive account of the processes that determine when and where fixations are made during reading. The eye movement patterns that the model predicts are remarkably similar to the observed data. Furthermore, the model is based on quite simple, fundamental principles. In this commentary, we would like to consider two central aspects of the model that we believe may require reconsideration if future versions of the model are to explain data that currently exist in the literature. First, there is evidence to show that processing beyond the level of orthographic identification can influence the duration of fixations. The second issue is that there is growing evidence to suggest that the orthographic characteristics of words can influence where they are first fixated and refixated. It is possible that future versions of the model could account for these additional phenomena and, therefore, our criticisms are intended to be constructive in nature.

Our first point is that Reichle et al. limit their model to explaining lexical and visual influences on fixation duration. In the E-Z Reader model, L_1 is a stage of orthographic identification that is influenced by word frequency and predictability. Completion of this stage of processing is the primary determinant of fixation duration. However, studies have shown that processing beyond orthographic identification does influence initial fixation durations on words (e.g., Murray & Liversedge 1994; Rayner et al. 1983a). To account for these higher-level influences on the duration of fixations whilst retaining the underlying mechanisms of the model, such processes must, it seems to us, modulate the time required to complete L_1. That is, L_1 must be redefined as being processing which includes full lexical access, syntactic processing, and perhaps even thematic and semantic processing.

However, it is not clear whether such depth of processing may be realistically achievable within existing L_1 time constraints. If not, then it may be necessary to extend the L_1 stage of processing, thereby providing sufficient time for higher order processing to occur during this period. Such a modification would result in more plausible timings for the occurrence and influence of higher order cognitive processes on fixation durations. Note, however, that since eye-movement programming can begin only after completion of L_1, this will necessarily reduce the time allocated to program a saccade (M_1 and M_2). As the authors note in section 3.1.4, given existing data (e.g., Rayner et al. 1983b), the mean eye-movement programming time cannot be much shorter than is currently specified in the model. Consequently, such a modification may not be viable. Note also that if this modification were made, it is then unclear what type of processing would occur during L_2 (the stage in which readers currently perform full lexical access and which triggers the attention shift). L_2 is central to the mechanism for decoupling of eye movements and attention, and abandoning this stage would constitute a major change to the model.

An alternative possible modification is to substantially alter the fundamental mechanism for the initiation of eye-movement programming. That is, completion of L_1 would not serve to trigger the initiation of an eye movement. In such a situation, higher-level processing could take place in parallel with the labile stage of saccadic eye movement programming. Cognitive processing could then affect this labile stage at any time to influence when the eyes move. Such an alteration may overcome some of the time constraint problems identified above; however, the nature of L_2 would still have to be respecified. Furthermore, the authors may consider such a modification to be too radical a departure from the existing mechanistic processes by which E-Z Reader 7 currently operates.

The second point that we wish to make about E-Z Reader 7 concerns what determines specifically where words are first fixated. Within E-Z Reader 7, the visual system extracts low spatial frequency information from the visual periphery and the oculomotor system uses this visual information, apparently exclusively, to target saccades. While the authors suggest that word shape information may be provided by the visual system and this in turn could affect saccade targeting, within their simulation, the only information that is used to guide saccade extent is word length. As noted by the authors (Note 5), a number of studies have now shown that the frequency of letter sequences at the beginning of words influences where words are first fixated (see also Radach et al. 2003; White & Liversedge, in press). Furthermore, evidence also suggests that the characteristics of words can influence the direction (White & Liversedge, in press) and length (Hyönä 1995a; White & Liversedge 2003) of refixation saccades. While it may be possible to explain such effects through processing of low spatial frequency word shape information, how such processes would operate is not currently specified. Moreover, studies using artificial tasks (Beauvillain & Doré 1998) and recent results from our laboratory investigating normal reading (White & Liversedge 2003) have shown that orthographic information influences where words are first fixated and refixated for upper case text. Uppercase text does not have visually distinctive differences in word shape to the same degree as lower-case text. Therefore, it is not clear how E-Z Reader 7 could explain such results on the basis of low spatial frequency information alone.

To conclude, E-Z Reader 7 impressively explains a wide range of eye movement behaviour during reading. In its present form, it makes no attempt to explain existing evidence for higher-level influences on fixation durations and growing evidence for orthographic influences on where words are initially fixated and refixated. We believe that these aspects of eye-movement control during reading are important and that an attempt to account for such oculomotor behaviour would strengthen future versions of the E-Z-Reader model.

Basic assumptions concerning eye-movement control during reading

George W. McConkie[a] and Shun-Nan Yang[b]

[a]*Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801;* [b]*Brain Science Institute, Riken, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan.* **gmcconk@uiuc.edu** **syang3@brain.riken.go.jp**

Abstract: Reichle et al. specify two assumptions as being basic to E-Z Reader: Words are sequentially attended during fixations, and saccades are triggered by a cognitive event. We point out that there is little evidence for the first assumption and counterevidence for the second. Also, the labile/nonlabile stage distinction in saccade preparation seems to be contrary to current evidence. An alternative explanation of saccade onset times in reading assumes that saccades are strategically generated, independent of language processing, but are delayed on a probabilistic basis by processing difficulties.

The development of E-Z Reader is a notable intellectual accomplishment. In its detail, scope, and ability to account for many established results, it sets a new standard for what a theory of eye-movement control should accomplish. At the same time, there is reason to doubt some of the basic assumptions that lie at its heart. Here we discuss the two that Reichle et al. list as being the "central assumptions" of the model: (1) words are attended sequentially during the fixation (here called the sequential attention assumption), and (2) the signal for moving the eyes is the occurrence of some stage in word identification (the cognitive saccade triggering assumption).

The sequential attention assumption, as proposed earlier by Morrison (1984), provides a way to employ the findings of "parallel programming" of saccades by Becker and Jürgens (1979) to drive the oculomotor engine. Becker and Jürgens observed that if a saccade target is displaced at critical times following its onset, the resulting saccade is modified in its amplitude or direction. By assuming that the critical event in those studies was a shift of attention to the new target location, Morrison had only to further assume that covert attention shifts occur from word to word during fixations in reading, to then invoke this mechanism as a model for driving the eyes during reading: Shifting attention occurs when the system is ready to consider the next word; this shift triggers preparation for a saccade; and if another shift occurs soon enough, it cancels or modifies the current saccade program.

Actually modeling this mechanism indicates that timing constraints within Morrison's model are contradictory, which has forced modifications in the more recent implementation of E-Z Reader; for example, a saccade currently is assumed to be initiated by some cognitive event that occurs before attention shifts. However, there is very little empirical support for the proposed sequential attention assumption during reading, in spite of direct attempts to experimentally reveal it (Blanchard et al. 1984). It is quite possible that Becker and Jürgens obtained their results only because of the stimulus displacement, and that no such discrete attention shifts occur during the reading of stable text. Reichle et al.'s model also postulates two distinctive stages of saccade programming, labile and nonlabile, that are used in accounting for the occurrence of word skipping. However, in parallel saccade programming, there is actually a dynamic modification of the amplitude of the initial saccade toward the second target position, indicating a summing of signals from both steps. Even after a saccade is in flight, it can still be modified by a later-programmed saccade, thus contradicting Reichle et al.'s distinction as they describe it.

Evidence is clearer with regard to the cognitive saccade triggering assumption. We (McConkie & Yang 2003; Yang & McConkie 2001; in press) have conducted a series of experiments in which the text is replaced by alternative stimulus patterns (strings of Xs, random letters, patterns with spaces filled, etc.) for occasional single fixations as people read, to observe the effect on the saccade onset time. The stimulus changes occur during saccades so that the stimulus motion that normally signals such changes is not perceived and does not produce the type of saccadic inhibition observed with changes during fixations (Reingold & Stampe 1999). A stimulus pattern consisting of wordlike random letter strings would be expected to produce characteristic changes in saccade initiation times (fixation durations) in an E-Z Reader mechanism or any other mechanism in which the saccade onset is triggered by some cognitive event at a level higher than simply achieving a clear visual image. Because this is a low-frequency stimulus pattern, its evaluation should be slow, and any cognitive event depending on word identification will fail to occur. Thus, we would expect that the durations of fixations would be increased because the triggering event is delayed or missing. Figure 1 shows frequency distributions of fixation durations for three conditions: a control condition, the random letter condition, and a condition in which wordlike units are removed by replacing spaces with letters (Yang & McConkie 2001; in press). The figure shows that if fixations are long enough, the following saccade is indeed delayed. However, nonwords have very little effect on saccade onset time (that is, the experimental condition curve is similar to that of the control condition) until 175 msec after the onset of the fixation; and many fixations after that appear unaffected. A large propor-

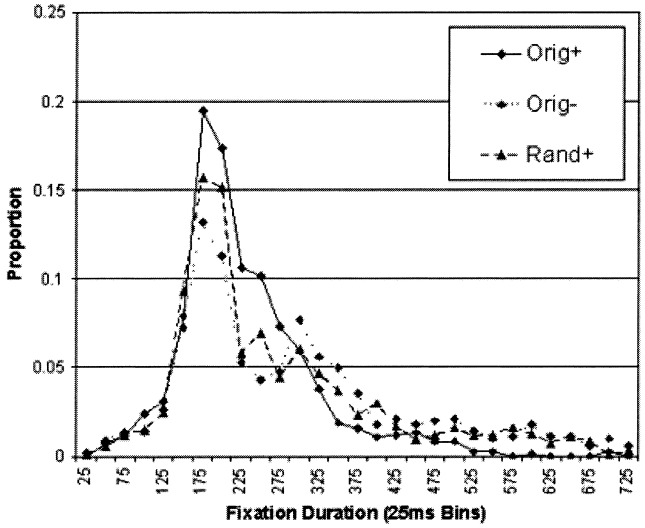

Figure 1 (McConkie & Yang). Fixation duration frequency distributions for people reading under three conditions defined by the nature of the text present during critical fixations: Orig+ (original text with spaces: control condition), Rand+ (letters replaced by random letters, but with spaces remaining in the text), and Orig− (original text but with spaces replaced by letters). Data from Yang & McConkie (in press).

tion, perhaps even a majority, of saccades occur at their normal time regardless of whether the text consists of words. Furthermore, as Figure 1 shows, a large proportion of saccades occur at their normal times even when there are no perceptual word units to which to shift attention or direct saccades. While E-Z Reader and other cognitive saccade triggering theories may expect a small set of preprogrammed early saccades, it is unclear how they can account for so many saccades occurring uninfluenced by such aberrant stimulus patterns. Triggering a saccade must not depend on word-based processes.

We have proposed an alternative explanation for saccade generation in reading, in which saccades are produced on a strategic basis, not directly determined by the language processes (see also Engbert & Kliegl 2001; Engbert et al. 2002). However, when processing difficulty occurs, the saccadic system is inhibited, reducing the likelihood of saccade generation. The nature of the processing difficulty determines the onset time and severity of the inhibition, with a given saccade being delayed only if it has not occurred by the time of the inhibition (a horserace model). Earlier and strong inhibition delays more saccades, resulting in a longer mean fixation duration. There is ample evidence that such autonomous saccades are generated in situations requiring repetitive or predictable eye movements (Basso & Wurtz 1997; 1998; Fischer & Ramsperber 1984; Vitu et al. 1995). Our proposal suggests that cognitive events inhibit saccades rather than triggering them during reading. We do not rule out the possibility that in reading, saccades can be triggered based on the guidance of cognitive information, but evidence suggests that these saccades will take much longer time to take place (Yang & McConkie, in press).

The primary strength of E-Z Reader is its ability to reproduce many previously observed phenomena. However, if further research confirms that its basic assumptions are incorrect, this may raise questions about the value of this type of evidence – and raise the interesting question of what the tests of E-Z Reader have actually been tests of. Much of the framework of the model could be preserved while changing these basic assumptions, but would it then be E-Z Reader? The future of this model will be interesting to watch.

The eye-movement engine

Wayne S. Murray

Department of Psychology, University of Dundee, Dundee DD1 4HN, Scotland. **w.s.murray@dundee.ac.uk**
http://www.dundee.ac.uk/psychology/wsmurray/

Abstract: E-Z Reader fits key parameters from one corpus of eye movement data, but has not really been tested with new data sets. More critically, it is argued that the key mechanism driving eye movements – a serial process involving a proportion of word recognition time – is implausible on the basis of a broad range of experimental findings.

Reichle et al. provide an appealing, explicit, and, to a first approximation, accurate model of eye-movement control. One problem with being fully explicit is that it encourages nit-picking concerning details of the formulation. But Reichle et al. do a pretty good job of justifying the existence of, and appropriate values for, most of their mechanisms and free parameters. However, it seems a terrible pity that this model (and others) have been developed and "tested" using a single, rather limited, corpus of data. A real test involves setting parameters and then testing the model on another corpus.

I will, however, focus here on whether the conceptual properties of E-Z Reader 7 seem appropriate, given what we know. Let's look at the engine that drives fixation timing in the model: a process that serially identifies successive words in text. Or, more specifically, a serial mechanism sensitive to some proportion of that process. This is a constant (67%) for unpredictable words but increases with predictability. In no case, however, is it greater than the time required for word recognition. Nor does it ever encompass the recognition – or even part thereof – of more than one word. Partial or even complete recognition of a subsequent word can proceed after that engine is engaged and while the gears drive the saccade into motion, but these factors cannot influence engagement of the engine.

How plausible is this? It leads to conclusions that appear to be at variance with much of what we know about eye movements. First, obviously, is the question of whether processing is strictly serial. The authors concede that two "attention gradient" models (SWIFT and GLENMORE) do a similarly good job of accounting for critical phenomena. (To say that these models "fail" to account for some phenomena that they have never attempted to model is no damning criticism.) Others, I am sure, will point out the accumulating mass of evidence pointing to the *possibility* of parafoveal difficulty influencing fixation duration. Indeed, the authors point to some of it themselves – while suggesting that it occurs only if the task "isn't quite reading." I'll content myself with the observation that well-established, uncontroversial phenomena often come and go across experiments that clearly involve "reading" (see below). The same is true of parafoveal effects in "reading-like" tasks: Murray and Rowan (1998; see also Murray 1998) found effects of the pragmatic plausibility of a yet-to-be-fixated word. The effect was replicated by Kennedy et al. (in press), but not by Rayner et al. (2003), using a different procedure. The effect was also not replicated with the same procedure but differently structured items, by Murray and Clayes (1998). However, using the same task, and items of the same structure as Murray and Clayes, but differing semantics, Murray et al. (1999) found evidence for a parafoveal pragmatic effect. It appears not to be either the task or the items that drive it, but a complex interplay of these factors – and possibly others. The same, as we will see, is true of other, less controversial, effects. To dismiss effects as coming from "non-reading tasks," as the authors are inclined to do, appears to be missing the point: The mechanism "normally" used for reading (whatever that means) is just as likely to be applied to the processing of a piece of text forming a paragraph, a line, a single sentence, a phrase, or indeed any concatenation of words; and it is just as likely to be applied when the task involves "understanding" a piece of text, as when it is to decide if this text is the same as another. If this mechanism sometimes shows evidence of parafoveal-on-foveal effects, that constitutes a priori evidence that such a process is possible (but not always necessary) in normal reading, just as we assume that syntactic or pragmatic effects sometimes reflected in first fixation durations on a critical word reflect a part of the normal reading process, despite the fact that such early effects are not always found.

It is no exaggeration to say that the time course of syntactic and pragmatic effects can be frustratingly variable. Some investigators, in some experiments (e.g. Traxler & Pickering 1996), tend to find them only "downstream" – in later measures, such as gaze duration, regional reading time, or probability of regressions. Yet other experiments demonstrate very early effects of the self-same phenomena, sometimes on the duration of the first fixation falling on the critical word. Two points are worth making: As mentioned above, variability in the time-course of these phenomena (or the fact that they sometimes show up in longer inspection times and sometimes only as increased regressions) has never been used to call into question the possibility of their (early) existence in normal reading. The second and more critical point is that early effects of this sort should not exist, according to E-Z Reader. It should not be possible for syntactic or pragmatic factors to influence the duration of the first fixation falling on a word, but there is plenty of evidence that they can (see Murray & Liversedge [1994] and Murray & Rowan [1998] for examples, but also many other studies).

When the engine that drives the saccade is 67% of word recognition, how can the timing of that saccade be affected by the nature of the syntactic or semantic combination of the identified form of that word and other words in the text? Even adopting the generous assumption that combinations of this sort start to be computed before complete recognition of the critical word, is it plausible that the consequences of that combination could then be used, within the time frame envisaged, to drive the saccadic mechanism?

The authors state that they wish to begin to incorporate an ability to account for other established linguistic phenomena into E-Z Reader. It is very difficult to see how results such as these could be incorporated, and indeed they call into question basic assumptions regarding the engine. It seems that it is driven more variably across tasks or texts, and sometimes by the properties of more than one word.

On the perceptual and neural correlates of reading models

Naoyuki Osaka

Department of Psychology, Graduate School of Letters, Kyoto University, Kyoto 606-8501, Japan. **osaka@psy.bun.kyoto-u.ac.jp**
http://www.psy.bun.kyoto-u.ac.jp/~osaka/

Abstract: The current model appears comprehensive but is probably not applicable to a writing system like Japanese, which has unspaced text, because the model is mainly based on English. The span size difference (smaller for Japanese than for English) may be a result of high-level working memory-based attentional processing and not of low-level processing. Further, neural correlates of the model are discussed in terms of central executive function.

In introducing the E-Z Reader model of reading, Reichle et al. reviewed the models that explain "the interface between vision and low-level aspects of language processing" (target article, sect. 1) in terms of eye-movement control and visuospatial attention. My first argument is based on the perceptual span and the second one is based on neural correlates of the model. The current model appears comprehensive but is probably not applicable to a writing system like Japanese, because the model is mainly based on English and other Roman alphabet-based script. Regarding perceptual span, for example, measured using the moving-window technique, the size of the span appears smaller for Japanese (about 3–4 character spaces to the right of fixation: Osaka 1992; Osaka & Oda 1994) than for English (about 14–15 character spaces to the right of fixation: McConkie & Rayner 1975). Does this difference in writing systems come from low-level eye-movement control or high-level processing involving attentional dynamics? Moreover, the model expects that the boundary of each word can be easily separated by blank spaces, as in English; that is why Reichle et al. hypothesize that the reader moves her/his eyes guided by the spaces under oculomotor control, as shown in Figures 3 and 5. However, writing systems like Japanese, Chinese, and other oriental languages lack the blank spaces between words in the text (causing a lower spatial frequency region, whereas languages like English involve high spatial frequency); this might introduce difficulty in interpreting eye-movement control tactics in Japanese in the same way as is done with English.

During eye-movement control while reading unspaced text, it was found that the eyes land on the Kanji characters (logographic symbols) more frequently than on Hiragana characters (phonetics symbols) (cf. Kajii et al. 2001) for extraction of meaning during reading. Furthermore, the systematic errors (SRE) estimated in Equation 4 of the target article were derived from English readers whose oculomotor systems "prefer" to make saccades that are seven character spaces in length, according to Reichle et al. However, this value would be influenced by differences in writing style, and most likely be different for different scripts, as described above. An alternative possible tactic under cognitive control is that the phonological loop in working memory determines when to move the eyes in the text. The identification of the currently fixated word may initiate the attentional spotlight (driven by phonological loop) to move to the next word, which in turn initiates the oculomotor system to begin programming a saccade to the next word (Morrison 1984). Further, a longer word takes a longer time to identify than a short word because the phonological loop takes longer for the former during reading, which is explicitly shown as parameter N in Equation 1. Therefore, the validity of a model applicable to a writing system *without* blank spaces might be expected to contribute toward a unified model of reading.

The second argument is based on the neural basis of visuospatial attention. Reichle et al. speak of a "low-level of language processing," not "high-level," when they refer to attention. However, visuospatial attention is not likely to be "low-level." Rather, it might involve more "high-level" processing based on the executive function of the prefrontal brain. Regarding the neural correlates of the model, the E-Z Reader model suggests an attentional neural network in the region around the intraparietal sulci and angular gyrus in the parietal brain; primary and extrastriate visual cortex in the occipital brain, inferior temporal gyrus in the temporal brain, and eye movement-related motor area (BA6 /8) in the frontal brain, are just described in Figure 14. However, the cognitive component of attentional control – that is, executive function, in the prefrontal region (i.e., BA 46/44/9 in the left brain), other than the motor component – seems more closely related to dynamic properties of visuospatial attention during reading. For example, the length of the span that is influenced by the dynamics of allocation of visuospatial attention appears to be increased for subjects with high working memory, with efficient attentional control, compared to that of subjects with low working memory (Osaka & Osaka 2002). This suggests that eye-movement control could also be influenced by attentional control by high-span subjects; in other words, working memory plays an important role in eye-movement control during reading.

Osaka et al. (2003) showed a strong functional connectivity between ACC (anterior-cingulate cortex) and left DLPFC (dorsolateral prefrontal cortex) for attention control during sentence reading: They reported that subjects with high working memory capacity (high reading span score) showed higher efficiency in controlling attention than did low capacity subjects. This was confirmed by a "focus word" experiment performed subsequently

(Osaka et al. 2002). Thus, it is likely that "higher-level" visuospatial attention appears to control optimal eye movement. Phonological store and phonological loop (each assumed to be located in the supramarginal gyrus and inferior frontal gyrus [Broca's area BA44], respectively) are subcomponents of the central executive during sentence reading that could be "interfaced" with the cognitive components of working memory. "Interfacing" refers, in my opinion, to a resource-limited attentional mechanism with executive function (Osaka et al. 2003). Therefore, it is likely that the phonological loop influences eye movements. These data suggest that the eye movement might be influenced both by the writing system and by individual working memory capacity.

Linguistically guided refixations

Jeremy Pacht

Department of Psychology, Glasgow University, Glasgow G12 8QB, United Kingdom. **jeremy@psy.gla.ac.uk**

Abstract: I discuss evidence for direct linguistic control of refixations and argue that the E-Z Reader model's account of refixations requires elaboration or revision.

What are the proximal causes of consecutive fixations on a word in reading? Four suggestions have been advanced: (1) Refixations may be due to oculomotor error in saccades targeted at another word (e.g., McConkie et al. 1989; Pollatsek & Rayner 1990); (2) Refixations may be guided by low-level, nonlinguistic information such as word length (e.g., O'Regan 1992a; Vergilino & Beauvillain 2000); (3) Refixations may reflect a trade-off between linguistically guided decisions to maintain fixation on the current word and to move the eyes to another word (e.g., Henderson & Ferreira 1990; Pollatsek & Rayner 1990); (4) Refixations may be linguistically guided movements targeting another region of the word (Hyönä & Pollatsek 1998; Pynte 1996). The E-Z Reader model allows for only the first two of these possibilities, although the model can account for some of the evidence of linguistic influence on refixation patterns indirectly. This is because the model supposes that dumb refixation decisions are less likely to win a race against linguistically based decisions to saccade to another word when the currently fixated word is (initially) easy to access.

Other evidence of linguistic influence on refixations is less easily reconciled with the E-Z Reader model. One example is evidence from Finnish that properties of a word's morphemes affect refixation location (Hyönä & Pollatsek 1998). The difficulties posed by this finding have been acknowledged in previous expositions of the model, but Reichle et al. (1999) suggest that a homologous adaptation of the current model, adopting the morpheme rather than the word-form as the fundamental lexical unit, might be capable of accommodating this result – in this case, linguistically guided word-form refixations would be reconstrued as linguistically guided intermorphemic saccades. A similar finding, not mentioned in any exposition of the E-Z Reader model, is Pynte's (1996) demonstration (using polymorphemic French words) that refixations may be preferentially directed to whichever region discriminates the word from similar words of higher frequency.

Incidental findings I obtained in a reading experiment using the boundary technique (Rayner 1975), pose further difficulties for the E-Z Reader model. In the experiment, participants read Dutch sentences for comprehension while their eye movements were monitored. Each sentence contained a monomorphemic target word primed by a parafoveal preview of varying orthographic similarity to the target word: The preview was either a higher frequency orthographic neighbor (HFN) of the target word, overlapping with the target at all letter positions but one (e.g., *spier-spies*), or an unrelated word preview, overlapping at zero letter positions (e.g., *jacht-spies*). To guard against the possibility that preview effects would be attributable to something other than the manipulated variable, the two preview groups were equated in terms of predictability from the preceding context, number of syllables and morphemes, word class, word frequency, summed bigram frequency, neighborhood size, number of higher frequency neighbors, familiarity, age of acquisition, imageability, polysemy, and (because the Dutch orthography is highly transparent) regularity. In addition, launch site distributions and the distributions of landing sites on the target word did not differ as a function of preview type. The primary aim of the experiment was to test predictions derived from the results of previous experiments, concerning the interaction of perceptual and lexical factors in visual word recognition. As expected, clear inhibitory effects of orthographic preview similarity were found in eye-movement measures such as gaze duration and total time on the target word, once well-known perceptual constraints were taken into account. The findings have been reported at a number of conferences (e.g., Pacht et al. 1999) and form the basis of a manuscript in preparation.

For present purposes, the most relevant findings concern the pattern of preview effects on the first fixation of refixated target words (FFR) and on target word refixation rates. Many studies have found that target word processing may benefit from the availability of a parafoveal preview sharing the first two or three of the target word's letters (for a review, see Rayner 1998). Consistent with these findings, I found that FFR was facilitated by the HFN preview, provided that the HFN preview and target word overlapped at the first 2–3 letter positions (255 msec vs. 273 msec, $F1(1,50) = 4.24$, $p < .05$, $F2(1,62) = 4.57$, $p < .05$). The E-Z Reader model accounts for this result (and other findings of preview benefit) by assuming the HFN preview facilitated the initial phase of target word lexical access (*L1*). By the same token, the model predicts planned refixations on the target word should have been canceled more often, given the HFN preview, resulting in fewer refixations in that preview condition. However, this was far from being the case: If anything, there was a tendency for target words to be refixated *more* often, given the HFN preview (16% vs. 14%, $Fs < 1$).

A plausible account of these findings is that the HFN preview initially facilitated target word access, by priming representations or form-neighborhoods shared by the target, but subsequently interfered with target word access by activating (or adding to the activation of) its own higher-frequency lexical representation. The initial facilitation elicited a relatively fast decision to move the eye, while later-emerging lexical competition elicited a decision to fixate the current word, which might be construed as the initial "where" decision, or as a supervening "where" decision to maintain fixation or to refixate. Two implications for models of eye-movement control in reading are that the execution of refixations may *follow* execution of linguistically guided saccades (or at least, "when" decisions), and that refixations may themselves be *proximally* (and not only indirectly) controlled by linguistic variables. Both of these implications are at variance with the assumptions of the E-Z Reader model.

In sum, while some refixations may be planned without reference to linguistic information, others appear amenable to direct linguistic influence. I will close by suggesting one way in which my findings might be reconciled with the E-Z Reader model. If refixations are defined not as consecutive fixations on a word but as consecutive fixations during which the current word is processed, then according to the E-Z Reader model some refixations are indeed proximally controlled by linguistic variables and follow execution of linguistically guided "when" decisions. Specifically, the immediate regressions, which the model assumes arise when an intended interword saccade is executed before the current word is fully accessed, may be viewed as refixations following on the heels of a prior but improperly executed attempt to refixate. That is, in such cases, the "intended interword saccade" is in fact intended as a refixation at the moment the movement is executed. This amounts to a proposal that in its labile phase, the interword saccade destined for word$_{n+1}$ may be modified in two ways. First, as the current model allows, in cases where *L1* is completed on word$_{n+1}$, the saccade may be replaced by a saccade targeted on the following word; second, however, in cases where difficulty is detected in accessing the current word, the saccade may be replaced by a saccade targeted on the same word. The model would then posit two types of refixation, one driven only by low-level factors, the other guided by cognitive constraints.

Regressions and eye movements: Where and when

Manuel Perea[a] and Manuel Carreiras[b]

[a]*Departament de Metodologia, Universitat de València, 46010-València, Spain;* [b]*Departamento de Psicología Cognitiva, Universidad de La Laguna, Campus de Guajara, 38205-Tenerife, Spain.* **mperea@uv.es**
mcarreir@ull.es **http://www.uv.es/~mperea**

Abstract: Reichle et al. argue that the mechanism that determines where to fixate the eyes is controlled mostly by low-level processes. Therefore, unlike other competing models (e.g., the SWIFT model), the E-Z Reader model cannot account for "global" regressions as a result of linguistic difficulties. We argue that the model needs to be extended to account for regressive saccades.

Two basic assumptions of the E-Z Reader model are that the mechanism responsible for *where* to fixate the eyes is controlled mostly by low-level processes, whereas the mechanism responsible for *when* to move the eyes is controlled mostly by cognitive processes. Although the model accounts for fixation durations, refixation/skipping probabilities, and initial landing positions in normal silent reading, it leaves regressive saccades unaccounted for. It is worth noting that a competing model, the SWIFT model (Engbert et al. 2002), can capture both short- (*local*) and long-range (*global*) regressions. Normal silent reading involves not only forward saccades, but also a number of regressions back to the previous word(s) when readers experience some difficulties with linguistic processing (or with oculomotor processes). Bear in mind that regressions represent around 14% of saccades for adults (and around 25% for children; Starr & Rayner 2001). The point we raise here is that, in regressions, the signal of where to send the eye does not seem to be controlled solely by oculomotor variables. Instead, cognitive processes can signal where to fixate the eyes next in order to resolve conflicting information from the text or to finish processing partially encoded information. We present two examples from recent research: one with sentences involving a target word with (or without) higher frequency *neighbors* (the neighborhood frequency effect; "local" regressions) and the other with sentences that include a mild garden path ("global" regressions).

Several eye movement experiments have shown that the number of regressions back to the target word in a sentence increases when the target word has higher frequency neighbors (see Perea & Pollatsek 1998; Pollatsek et al. 1999a). For example, in the sentence "The store didn't sell John's favourite [spice, sauce] any more," readers make more regressions back to the target word *spice* than to the target word *sauce.* (Note that *spice* has *space* or *spite* as higher frequency neighbors; *sauce* does not have any higher frequency neighbors.) Under these conditions, the target word may have been misidentified as the higher frequency candidate (*space* instead of *spice*) or, alternatively, the higher frequency neighbor could have slowed down the final stage of lexical processing (e.g., in an interactive activation system). This actually provokes an increased number of regressions back to the target word for words with higher frequency competitors. In the E-Z Reader model, the signal that word recognition is imminent (*L1* stage) causes the preparation of the saccadic movement on the word$_{n+1}$ before lexical access (*L2* stage) is completed. A regressive saccade may occur when the *L2* stage is long and the reader is still processing the target word. In that case, the target of this saccade is the difficult-to-process word$_n$. Thus, the E-Z Reader model, despite not having a specific mechanism for regressive saccades, can predict the presence of these "local" regressions as a special type of refixation. It is important to note that the SWIFT model (Engbert et al. 2002), which borrows the two word identification stages from the E-Z Reader model, can also capture these local regressions as a result of incomplete lexical processing.

The E-Z Reader model can accommodate short, local regressive saccades as a special type of refixations. But what about global regressive saccades? Are they simply triggered by high-level processes blindly, in the sense that they do not indicate exactly which part of the sentence the eyes should be directed to? This does not seem to be the case. The pattern of regressive eye movements while reading mild garden-path sentences strongly suggests that readers perform an overt selective reanalysis process (see Meseguer et al. 2002). This process seems to direct the regressive saccade to specific points of the sentence in which relevant information can be picked up (see also, Kennedy et al. 2003). In other words, the reader's eye seems to be intelligently led to the critical part of the sentence. In the E-Z Reader model, only one word can be attended to at a time, and the model has no straightforward means to redirect the eye to the relevant area of the information in the sentence. (These regressive saccades are beyond the scope of the current implementation of the model.) One possible way to accommodate these regressions is to assume that readers have access to some form of spatially coded information (Kennedy 2001). Alternatively, in the framework of a "guidance by gradient" model (i.e., more than one word can be attended to at a time) like SWIFT, it is possible to send the eye back to the critical point of the sentence where the reader experienced some linguistic difficulties (global regressions; see Engbert et al. 2002, Fig. 7).

Therefore, one challenge of a sequential attention-shift model like the E-Z Reader is to specify in detail how regressions are made without violating the "when/where" principle. We agree with Reichle et al. that it may be difficult to make precise predictions in parsing experiments. However, inclusion of an explicit mechanism for regressions is not an obstacle. As stated above, the SWIFT model captures the presence of global regressive saccades by assuming that the gradient of attention is not confined to individual words, but rather, to a wider attentional window. We should also note that this issue may be linked to the fact that readers seem to extract information from more than a word at a time (see Inhoff et al. 2000). Whether these are critical limitations for attention-shift models (note that these models can be considered extreme cases of "guidance by gradient" ones) is a matter for future research.

ACKNOWLEDGMENTS
Support was provided by grants BSO2001-3492-C04-01 and BSO2002-03286 from the Spanish Ministry of Science and Technology.

Attention, saccade programming, and the timing of eye-movement control

Ralph Radach[a], Heiner Deubel[b], and Dieter Heller[a]

[a]*Institute of Psychology, Technical University of Aachen, 52056 Aachen, Germany;* [b]*Department of Psychology, Ludwig-Maximilians-Universität München, D-80802 Munich, Germany.* **ralph@psych.rwth-aachen.de**
http://www.psych.rwth-aachen.de **deubel@psy.uni-muenchen.de**
http://www.paed.uni-muenchen.de/~deubel
dieter.heller@psych.rwth-aachen.de
http://www.psych.rwth-aachen.de

Abstract: E-Z Reader achieves an impressive fit of empirical eye movement data by simulating core processes of reading in a computational approach that includes serial word processing, shifts of attention, and temporal overlap in the programming of saccades. However, when common assumptions for the time requirements of these processes are taken into account, severe constraints on the time line within which these elements can be combined become obvious. We argue that it appears difficult to accommodate these processes within a largely sequential modeling framework such as E-Z Reader.

In this commentary, we address three aspects that are relevant for the time line of word processing and eye-movement control in continuous reading: (1) the time it takes to lexically process a word, (2) the reprogramming time needed to alter the amplitude of a saccade, and (3) the question of whether "attention shifts" may also take time to be prepared and executed.

(1) Sereno et al. (1998) found in event related potential (ERP) studies on single-word recognition that N150 responses are sensitive to differences in word frequency. The responses for low and high frequency words start to diverge at about 130 msec, providing an indication for the minimal time required for any substantial lexical analysis. This roughly corresponds with the figure given in the target article "that the mean time needed to identify the word 'the' (the most frequent word in English text) when it is centrally fixated and in a completely predictable context is 148 msec" (sect. 3.1.2, last para.). In this specific case the time for L_2 is assumed to be zero, hence 150 msec is the time needed for L_1 under the most favourable circumstances.

(2) Looking at the other end of the time line, the question arises: How long, minimally, does it take to program or reprogram a saccade to a specific target word? This question can be discussed on the basis of the findings from the so-called double-step paradigm (Becker & Jürgens 1979), which have laid the foundation for the distinction between a labile and a nonlabile stage of saccade preparation (Morrison 1984). In a typical double-step experiment, a fixation target is shown at an eccentric location and, before a saccade can be executed, a second target is presented while the first disappears. Depending on the size and direction of the second target step, two basically different types of responses can be observed (Ottes et al. 1984). In the first, *averaging* mode, there is a continuous transition of the primary saccade amplitude from landing positions near the first target to positions close to the second target. This amplitude transition is a function of the available reprogramming time between the occurrence of the second target and the execution of the primary saccade. Importantly, the critical temporal window for saccade modification closes at 70–90 msec before saccade execution. The second response mode is characterized by *bistable* responses, which can be observed when the distance between the two stimuli is large or the direction of the saccade needs to be changed. In this case, landing positions of primary saccades cluster at both target locations. The succession of progressive saccades in reading appears consistent with the averaging response mode (note that sentences with regressions are removed from the data base E-Z Reader is tested with), suggesting that the absolute minimum time for the nonlabile stage of saccade programming is 70 msec. Alternatively, interpreting the nonfixation of words (skipping) in analogy of a bistable response mode would be consistent with the fixation duration on the origin word being increased (see below). In this response mode, the minimum reprocessing time is assumed to be 120 msec (see Deubel et al. 2000 for further detailed discussion).

Empirically, the question of whether fixation durations before word skipping are inflated is under dispute. It appears that some studies have found this effect and others have not. Critically, Radach and Heller (2000), in addition to reanalysing a sentence reading experiment, examined a very large corpus of reading data. Carefully controlling for factors like the fixation pattern on the origin word and launch site relative to the target word, they found no evidence in favour of such a phenomenon. It may thus appear premature to list the effect in Table 1 of the target article. Reichle et al. have noted with respect to the Glenmore model by Reilly and Radach (2003a) that "it remains an open question as to whether the model can predict the costs that have been observed for skipping" (target article, sect. 4.6). It is true that the phenomenon would not fit well with the mechanics of Glenmore. However, given the present state of affairs, we see no need to account for it in the model and look forward to seeing how the empirical debate on the issue will develop.

(3) We are in sympathy with the addition of a preattentive processing stage to the architecture of E-Z Reader and welcome the clear separation of visual selection for the purpose of saccade generation from selection preceding cognitive (lexical) processing (see Schneider & Deubel 2002 for a recent discussion in a more general context). Specifically, Reichle et al. reserve the term "attention" for "the process of integrating features that allows individual words to be identified" (sect. 3.1.3). In the description of the model, the authors have asserted many times that attention shifts from word to word as a result of completing lexical access. This raises a fundamental question. If the shifting constitutes a *movement* of attention, would this movement itself not need to be programmed, and would its preparation and execution not take a certain amount of time? If the answer to this question is that the shifting is merely equivalent to starting the lexical processing of a new word, then using the term *attention* in this context becomes rather meaningless. If however, the shift is seen as an obligatory stage that constitutes a precondition for the start of linguistic processing, then this process will have a latency and it will need time to be executed. Indeed, this is a major issue in the attention literature. The respective time interval is often referred to as *attentional dwell time,* and usual estimates of its duration are on the order of at least 50 msec (Duncan et al. 1994; Treisman & Gelade 1980).

Together, these considerations imply the following constraints to a tentative time line: Take 130 msec as a conservative estimate for the duration of L_1 on word$_n$ and 70 msec as a conservative estimate for the minimal duration of the non-labile stage of saccade programming. Given a fixation duration of 250 msec, the summed duration of these two stages in a sequential time line leaves a time of only 50 msec for all remaining processes. In the case of skipping word$_{n+1}$ this time would have to include the attention shift to $n+1$, the completion of some lexical processing (L_1) of this word, and the reprogramming of a saccade to word$_{n+2}$. Given the commonly observed phenomenon of skipping words that are relatively difficult to process, it is hard to conceive a scenario such as in Figure 5C of the target article, where word$_{n+2}$ becomes the target of the next saccade after less then 150 msec. Finally, in the case depicted in Figure 5B where word$_{n+2}$ becomes the target of the saccade after less than 100 msec, the question arises how this pattern could have emerged in the computational implementation of E-Z Reader. In any case, it appears incompatible both with the verbal description of the model and the time line constraints discussed above.

E-Z Reader 7 provides a platform for explaining how low- and high-level linguistic processes influence eye movements

Gary E. Raney

Department of Psychology, University of Illinois at Chicago, Chicago, IL 60707. **geraney@uic.edu** **www.uic.edu/~geraney**

Abstract: E-Z Reader 7 is a processing model of eye-movement control. One constraint imposed on the model is that high-level cognitive processes do not influence eye movements unless normal reading processes are disturbed. I suggest that this constraint is unnecessary, and that the model provides a sensible architecture for explaining how both low- and high-level processes influence eye movements.

Reichle et al. describe E-Z Reader 7 as a processing model of eye-movement control in reading. This reflects the assumption that ongoing cognitive processes influence when and where the eyes are moved. Despite this assumption, Reichle et al. make the strong claim that "higher-order processes intervene in eye-movement control only when 'something is wrong'" (sect. 3.1). The justification for this claim is that the process of integrating semantic and syntactic elements of a text occurs too late in the processing stream to influence decisions about when and where to move the eyes. This claim seems inconsistent with the word identification process within E-Z Reader 7, and, if anything, understates the capabilities of the model. Here I discuss what I believe is a minimized strength of E-Z Reader 7, namely, that the model provides a natural framework for explaining how high-level cognitive processes influence eye-movement control. I begin by addressing the question: What is a high-level process? I then discuss how high-level processes might be explained within E-Z Reader 7.

What constitutes a high-level process is partly an issue of definition. High-level processes may be defined as processing based on information not contained within the lexical representation of a word. This is similar to the description of top-down processes in models of word processing (e.g., McClelland & Rummelhart 1981). Note that E-Z Reader 7 includes predictability as an element of word identification. Processes based on word predictability qualify as high-level processing in the sense that predictability accumulates across words and sentences. Monitoring predictability to enhance word identification appears to be a normal component of reading. I suspect that Reichle et al. would agree given that the predictability is a component of both the L_1 and L_2 stages of word identification in E-Z Reader 7. Based on the above definition, predictability represents an example of high-level information that directly influences eye-movement control during normal reading. Therefore, the conclusion that high-level processes influence eye movements only when "something is wrong" seems inconsistent (and unnecessary) with the structure of the model.

High-level processes may also be defined as based on later-occurring semantic processing, thereby excluding early-occurring visual processing. Determining whether high-level processes occur too late in the processing stream to influence eye movements becomes a critical issue. Although there is evidence for many high-level processes being slow, in the sense that they occur in late stages of word processing or even after a problematic word has been read (e.g., the garden path sentences used by Frazier & Rayner [1982]), there is growing evidence that high-level processes can influence early stages of word identification (Morris 1994; Sereno 1995; Wiley & Rayner 2000). This evidence again calls into question the necessity of the claim that high-level processes do not influence eye movements unless something goes wrong. My purpose here is not to resolve this definition issue but to suggest that Reichle et al. might be constraining their model unnecessarily. An untapped strength of E-Z Reader 7 is that it provides a transparent (i.e., definable) architecture for explaining how high-level processes influence eye movements (at least the decision of when to move the eyes). This contrasts with other models in which the architecture is not always transparent (such as, how hidden layers operate in connectionist models). Thus, my criticism of E-Z Reader 7 is that the architecture of the model is not fully utilized. Below I provide two examples of how the model may be applied.

Including a two-stage word identification system provides a natural architecture for separating the locus of low- and high-level processing influences on eye movements. Recent studies from my own lab support this conclusion. In one study (Raney et al. 2000), I recorded subjects' eye movements while they read a text once and then read either the same text a second time or a paraphrased version of the text. Paraphrases were created by replacing words with synonyms. For identically repeated target words, both first fixation duration and gaze duration were reduced during the second reading. For synonyms, only gaze duration was reduced during the second reading. For synonyms, early-occurring orthographic processing was not facilitated whereas later-occurring semantic processing was facilitated. This makes sense because no orthographic repetition occurs for synonyms, but semantic repetition does occur. In terms of the E-Z Reader model, the results for synonyms reflect no facilitation of the L_1 stage of word identification, but facilitation of the L_2 stage (a reduction in gaze duration, which reflects more later-occurring processes than first fixation duration).

In a similar study (Raney et al. 1996), fluent and nonfluent bilinguals read a text in one language and then reread either the same text or a translation. Embedded in the texts were cognate and noncognate target words. For fluent bilinguals, fixation durations were equivalent for cognates and noncognates during the second reading. For nonfluent bilinguals, fixation durations were shorter for cognates than for noncognates during the second reading. The low-level benefit of repeating the orthographic form (for cognates) interacted with high-level processes associated with comprehension level (fluency). These findings also map onto the model. Specifically, only semantic processes influenced fixation duration for fluent bilinguals (L_2), but both orthographic (L_1) and semantic processes (L_2) influenced fixation duration for nonfluent bilinguals.

To summarize, Reichle et al. describe E-Z Reader 7 as a processing model of eye-movement control. One constraint they impose on the model is that high-level cognitive processes do not influence eye movements unless normal reading processes are disturbed. This constraint makes the model conservative regarding what forms of information are allowed to influence eye movements. My own view is that there is enough evidence that high-level processes influence early and late stages of eye movements, for models of eye-movement control to incorporate these processes. E-Z Reader 7 provides a sensible architecture for explaining how high-level processes influence eye movements. Constraining the impact of high-level processes reduces the explanatory power of the model.

ACKNOWLEDGMENTS

I wish to thank John Murray and Bette Bottoms for their helpful comments on drafts of this commentary.

Methodologies for comparing complex computational models of eye-movement control in reading: Just fitting the data is not enough

Ronan Reilly[a] and Ralph Radach[b]

[a]*Department of Computer Science, NUI Maynooth, Maynooth, Co. Kildare, Ireland;* [b]*Institute of Psychology, Technical University of Aachen, 52064 Aachen, Germany.* **ronan.reilly@may.ie** **http://cortex.cs.may.ie** **ralph.radach@psych.rwth-aachen.de** **http://www.psycho.rwth-aachen.de**

Abstract: As the number of computational models of eye-movement control in reading increases, so too will their coverage and complexity. This will make their comparison and testing increasingly challenging. We argue here that there is a need to develop a methodology for constructing and evaluating such models, and outline aspects of a possible methodology.

In recent years, research on eye movements in reading has made substantial progress. A key new development in the field is the emergence of computational models of eye-movement control during reading. The target article is a timely evaluation of this branch of reading research. The modeling principles and algorithms that different computational models embody reflect the theoretical viewpoints of their authors. In the case of E-Z Reader, sequential lexical processing is proposed as the obligatory trigger for the generation of all eye movements made in normal reading.

In contrast, Reilly and O'Regan (1998), following the theoretical framework developed by O'Regan (1990), demonstrated that a good account for the positioning of fixations in reading can be achieved by using a set of rather simple oculomotor heuristics. We believe that both of these positions have their merits and can account for important aspects of eye behaviour during reading. On the other hand, both approaches also have serious limitations. Therefore, the question of interest is not whether eye movements are determined by visuomotor factors or by linguistic processing, but to what degree these two factors are involved and how they interact.

Recent modeling work by Engbert and Kliegl (2001) and Reilly and Radach (2003a) can be seen as attempts at reconciling these two views of the reading process. As things stand, however, all current computational models of eye-movement control in reading deal with the process at a relatively shallow level. As pointed out in the target article, one of the real challenges for the next generation of models will be to broaden their coverage to include cognitive and linguistic factors. Unfortunately, as models become more complex, their comparison will become more problematic. The main point of this commentary, therefore, is to make the case for the development of a methodology for the comparison of computational models of eye-movement control in reading.

Our methodological proposals fall under three headings: (1) the facilitation of the comparison of the structural and functional assumptions of competing models; (2) the grounding of models in the neuroscience of vision and language; and (3) the establishment of data sets for model comparison and benchmarking. With regard to the comparison of the structure and function of models, this could be facilitated by using a common implementation framework comprising a set of reusable software components (Schmidt & Fayad 1997). In software engineering terms, a framework is a reusable, "semicomplete" application that can be specialised to produce particular applications or, in this case, particular models. The components would need to be fine-grained enough to accommodate the range of model types and model instances described in the target article. If one could develop an acceptable and widely adopted modeling framework, it would be possible to establish a common basis on which to implement a variety of models. This would make the models more directly comparable not only in terms of their ability to account for data, but also in terms of their underlying theoretical assumptions. The modeling environment could provide a semi-formal language with which a model's structures and processes function could both be unambiguously articulated. This would aid the task of both designing the models and communicating the design to other researchers.

Functionalist computational models, of which E-Z Reader is an excellent example, are inherently underdetermined in terms of their relationship to the brain mechanisms that underlie them. For example, one could envisage a family of E-Z Reader-like models with quite different combinations of parameters and/or parameter values that would be capable of providing an equally good fit to the empirical data (e.g., Engbert & Kliegl 2001). One way to reduce this lack of determinism is to invoke a criterion of biological plausibility when comparing models. We agree with the authors that there is an increasingly rich set of data emerging from the field of cognitive neuroscience which could be used to augment the traditional behavioural sources of constraint on computational models. We believe that models of reading can no longer avoid scrutiny from this perspective. Another, not unrelated, factor in assessing competing models is to take due account of the evolutionary context in which our vision system evolved. Because it evolved for purposes quite different from reading, we need to beware of too-easy recourse to arguments of parsimony, particularly when they are couched solely in terms of the reading process itself. A model with the minimum of modifiable parameters may be parsimonious on its own terms but fail the test of biological realism when compared with, say, a model that comprises an artificial neural network with many hundreds of adjustable parameters. While evolution is parsimonious in the large, when we look at brain subsystems in isolation, such as those involved in reading, we need to be careful how we wield Occam's razor.

Finally, the issue of appropriate data sets with which to test and compare computational models of eye-movement control needs closer attention than has been given to date. The Schilling et al. (1998) data set used to parameterise and test E-Z Reader and several other models discussed in the target article is not particularly extensive. A good case can be made for establishing a range of publicly accessible data sets against which any proposed model can be tested. This would be similar to what has been done, for example, in machine learning, in data mining, and, most notably, in the field of language acquisition (MacWhinney 1995). Furthermore, the corpus of benchmark data should be extended to include a variety of languages, alphabets, and scripts. The more successful models will be those that can readily generalise beyond just one language and one writing system.

Eye-movement control in reading: Models and predictions

Eyal M. Reingold

Department of Psychology, University of Toronto, Toronto, Ontario M5S 3G3, Canada. **reingold@psych.utoronto.ca**
http://psych.utoronto.ca/~reingold/

Abstract: It is argued here that a critical prediction of the E-Z Reader model is that experimental manipulations that disrupt early encoding of visual and orthographic features of the fixated word without affecting subsequent lexical processing should influence the processing difficulty of the fixated word without producing any processing effect on the next word. This prediction is explained and illustrated.

In the target article, Reichle et al. introduce a comprehensive framework for evaluating models of eye-movement control during reading. The authors also provide an updated version of the E-Z Reader model (Reichle et al. 1998; 1999) and argue that the qualitative and quantitative predictions that follow from this model closely match empirical findings concerning a wide range of reading phenomena. Consequently, they contend that the new version of their model, E-Z Reader 7, constitutes the best available computational framework for modeling eye-movement control during reading. The purpose of this commentary is to derive and illustrate a critical and as yet untested prediction that is unique to the E-Z Reader model. The proposed empirical strategy is illustrated in Figure 1 and will be outlined below.

As illustrated in Figure 1, three core aspects of the E-Z Reader model are central to the present proposal: (1) The E-Z Reader model introduces a distinction between two stages of lexical processing: an early lexical processing stage corresponding to the extraction and identification of the orthographic form of the word (L_1), and a late stage involving access to the phonological and semantic forms (L_2); (2) the programming of a saccade to the next word (word_{n+1}) is initiated following the completion of L_1 of word_n; and (3) parafoveal preview of word_{n+1} begins following the completion of L_2 of word_n. Therefore, according to the E-Z Reader model, variation in the duration of L_2 of word_n – $t(L_2)$ – critically determines the duration of parafoveal preview of word_{n+1}.

As shown in Figure 1, the duration of the parafoveal preview of word_{n+1} equals the duration of the interval between the initiation and execution of the saccade to word_{n+1} minus $t(L_2)$ of word_n. In the current implementation of the E-Z Reader model, two variables, word frequency and contextual constraint or predictability, influence the duration L_2 of word_n and consequently should also control the duration of the parafoveal preview of word_{n+1} and the magnitude of any benefit when word_{n+1} is later fixated (e.g., shorter fixations on word_{n+1}, greater probability of skipping word_{n+1}). Consistent with this prediction, greater parafoveal preview benefit on word_{n+1} has been demonstrated when word_n is a high frequency word (Henderson & Ferreira 1990; Kennison & Clifton 1995) and when word_n is highly predictable from the preceding text (Balota et al. 1985). These findings are typically taken to suggest that as the difficulty of foveal processing increases, parafoveal preview benefit decreases.

However, the E-Z Reader model dictates more precise inferences concerning any effects of experimental manipulations of the characteristics of word_n on the subsequent processing of word_{n+1}

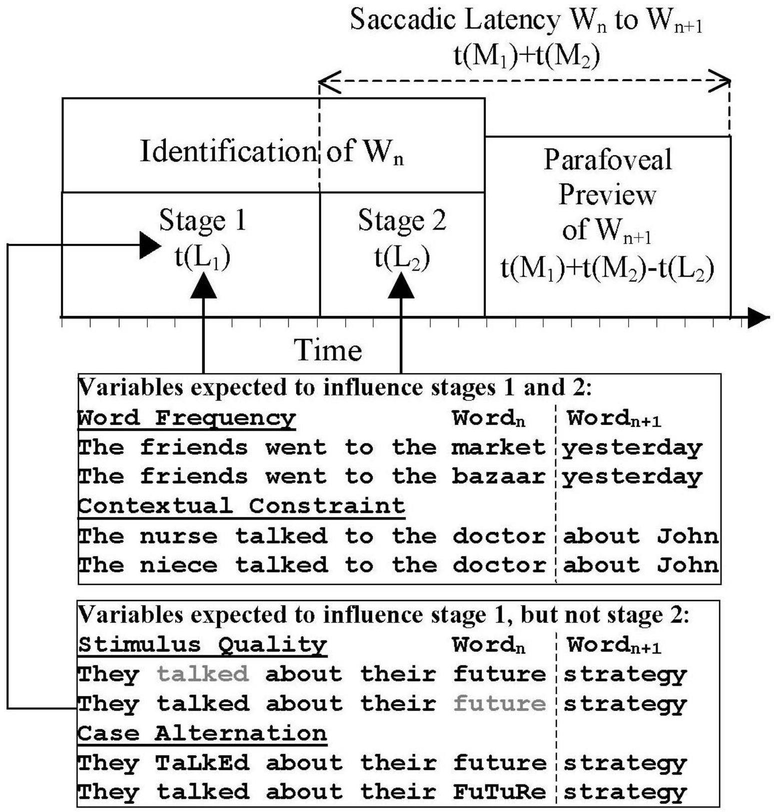

Figure 1 (Reingold). An illustration of the core assumptions of the E-Z Reader model and the methodology proposed for testing this model. The completion of the first stage of lexical processing (L_1) of word$_n$ (W_n) signals the oculomotor system to begin programming a saccade to word$_{n+1}$ (W_{n+1}). On the average, the duration prior to the execution of this saccade (i.e., the saccadic latency) is equivalent to a constant representing the mean duration of two saccadic programming stages ($t[M_1] + t[M_2]$). The duration of the parafoveal preview of word$_{n+1}$ equals the duration of this saccadic latency minus the duration of the second stage of lexical processing $t(L_2)$ of word$_n$ ($t[M_1] + t[M_2]-t[L_2]$). Consequently, experimental manipulations that are expected to influence stage 1 of the lexical processing of word$_n$, but not stage 2 (e.g., stimulus quality, case alternation), are predicted to affect the fixation duration on word$_n$ without influencing the duration of the parafoveal preview of word$_{n+1}$. In contrast, experimental manipulations that are theorized to influence both stages 1 and 2 of the lexical processing of word$_n$ (e.g., word frequency, contextual constraint) have been shown to influence both the fixation duration on word$_n$ and the parafoveal processing of word$_{n+1}$.

(i.e., parafoveal preview effects or spillover effects). Specifically, such effects on word$_{n+1}$ are predicted *if and only if* an experimental variable influences L_2 of word$_n$. This has two vital implications for the empirical validation of the E-Z Reader model. First, variables other than frequency and predictability, such as syntactic difficulty, that have been shown to modulate the magnitude of the parafoveal preview benefit (Henderson & Ferreira 1990) can be inferred to influence the duration of L_2, suggesting possible extensions to the E-Z Reader model. Second and more important, variables influencing L_1, but not L_2 of word$_n$, while modulating the difficulty of the lexical processing of word$_n$, should not affect the magnitude of any processing benefit when word$_{n+1}$ is later fixated. This marks a clear departure from the hypothesis that a substantial increase in the difficulty of foveal processing invariably results in a decrease in the parafoveal preview benefit (Henderson & Ferreira 1990).

I argue here that searching for variables that influence the processing difficulty of word$_n$ without producing any processing effect on word$_{n+1}$ is a critical test of the E-Z Reader framework. A closer examination of the L_1 versus L_2 distinction proposed by E-Z Reader 7 suggests several potentially promising experimental manipulations. Essentially, the L_1 versus L_2 distinction assumes an early lexical processing stage corresponding to the extraction and identification of the orthographic form of the word and a late stage involving access to the phonological and semantic forms. Accordingly, a disruption early in the word recognition system when visual features are encoded and abstract letter identities are computed should be expected to influence L_1, but not L_2. Two manipulations that are generally believed to disrupt early encoding of visual features are illustrated in Figure 1: *stimulus quality* (for a review see Borowsky & Besner 1993) and *case alternation* (for a review see Mayall et al. 1997). However, it has been recently suggested that unlike the former, the latter variable may also influence post-encoding lexical processing (Herdman et al. 1999) or attentional processing (Mayall et al. 2001). It is important to note that the description of the manipulations of stimulus quality and case alternation is meant to merely illustrate, rather than exhaustively detail, the potential value of the proposed research strategy for the of study eye-movement control during reading in general, and the empirical validation of the E-Z Reader model in particular.

ACKNOWLEDGMENTS

This research was supported by a grant to Eyal Reingold from the Natural Science and Engineering Research Council of Canada (NSERC). The author is grateful to Keith Rayner, Elizabeth Bosman, Ava Elahipanah, and Colleen Ray for their helpful suggestions and input. Correspondence concerning this commentary should be sent to Eyal M. Reingold, Department of Psychology, University of Toronto, 100 St. George Street, Toronto, Ontario, Canada M5S 3G3. E-mail: reingold@psych.utoronto.ca.

Neural plausibility and validation may not be so E-Z

Sara C. Sereno[a], Patrick J. O'Donnell[a], and Anne B. Sereno[b]

[a]*Psychology Department, University of Glasgow, Glasgow, G12 8QB, Scotland, United Kingdom;* [b]*Department of Neurobiology and Anatomy, UT-Houston Medical School, Houston, TX 77030.* **s.sereno@psy.gla.ac.uk p.odonnell@psy.gla.ac.uk anne.b.sereno@uth.tmc.edu**

Abstract: Although the E-Z Reader model accounts well for eye-tracking data, it will be judged by new predictions and consistency with evidence from brain imaging methodologies. The stage architecture proposed for lexical access seems somewhat arbitrary and calculated timings are conservatively slow. There are certain effects in the literature that seem incompatible with the model.

The E-Z Reader model is successful in handling a number of eye-movement phenomena that have been observed during reading. It fares better than other models, although it is not strikingly better than, for example, Glenmore or SWIFT. Notably, there is no comparison across models concerning the functional assumptions about attention shifts or the plausibility of neural mechanisms. It is unlikely that eye-tracking alone and changing parameters of the E-Z Reader model will validate these assumptions. The model has some limitations in its empirical and theoretical justifications (e.g., separation of word identification into two stages, separation of the attentional shift from eye-movement programming), has questionable time frames, and has difficulties with some observations (such as fixation time and word length).

The stage distinction of word identification is crucial to the model because completion of Stage 1 (access to orthographic form) drives the oculomotor system to program the next saccade, while completion of Stage 2 (access to phonological and semantic forms) shifts the spotlight of attention to the next word. A disconnection with independently verifiable temporal estimates for word processing beyond eye-movement measures is a limitation. The danger here is that in setting out to establish a model of eye-movement control, the result may be a model of eye-movement experiments. In the end, a theory of reading will draw on our knowledge of the temporal flow of information across visual and higher-order brain areas, beginning with the first afferent volley that reaches frontal cortex 80 msec post-stimulus and continuing through the top-down feedback loops, which modulate further processing in sensory areas (e.g., Foxe & Simpson 2002). Dense-mapping event-related potential (ERP) and magnetoencephalogram (MEG) recordings provide the requisite spatiotemporal granularity necessary to trace the network of cortical activations millisecond by millisecond. Such research has consistently pushed back estimates of when lexical access, semantic, and top-down contextual effects are first manifest (e.g., Pulvermüller et al. 2001; Rousselet et al. 2002; Sereno et al. 1998; 2003). Assumptions about the temporal course of lexical access should incorporate such finer-grained analyses and use them to test specific predictions that emerge from imputed discrete stages.

One concern about the E-Z Reader model relates to the role of attention. The model decouples saccadic programming from certain shifts of attention. That is, saccadic programming to word$_{n+1}$ begins after the first stage of lexical processing, but an "attention shift" occurs only after a second stage, when the attended word$_n$ has been identified. However, the authors then explain that they are not talking about spatial attention (i.e., attention to spatial orientation), but rather, about attention to "feature," as spatial attention shifts with saccade programming (cf. Sereno 1996). Specifically, they argue that only when an attended word has been identified will attention "shift" to the next word (meaning a shift in some feature space or "analyzer selection"). However, much work in both behavioral studies and physiology has shown that spatial attention improves sensitivity to features at the spatially attended location (e.g., Reynolds et al. 2000; Yeshurun & Carrasco 1998). By artificially separating spatial attention from attention to features, it is not clear or "E-Z" to see how features would be kept separate for proper integration, much less how enhanced stimulus processing from two spatially separate regions is handled in a "strictly serial fashion."

Another concern about the model arises from the implications of its various specifications. For example, Equation 1 can be used to calculate fixation time with regard to word length and fixation location. Given a central fixation point, fixation times for 3-letter and 13-letter words are 95 msec and 115 msec, respectively. It seems implausible that such disparity in word length amounts to only a 20 msec cost in processing. Furthermore, the division of lexical access into two discrete stages seems rather arbitrary: From stage 1 to stage 2, frequency steps down from full to half strength while predictability steps up from half to full strength. Even if one assumes this captures the totality of lexical processing, it would be more parsimonious and neurologically valid to express the decay and growth of these factors as graded functions over time within a single stage. Nonetheless, by summing 90 msec of early visual processing (apparently not accounting for differences in word length) with Stages 1 and 2 (Equations 2 and 3), "word identification time" can be calculated given a word's frequency and predictability. For words that differ substantially in frequency (5 vs. 150 per million) but only slightly in predictability (0.4 vs. 0.6 cloze probability), the resulting times are somewhat vexing: The size of the frequency (34 msec) and predictability (39 msec) effects are comparable, a result that is at odds with the literature in which more robust effects are evident for frequency. In the model, word identification times begin at 148 msec (highest frequency and predictability) and extend to 432 msec (lowest frequency and predictability), a range of 284 msec. However, given that the saccadic programming duration of 240 msec commences at the end of Stage 1 of access and that the longest time to complete Stage 2 is 114 msec, programming duration will always subsume that of Stage 2. Summing Stage 1 and saccadic programming times (across all conditions of frequency and predictability) presumably yields fixation times and these extend from 388 to 558 msec, a range of only 170 msec. Such times seem debatable both in terms of their inflated duration and lack of variability in comparison to fixation times in reading.

Finally, the model seems unable to account for various effects within the literature. Some of these include the interactions of frequency and predictability and of frequency and spelling – sound regularity (e.g., Sereno & Rayner 2000). Also, the "fast priming" eye-movement literature has shown that activation of phonological, semantic, and contextual information occurs very early in lexical processing (Rayner et al. 1995; Sereno 1995; Sereno & Rayner 1992). In a final example, the authors simulate certain lexical ambiguity effects by asserting that the subordinate sense of a (high-frequency) ambiguous word can be treated like a low-frequency word (cf. Sereno et al. 1992). While the word's meaning is of low frequency, its orthographic form is not. And though early access to word meaning may very well drive eye movements, this violates a basic assumption of their model.

We recognize that alternative methods of investigating the time course of semantic processing or the spatiotemporal allocation of attention, such as ERPs or cuing experiments, while offering greater control and finer-grained results, are handicapped with respect to normal reading. However, models of normal reading also make assumptions about lexical access and attention not readily revealed by eye-movement data. The task for the future is to integrate data from complementary, neurally based methodologies.

ACKNOWLEDGMENT

We thank Charles Sereno for his helpful comments.

Reading and the split fovea

Richard Shillcock[a], Scott McDonald[b], Padraic Monaghan[c]
[a]*Department of Psychology and Division of Informatics, University of Edinburgh, Edinburgh, EH8 9LW, Scotland;* [b]*Psychology Department, University of Edinburgh, Edinburgh, EH8 9JZ, Scotland;* [c]*Department of Psychology, University of Warwick, Coventry CV4 7AL, United Kingdom.*
rcs@inf.ed.ac.uk http://www.iccs.informatics.ed.ac.uk/~rcs
scottm@cogsci.ed.ac.uk
http://www.iccs.informatics.ed.ac.uk/~scottm
padraic.monaghan@warwick.ac.uk
http://www.warwick.ac.uk/~pssbab

Abstract: We argue that models of reading should be based on anatomical reality, namely, the fact that both eyes are used in reading; and the observation that the human fovea is precisely vertically split, and projects each half of a fixated word to the contralateral hemisphere.

Although E-Z Reader very effectively organizes a complex literature and has the advantages of an implemented model, we take issue with it for being grounded at a functional level of description. The best starting point for modeling reading and visual word recognition is with the anatomical givens. In a series of statistical and connectionist simulations, we have shown the productivity of incorporating the fact that the human fovea is precisely vertically split and initially projects each part of a fixated word to the contralateral hemisphere (Monaghan et al. 2003; Shillcock et al. 2000). When this anatomical fact is incorporated into a computational model of the reading of isolated words, a range of behaviors emerges – from those core behaviors captured by the different developments of Seidenberg and McClelland's (1989) triangle model (with various versions of its abstract orthographic input), through to behaviors that are either problematic for the triangle model (such as Coltheart & Rastle's [1994] demonstration of apparent serial processing in word naming), or are currently outside its scope (such as laterality effects in reading). We argue that the initial division of a fixated word between the two hemispheres goes on to condition the normal reading of that word; there is no early, seamless integration of the contents of the visual cortex in the two hemispheres, as implicitly assumed in most models of visual word recognition. Reichle et al.'s Figures 13 and 14 suggest a unidirectional, cascaded flow of information processing, from lower/peripheral levels to higher/central levels, consequently transcending the initial splitting of the word. However, recurrent connectivity is pervasive in the brain; the fixated word is precisely split between the two primary visual cortices, and it is this representation that provides the reader with the most authoritative record of what is on the retina.

Visual word recognition is central to the reading of text. We suggest that a model of eye-movement control in reading is best built on the same anatomical foundation, so that there is a degree of autonomous representation and processing of the information in the two hemifields. This approach is in contrast to Reichle et al.'s assumption of a single window onto the text. Evolution has devoted only the posterior part of the human corpus callosum to a direct connection with visual information, compared with the whole of the corpus callosum for lower primates. It is reasonable to infer that interhemispheric coordination in the human brain is mediated substantially by information that is not primarily visual. For reading, this assumption means that there is an option of allowing the autonomous processing of visual information in each hemisphere, with interhemispheric communication based on some of the output of that processing, such as partial semantic activation of words. It also means that some of the information driving eye movements might be best captured in terms of the visual information initially available to each hemisphere, and its immediate consequences.

Some of our current research into characterizing the control of eye movements is based on definitions of the information available to each hemisphere, operationalized in terms of the probability of identifying the word(s) present in the half of the foveal window on either side of the fixation point. We suggest that building up a computational model from these anatomical foundations will demonstrate that some of E-Z Reader's architectural distinctions and some of its parameters will emerge from the interaction of the divided architecture with the informational structure of the task.

The divided architecture of the split-fovea model allows attentional processing to be implemented in terms of quasi-autonomous, contralaterally oriented opponent processors located in the respective hemispheres. This approach has a long history in the modeling of normal and impaired attention (see, e.g., Cohen et al. 1994; Kinsbourne 1970; 1977).

A further anatomical given is the fact that most reading happens with two eyes. (The single eye in Reichle et al.'s Fig. 3 is labelled "Eyes.") The two eyes do not always seem to focus on exactly the same point in the text, with the disparity sometimes being non-trivial (e.g., 1–3 characters); furthermore, monocular occlusion slows reading down (Heller & Radach 1999). Perception is robust in the face of binocular disparity: Minimally different images are fused; more disparate images rival each other and one is suppressed. Both of these processes seem to occur similarly in each hemisphere (O'Shea & Corballis 2001), but there is also the possibility that binocular rivalry has at least a higher-level component; for example, it is possible for complementary checkerboard pictures to be resolved into single, coherent pictures when they are presented to different eyes (Ngo et al. 2000). Clearly, the resolution of the fact that the two eyes may often have different fixations is important to a model of reading, if speed of reading can be affected; there is every indication that the role of the two eyes can be incorporated into a model based on foveal splitting.

ACKNOWLEDGMENTS
The authors are supported, respectively, by ESRC fellowship R/000/27/1244, Wellcome Trust grant GR064240AIA, and a Human Frontiers Science Program grant.

The effects of frequency and predictability on eye fixations in reading: An evaluation of the E-Z Reader model

Laurent Sparrow, Sébastien Miellet, and Yann Coello
URECA–UPRES EA 1059, Université Charles de Gaulle, F.59653 Villeneuve d'Ascq, France. **sparrow@univ-lille3.fr miellet@univ-lille3.fr coello@univ-lille3.fr http://www.univ-lille3.fr/ureca**

Abstract: We tested whether the E-Z Reader model can be generalised to the French language. The simulation showed that the model can account for the frequency effect. The predictability effect is moreover accurate for word skipping, but not for fixation times. We think that this model is psychologically plausible for certain aspects of reading and we have used it to evaluate the performance of dyslexic readers.

The E-Z Reader 7 model was applied to data that were obtained in an eye-tracking experiment. The purpose of the study was twofold. Its first aim was to test whether the model can be generalised to the French language. The second objective was to test whether the model can satisfy to contextual constraint influencing reading, namely, predictability. Fifteen participants were requested to read a French text that contained 134 words varying in frequency and predictability (the latter factor was determined in a previous task by asking participants to guess word$_{n+1}$ when provided with the sentence up to word$_n$). As in Reichle et al., the words were divided into five frequency classes and eye movement was recorded. Furthermore, we ran a simulation to determine how well the model could predict the observed data (i.e., means duration and distributions of fixations). To be in agreement with E-Z Reader 7, the first and last words of each sentence were not included in the data analysis.

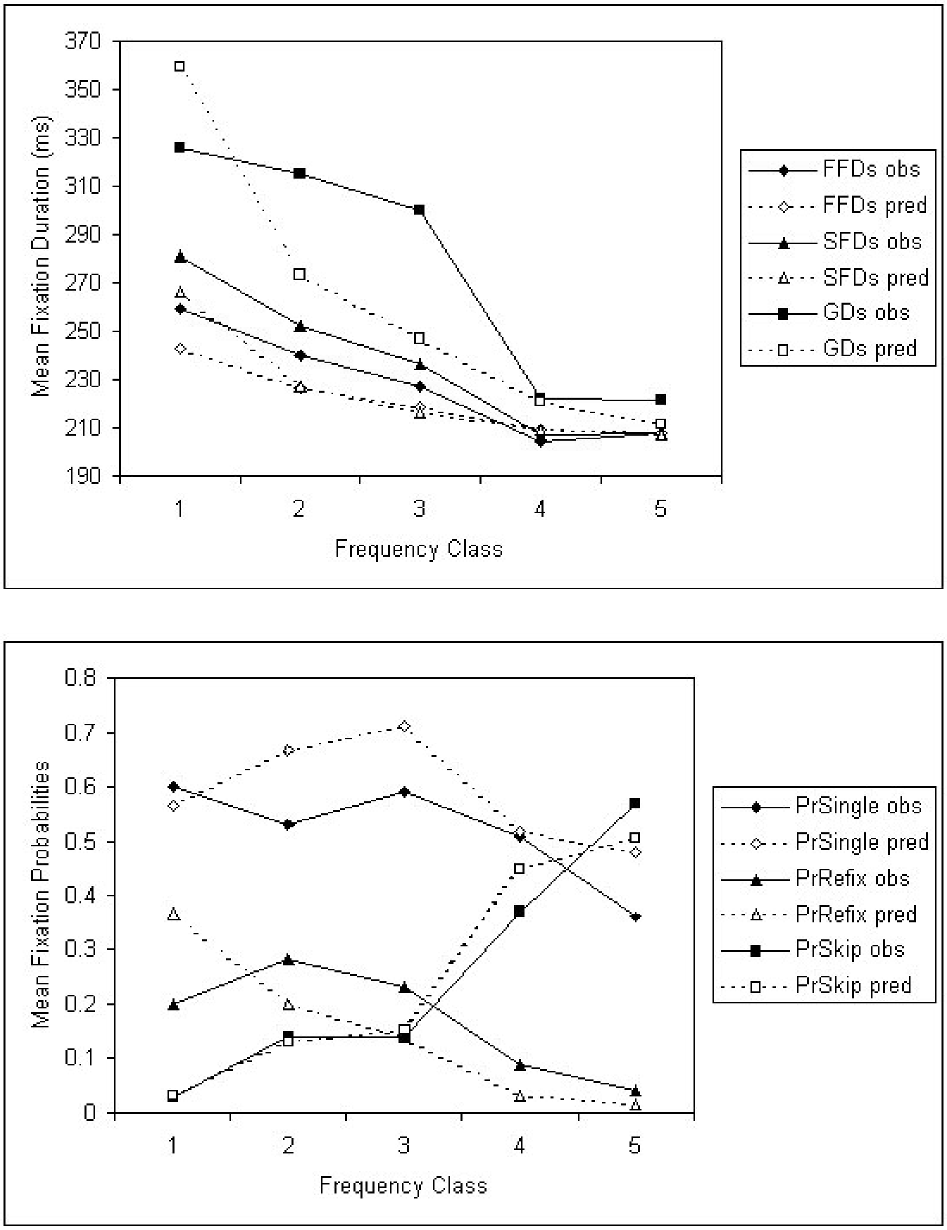

Figure 1 (Sparrow et al). The top panel shows the mean observed (obs) and predicted (pred) first-fixation (FFD), single-fixation (SFD), and gaze duration (GD) for five frequency classes of words. The bottom panel shows the mean observed and predicted single-fixation (PrSingle), refixation (PrRefix), and skipping probabilities (PrSkip) for five frequency classes of words (RMSD=0.452).

As shown in Figure 1, we obtained a satisfactory matching between observed and predicted values for all dependent variables. E-Z Reader 7 correctly predicted the negative monotonic relationship between first fixation duration and word frequency. However, the predicted values were slightly smaller than the observed values for gaze duration. But gaze duration, which includes refixation, is considered a global indicator of word processing as it is influenced by postlexical processing, although first fixation is supposed to be more sensitive to low-level word processing. Contrary to the reference experiment previously used to test the model (Schilling et al. 1998), a crucial aspect of our study was that subjects read a meaningful text, which suggests that postlexical processings were more deeply probed. In the Schilling et al. study, subjects read sentences that tended to be short, stereotyped expressions that were very simple to comprehend; and even though reading was involved with this paradigm, it did not obviously draw on the kinds of complex comprehension processes that are part of full-blown reading. However, it is noteworthy that the E-Z Reader model provides a theoretical framework for understanding the interface between vision and low-level aspects of language processing.

In addition to the frequency effect, an important finding was the predictability effect on eye movement. Generally, the effect of predictability is not as high as the frequency effect on fixation duration but is pronounced on word skipping. In our study, readers skipped 31% of predictable words compared with 22% of the unpredictable words. The predictions of E-Z Reader were very close to those values: 37% of predictable words and 25% of unpredictable words were skipped. The difference between predictable and unpredictable words (9% in our study and 12% for E-Z Reader) is quite consistent with prior research (Rayner & Well 1996; Rayner et al. 2001). However, inconsistency was observed

Table 1 (Sparrow et al.). *Mean observed (obs) and predicted (pred) first-fixation (FFD), single-fixation (SFD), and gaze duration (GD), in milliseconds, for predictable and unpredictable words*

	FFDs		SFDs		GDs	
Predictability	Obs	pred	obs	pred	obs	pred
Low	226	213	232	212	280	240
High	207	209	212	209	235	223
Difference	19	4	20	3	45	17

in the pattern of results for fixation times (Table 1). E-Z Reader did not predict the effect of predictability on first fixation duration and single fixation. With respect to gaze duration, the effect predicted by E-Z Reader was comparable in size to the effect obtained in other studies (Rayner & Well 1996; Rayner et al. 2001), though it was lower than in our study.

A closer examination of the data indicated that the prediction of E-Z Reader was reversed for high frequency words: Fixation duration was longer for predictable than for unpredictable words (Table 2). For low-frequency words, the prediction of E-Z Reader appears to be larger for first and single fixations duration. However, the effect predicted by E-Z Reader on gaze duration was comparable to the effect obtained in our study.

This difference in the pattern of results for fixation duration versus word skipping can be accounted for by the different mechanisms that might be involved with regard to the decision about when and where to move the eye. With E-Z Reader, the time required to complete the first stage of lexical access (i.e., when to move the eye) is principally a function of word frequency and a free parameter (theta; see Equation 1 in the target article) reduces the extent to which the predictability of a word attenuates the lexical processing time. Our data suggests that, in normal reading, predictability can play a more important role during the first stage of lexical access (i.e., L_1).

Finally, the latest version of E-Z Reader appears to be psychologically plausible and gives an accurate account of various phenomena in reading. It is, however, incomplete, as it preferentially takes into account "low-level" aspects of the reading process. This model nonetheless provides a valuable analytical tool to examine some key assumptions about eye-movement and language processing. As an example, we used the model to simulate how individual differences would affect the pattern of eye movements in reading. For this purpose, we compared observed eye movements of dyslexic subjects with the E-Z Reader-predicted data. The observed and predicted values were very close for the duration of first fixation (224 msec and 213 msec respectively for dyslexic subjects and E-Z Reader). However, gaze duration was considerably longer for dyslexics (384 msec) than for E-Z Reader (256 msec). This pattern of results can suggest that later stages of lexical access were impaired in dyslexic subjects, but not the low-level aspect of reading process. Of course, further investigations are required to corroborate this conclusion.

Table 2 (Sparrow et al.). *Effect of predictability for low frequency (LF) and high frequency (HF) words (in msec.)*

	FFDs		SFDs		GDs	
	obs	pred	Obs	pred	obs	pred
LF	15	29	15	29	55	48
HF	22	−23	24	−23	35	−14

Where to look next? The missing landing position effect

Geoffrey Underwood

School of Psychology, University of Nottingham, Nottingham NG7 2RD, United Kingdom. **geoff.underwood@nottingham.ac.uk**

Abstract: The E-Z Reader 7 model is powerful but incomplete. When programming the saccade to the next word, we take into account the familiarity of the letter sequences at the beginning of that word. This landing position effect is well established, but is neglected in the model. A possible locus for the effect is suggested within the E-Z Reader framework.

The E-Z Reader model is clearly our most advanced model of eye-movement control in reading, and this version of the model is particularly welcome as a refinement of an already powerful description. The success with which the model accounts for variations in fixation durations and locations is impressive. The model is incomplete, however, and, like any model, it cannot be considered to be accurate unless it takes account of the full range of phenomena that can be observed. The starting point in the model-building enterprise is to know what it is that is being modeled – what are the phenomena to be taken into account? The E-Z Reader 7 model is impressive in its predictions about an extensive range of effects, but it fails to make predictions about the *landing position effect.* Although Reichle et al. are aware of this effect, their mention of it in the target article is restricted to Note 5. It is unclear whether this is because the authors do not regard it as an effect that requires explanation, or because they cannot see how the model could explain it.

There are two classes of *landing position effects* – the phenomena that reflect the output of the system that decides where we should look next when reading; and Reichle et al. account for the effects of (1) low-level visual factors, such as word length, comfortably. Another class of effects is (2) the influence of the distribution of information within the word that is to be fixated; and this is the effect that Reichle et al. appear to be sweeping under the carpet. Words like *awkward* and *coyote* can be described as having informative beginnings, because they start with uncommon trigrams. In contrast, the words *author* and *compact* start with trigrams that are shared with large numbers of other words. A number of published studies have demonstrated that the reader's eyes land slightly closer to the beginning of the word if the word starts with an uncommon trigram (Beauvillain & Doré 1998; Beauvillain et al. 1996; Everatt & Underwood 1992; Everatt et al. 1998; Hyönä 1995b; Hyönä et al. 1989; Inhoff et al. 1996; Underwood et al. 1990; Vonk et al. 2000). Oculomotor programming is influenced by the information in the word that is the target of the next saccade. This landing position effect has now been demonstrated with a number of different languages and is undoubtedly real, albeit small. Information within the word modulates the tendency to fixate slightly to the left of the centre of a word.

There is a question about the type of information that causes the modulation in the landing position. When we first reported the effect, we considered all options, including the possibility that it was lexical or morphemic information from the beginning of the word that was responsible (e.g. Hyönä et al. 1989; Underwood et al. 1990); but it is now clear that it is the orthographic content of the first few letters that is important. Note 5 of the target article acknowledges this debate but fails to incorporate the effect in the data set that the model should take into account. One reason for this might be that there is a question about the reliability of the effect. Indeed, there are reports of studies that have failed to find the effect for all of the words tested (Liversedge & Underwood 1998; Rayner & Morris 1992; Underwood et al. 1989). This inconsistency is possibly attributable to variations in our definitions of an informative or redundant beginning to a word (bigram/trigram frequency vs. predictability by a cloze task), or to variations in cognitive load imposed by the specific reading task, because we know that foveal load influences the effectiveness of parafoveal information (Henderson & Ferreira 1990), or to variations in the reading skill of the participants being tested (see Everatt et al. 1998). The inconsistency of an effect simply means that we have not yet determined the conditions in which it will appear. Not all words are skipped, and not all short words are skipped all of the time, but there is no suggestion that we should ignore this effect because it does not appear with total predictability. The same should hold for the landing position effect.

The conditions for an information-based landing position effect are that (i) oculomotor programming can be modulated by visually available information, and that (ii) orthographic information can be extracted from words currently in parafoveal vision. The evidence supporting both of these conditions is well established. First, landing positions are sensitive to word length, and some short words sometimes receive no fixation at all (target article, sect. 2.8.1). Second, the parafoveal preview effect has been demonstrated for orthographic information (sect. 2.6). The E-Z Reader model accounts very well for these demonstrations of modulation, and when they are viewed together they make plausible the modulation of the landing position by orthographic information.

The landing position effect is plausible on the basis of lability of oculomotor programming and on the basis of the parafoveal processing of orthographic information. It has been demonstrated in a number of experiments that have used a number of alphabetic languages. How then might the E-Z Reader model be developed to account for it? One possibility is that the first stage of word identification (L_1 in Fig. 3), which in earlier E-Z Reader models was described as performing a familiarity check, could identify predictable sequences of letters at the beginning of word$_{n+1}$. If a highly unfamiliar bigram or trigram started word$_{n+1}$, then a shorter saccade may be programmed. If a predictable sequence were detected, then the signal to the oculomotor system would be to start programming a saccade of increased amplitude. And if the word contained a predictable letter sequence and early visual processing had recognised that it was a short word, then the saccade would skip the word altogether. The model could incorporate this effect and make predictions that would help describe it further. The first task of all model-builders is to identify the essential evidence, and here the E-Z Reader model does not so much fall over as turn its back on the data.

The basic assumptions of E-Z Reader are not well-founded

Françoise Vitu

CNRS, Université René Descartes, Laboratoire de Psychologie Expérimentale, 92774 Boulogne-Billancourt, Cedex, France.
francoise.vitu@psycho.univ-paris5.fr
http://lpe.psycho.univ-paris5.fr/page%20perso/Vitu-Thibault/FEVitu.htm

Abstract: This commentary focuses on the two basic assumptions of the E-Z Reader model, discussing the possibility that adjacent words in reading may be processed in parallel rather than serially, and presenting evidence against a cognitive control of eye guidance in reading.

Like several recent models of eye guidance in reading, the E-Z Reader model provides a fairly good account of the variability of eye behavior that characterizes reading. From a psychological perspective, however, the performance of a computational model, or how well simulations fit behavioral observations, is not as critical as the reality of the mechanisms and the processes that lead to such performance. From that perspective, E-Z Reader may not be a serious candidate. In this commentary, I present several findings that challenge the two basic assumptions of the model, suggesting, first, that reading may not proceed through sequential attention shifts, and second, that ongoing processing is not the main driving force behind eye guidance.

The first basic assumption of E-Z Reader posits that for the temporal order of the words to be preserved in reading, adjacent words must be processed serially through covert shifts of attention from the fixated word to the parafoveal word(s). This assumption cannot, however, be asserted based on current empirical evidence. First, as noted by Schroyens et al. (1999), the negative influence of the foveal processing load on parafoveal preview benefit is in no way proof that processing of the parafoveal word starts only when the fixated word is identified. As the difficulty of the foveal word increases, a tendency for greater spillover effects in the visible compared to the masked preview condition can be observed, which suggests that the effect may simply originate from an interaction between processing associated with the foveal and parafoveal words. Furthermore, parafoveal preview benefit becomes greater rather than smaller as the fixation duration increases, thereby indicating that parafoveal processing lasts for the whole fixation time, rather than being limited to the delay that remains from the moment the fixated word is identified until the saccade is ready to go.

On the other hand, the fact noted by the authors, that processing of the foveal word remains unaffected in most instances by the lexical and semantic properties of the parafoveal word, does not necessarily mean that adjacent words are processed in sequence. In the framework of a pure parallel hypothesis that makes no recourse to the notion of attention, the presence of parafoveal-on-foveal effects is indeed not obligatory, but rather is conditional on the respective time course of processing associated with both foveal and parafoveal words (see Vitu et al., in press). As more letters from the parafoveal word fall into a region of lower visual acuity, parafoveal processing is relatively slow compared with that associated with the fixated word; and the likelihood that it influences processing associated with the fixated word may be relatively low, unless the word is very easy to process.

Parallel processing therefore remains a serious alternative to the sequential attention shift assumption, and, given that it relies strongly on visual acuity, it may as well preserve the temporal order of the words. Now, it remains an open question whether keeping word order straight is critical to reading. Indeed, several studies suggest that the eye-movement pattern does not always respect word order. This happens, for instance, when a word that is not yet identified is initially skipped. In that case, the word will be fixated only after the execution of a regressive saccade from one of the following words (Vitu & McConkie 2000; Vitu et al. 1998). These particularities of the eye-movement pattern, which are actually not accounted for by E-Z Reader, bring us to the second assumption.

The second basic assumption of E-Z Reader states that ongoing processing is the main driving force behind eye guidance. Ongoing processing would determine when the eyes move, and which word to send the eyes to, while visuomotor processes would produce systematic errors in sending the eyes to specific locations. Three major objections can be raised against this notion. First, the duration of individual fixations, although being correlated with the difficulty of processing associated with the encountered words, is also strongly influenced by the fixation position, with fixation duration being greatest when the eyes are at the center of words (Vitu et al. 2001). This phenomenon, which is opposed to Rayner et al.'s (1996) finding, is very robust since it was found in three different corpora of eye-movement data (based on a total of 153,855 fixations). Besides being interesting, it seriously questions the cognitive assumption, as it is about twice as large as the effect attributable to word frequency and as it cannot be interpreted in terms of ongoing processing. Because a word is most easily identified when the eyes initially fixate its middle, fixation duration should be shortest rather than longest at that position.

A second objection to the cognitive control hypothesis relates to the postulate that ongoing processing determines which word to send the eyes to. A recent review confirms that the frequency and the predictability of the words influence the likelihood of word skipping (Brysbaert & Vitu 1998). However, it also indicates that these effects are rather rare, and that when they occur they are smaller than the effects of both word length and launch site. As words get shorter, and the eyes are launched from closer to the beginning of the words, skipping rate increases drastically up to about 90% in the most extreme conditions. Undoubtedly, shorter and closer words are more likely to be identified in parafoveal vision, but this is not what causes them to be skipped more often. Indeed, a comparison between text reading and the scanning of meaningless material (or z-letter strings) reveals that string length and launch site influence string-skipping in the same manner as word length and launch site influence word skipping (Vitu et al. 1995). Such great regularities in the pattern of eye movements – present even when no cognitive processing is involved – strongly suggest that the decision of which word to fixate next is not cognitively determined, but rather, is primarily a function of low-level visuomotor processes.

As a final objection, I question the need to distinguish between two separate *Where* mechanisms, one that selects the next saccade target word and another sending the eyes at about the required location. Given that word length and launch site similarly influence both skipping rate and initial landing sites in words, a single mechanism may determine where the eyes move next. Furthermore, since visuomotor influences strongly deviate the eyes from the center of words, there may be no need to posit that they aim for the center of preselected words. Our assumption is that forward eye movements in reading are not directed to specific target locations, as the eyes would be naturally pulled by peripheral visual information in a center-of-gravity-type manner (see Vitu 1991a). Ongoing word identification processes would only intervene occasionally to modulate the oculomotor scanning pattern, favoring either the lengthening of planned saccades (see Lavigne et al. 2000), or the execution of regressive saccades to specific word locations (see Vitu & McConkie 1998; Vitu et al. 2000).

Authors' Response

Eye movements in reading: Models and data

Keith Rayner[a], Alexander Pollatsek[a], and Erik D. Reichle[b]

[a]*Department of Psychology, University of Massachusetts, Amherst, Amherst, MA 01003;* [b]*Department of Psychology, University of Pittsburgh, Pittsburgh, PA 15260.* **rayner@psych.umass.edu** **pollatsek@psych.umass.edu** **reichle@pitt.edu**

Abstract: The issues the commentators have raised and which we address, include: the debate over how attention is allocated during reading; our distinction between early and late stages of lexical processing; our assumptions about saccadic programming; the determinants of skipping and refixations; and the role that higher-level linguistic processing may play in influencing eye movements during reading. In addition, we provide a discussion of model development and principles for evaluating and comparing models. Although we acknowledge that E-Z Reader is incomplete, we maintain that it provides a good framework for systematically trying to understand how the cognitive, perceptual, and motor systems influence the eyes during reading.

R1. Introduction

Before we delve into specific issues and controversies about the issues raised by the commentaries, it might be of some value to try to characterize what we had hoped to accomplish with our model and thus indicate what we feel is the proper context in which it is best discussed and evaluated. Perhaps a good place to start is with a quote from E. B. Huey in the introduction to his classic book on the psychology of reading:

> And so to completely analyze what we do when we read would almost be the acme of a psychologist's achievements, for it would be to describe very many of the most intricate workings of the human mind, as well as to unravel the tangled story of the most remarkable specific performance that civilization has learned in all its history. (Huey 1908)

We hope no one misunderstood our modeling effort as pretending to be an attempt "to completely analyze what we do when we read." We indeed agree with Huey that to completely analyze what we do when we read would be a monumental achievement well beyond our efforts – and certainly well beyond what our field has learned, to date, about reading, the control of eye movements, and language processing.

Instead, we had hoped to develop a model that was a reasonable "zero-order approximation" of how cognitive processes control the eyes in reading. Our first goal was to come up with a model that could explain how ongoing cognitive processes *could* control eye movements during reading, subject to the constraints that it do so in real time and be plausible in terms of what is known about the underlying cognitive processes and the control of eye movements. We think we succeeded at this goal, and that this was a far from trivial task.

Our second goal was to make the model as simple as possible, so that it would be reasonably transparent with respect to how and why it worked. If we succeeded at this goal, then when the model failed to account for a phenomenon properly, it would be clear why the model did not work and, therefore, whether it was merely too simplistic and needed to be further elaborated, or whether it was fundamentally flawed. Put differently, if a model is simple enough, it can be a useful heuristic device, but when it gets too complex, it risks merely being a data-fitting algorithm. Put a third way, if the model is simple enough, it has the chance of generating both testable and interesting predictions about reading (i.e., predictions that can be falsified and that test basic assumptions of how the model explains the reading process).

Consistent with these goals, we tried to delimit our problem carefully so that we would have some chance of success. Most notably, we did not try to explain interword regressions (i.e., regressions to words that had been either previously read or skipped over). The reason for this is that there is ample evidence that many of these regressions are a result of various "higher-level" processing failures (e.g., attributable to general comprehension problems or temporary misparsing of sentences). Another reason is that modeling these kinds of regressions in any detail is several orders of magnitude more difficult than what we have attempted, both because we have a far from adequate cognitive theory of the processes that fail, and because it is quite unclear – given such a failure – exactly what should be targeted by an eye movement responding to this failure. We also did not pretend to have a serious or detailed model of word identification, both because how words are encoded is still a contentious issue that we hoped to bypass, and because a model of word identification (even without worrying about eye movements) that attempted to explain all the phenomena of word encoding that have been studied to date would have required an effort at least as large as our present modeling effort. Instead, we merely inserted a functional account of word encoding into our model that had reasonable processing characteristics.

In the prior paragraph, we have noted that a major goal of our model was to explain how cognitive processes can control eye movements in normal silent reading. Some of the commentaries, however, appear to take us to task for this, their essential claim being that cognitive processes have minimal control over eye movements in reading. This criticism is interestingly in contrast with several other commentaries that argue our model is quite limited, in that there are phenomena in reading indicating cognitive control over eye movements that the model does not account for (and that the commentators are skeptical can be accounted for by a model such as ours).

We will fill in the details below, but we remain convinced that our model is on the right track, in that it gives a good account for much of the forward progress of the eyes through the text and that this progression is importantly guided by cognitive control. Nonetheless, as we have indicated above and in the target article, our model is definitely incomplete, and we are trying to elaborate on the details to develop a more comprehensive model (e.g., we are currently trying to account for how subword units like morphemes can influence eye movements).

We wish to close this section with two disclosures. The first is that there was a conceptual oversight in the E-Z Reader 7 model presented in the target article: We assumed that the early visual processing stage does not have to be reinitiated when a previewed word is subsequently fixated. This assumption is clearly wrong, as it goes against the large amount of evidence that the integration of information across saccades in reading is *not* at the "visual" level (Rayner 1998). It also causes the model to predict far too much preview benefit, a fact that alerted us to the problem and that was made apparent to us by the excellent analysis in the **Findlay & White** commentary. A subsequent simulation in which the initial stage of processing had to be repeated on every fixation led to fits that were virtually as good as the original simulation; and furthermore, reasonable estimates of preview benefit emerged from this simulation. The second point is that our description of the model led some commentators to believe that we thought that the targeting of saccades was, *in principle,* completely controlled by word boundary information (i.e., the location of spaces between words). In keeping with our attempts to be parsimonious, we did think that such an assumption was not unreasonable in explaining the phenomena discussed in the target article. However, in our first attempts to model the processing of Finnish compound words (Pollatsek et al. 2003), it soon became apparent that more "cognitive" control was needed to explain the process of targeting refixations on these long words. In the next few sections of this reply, we will address what we see as the central issues that were raised by the commentaries.

R2. Parallel versus serial allocation of attention

Our assumption that attention is allocated serially should, like virtually all the assumptions in our model, be taken as a working hypothesis. However, it's an assumption we are reluctant to part with unless very strong evidence forces us to give it up. A major reason for wanting to hold onto it is that giving up this assumption would make the model significantly less transparent – and therefore much less useful as a heuristic device. Our general working assumption is that most people studying reading are interested in the cognitive processes underlying this activity. Many of these researchers, however, do not use the method of recording eye movements of people who are reading text, but instead rely on relatively artificial single-word paradigms (such as the lexical decision task) or more contrived reading paradigms (such as having the reader advance the text by pushing buttons). One of the reasons researchers resort to such techniques is economic: Doing good eye-movement research is expensive. However, it seems clear that another reason people shy away from engaging in eye-movement research is that they are afraid that the relationship between the cognitive processes they wish to study and the eye-movement record is hopelessly complex. As a result, a major motivation for developing the E-Z Reader model was to create a tool that would help demystify eye-movement measures and provide a reasonable rationale for their interpretation in the study of how readers comprehend written language.

Clearly, one aspect of the eye-movement record that our model is building on is the reasonably clear relation between the variables that indicate cognitive processing of language (e.g., the frequency of a word) and the fixation time on that word. (We discuss this relationship in detail in sect. R3, on lexical processing.) However, it seems equally clear that one can't merely assume that the time spent on a word is a direct measure of the time to process that word. First, the onset of the first fixation on a word can't coincide with the start of processing that word because – among other things – there is clear evidence that the processing of many words begins before they are fixated (Rayner 1998). Second, unless one assumes that eye movements are programmed instantaneously, the end of a fixation can't coincide with the termination of processing of that word. However, our assumption that words are processed serially allows one to draw the inference that the gaze duration on a word can be inferred to be the duration of a stage of word processing (the one that triggers the eye movement off the word). Therefore, our model does provide a good rationale for taking gaze durations on words seriously, and making reasonably simple inferences from them.

This inference rests quite heavily on the assumption that attention moves serially. Can we guarantee that this assumption is true? Clearly, we cannot. However, there is a reasonable amount of evidence in the attention literature (see, e.g., Sieroff et al. 1988) to suggest that attention is importantly object-based, and that – at least in certain cases – it moves in a serial fashion from object to object. Moreover, much of the evidence against this serial assumption in reading (see sect. R4, on parafoveal-on-foveal effects) seems quite questionable. We therefore see no reason to change this admittedly strong assumption about attention.

We should also point out that, from a modeling standpoint, it would be easier to adopt the view of attention espoused in the SWIFT model (Engbert et al. 2002): In this model, attention is a gradient that is allocated in parallel to several words, so that, instead of positing two stages of word identification (as in E-Z Reader), the model has many stages of word processing. In essence, an assumption that words are processed in parallel allows many more degrees of freedom in fitting the data.

We would like to make one more point about our serial-processing assumption. **Kennedy** takes us to task for our suggestion that this assumption is a good one because readers "need to keep word order straight." He then goes on to argue that written language, unlike spoken language, "can be inspected at will," so that text can be reprocessed. (This is presumably necessary to correct any initial misencodings of word order, among other things.) We certainly did not intend to argue that it was absolutely necessary for readers to correctly encode word order on their first pass to have any chance of understanding what was said in the text. However, we find Kennedy's argument a bit puzzling. Surely, he would agree that it benefits the reader quite a bit to get the word order correct on the first pass through the text, and that, if words are processed serially, this then helps readers encode words in their correct order. (Of course, even if the original encoding of order was correct, as it would be in a serial-encoding model, there could still be later memory errors.) We think an interesting question that Kennedy's comment raises, however, is whether there might be significant differences between languages in this regard. That is, word order is particularly important in conveying information in English because there is little in the way of case marking; so perhaps one might find that readers do attempt to process more than one word at a time in heavily case-marked languages, such as Russian.

R3. A more detailed model of lexical processing

As **Huestegge, Grainger & Radach** (**Heustegge et al.**) indicate, there are two general strategies that one might use to try to understand the relationship between word identification and eye movements during reading. The first is to modify an existing model of word identification so that it includes those assumptions necessary to explain how this process affects (and is affected by) the movement of the eyes through text. The obvious advantage of this approach is that it can potentially explain phenomena that have been observed in both single-word tasks (e.g., frequency effects in lexical decision) and natural reading. The drawback of this approach is that the resulting models may become so complex that they are difficult to evaluate; as Huestegge et al. point out, the model that has come closest to realizing this approach (the Glenmore model) has not yet been evaluated on a sentence corpus.

The second modeling strategy (which, as we noted previously, is the one that we have adopted) is to develop a minimal set of assumptions about word identification and eye-movement control, and then use these assumptions as a framework for examining both cognitive processes and eye-movement control in reading. The main virtue of this minimalist approach is that the inner "workings" of this type of model are transparent enough to serve as heuristics for thinking about the reading process and simple enough to be tested on large sentence corpora. The major drawback of this approach is that the models of lexical processing and eye-movement control are likely to be oversimplified and/or incomplete.

We agree with many of the commentators that our explanation of word identification is not fully specified. For example, **Andrews** and **Perea & Carreiras** correctly note that our model cannot handle orthographic neighborhood effects. Similarly, we agree with **Hyönä & Bertram** that our attempts to simulate the complex effects of morphology (Pollatsek et al. 2003) are preliminary and cannot account for the full range of effects that have been observed (e.g., Bertram & Hyönä 2003). Further, we agree with **Beauvillain & Pouget**, **Liversedge & White**, and **Underwood** that the inclusion of a preattentive visual processing stage in E-Z Reader does not (at present) provide an explanation for how the characteristics of a word's orthography affect where it is fixated (however, Underwood does suggest how our model might address this issue).

Clearly, our assumption that the speed of (both stages of) word processing can be completely accounted for by the written frequency of a word in the lexicon and its predictability from the prior text was intended to be a convenient initial step in attempting to relate cognitive processes in reading to the movement of the eyes through text. That is, when we proposed the model, we were fully aware that this assumption was an oversimplification as, among other things, we were involved in several of the studies that indicated that word parts, as well as word wholes, influence the time a word is fixated in text. However, we think the way to proceed is to consider each of these phenomena carefully and then to think about the minimal assumptions necessary to explain each of them. The other approach, which often involves using two or more processing modules operating in parallel (e.g., Reilly & Radach's [2003] Glenore model), may successfully predict the effects, but often not in a manner that suggests new or interesting phenomena, or in a manner that can be "unpacked" so as to reveal the assumptions that are responsible for making the predictions.

Several of the commentaries were more specific and argued that our assumption that words are identified in two stages is both vague and unjustified (**Huestegge et al.**; **Sereno, O'Donnell & Sereno** [**Sereno et al.**]; and **Shillcock, McDonald & Monaghan** [**Shillcock et al.**]). That is, a core principle in E-Z Reader is that the processing necessary to identify printed words is done in two stages: L_1 and L_2. In our earlier papers (Reichle et al. 1998; 1999), we described L_1 as a "familiarity check" that indicated full identification of a word was imminent, and L_2 as the "completion of lexical access," or the point at which a word's meaning was available to the language system for further processing. This distinction is important, because in our model, the completion of L_1 causes the oculomotor system to begin programming a saccade, whereas the completion of L_2 causes attention to shift to the next word. **Andrews** raises the question of whether this change in labels reflects a deeper conceptual change. Our answer is that there are at least two ways the distinction between L_1 and L_2 can be conceptualized, and that our goal in introducing these new labels was to make our model agnostic with respect to these different conceptualizations.

In earlier papers (Reichle et al. 1998; Rayner et al. 1998c), we referred to the global-activation memory models (e.g., MINERVA 2; Hintzman 1984) and pointed out that, according to these models, cues such as a word's orthographic pattern can be used to gain access to additional information in memory in two ways. The first is by matching the cue to the collective contents of memory. This process provides a rapid index of the cue's overall familiarity and serves as a basis for recognition decisions (Hintzman 1987; 1988). It is also consistent with our early conceptualization of the familiarity check. The second way of accessing information in memory is via a slower pattern-completion process that provides a basis for recall (Hintzman 1986). For example, a word's letters might be used to "retrieve" its pronunciation and/or meaning. This second process is consistent with our early conceptualization of the completion of lexical access. This distinction between a rapid sense of familiarity and a less rapid "retrieval" process has a venerable history in cognitive psychology (*dual-process theory:* Atkinson & Juola 1973; 1974; Mandler 1980; see Yonelinas 2002 for a review) and has recently been supported by physiological evidence (Davachi et al. 2003).

The second manner in which the two stages of word identification can be conceptualized is in terms of the time course over which orthographic, phonological, and semantic information becomes available as a printed word is identified. We think it is still an open question both about how the phonological code is activated (e.g., the roles of "assembled" and "addressed" phonology) and whether orthographic information accrues appreciably faster than phonological information. That is, there is now research from our lab (Lee et al. 1999; Pollatsek et al. 1992; Rayner et al. 1995; 1998b) and other labs (Perfetti & Bell 1991) indicating that phonological information is extracted early in word encoding – even before the word is fixated. Therefore, we didn't intend to posit, as **Huestegge et al.** claim (see also **Sereno et al.**), that phonological encoding is restricted to the second stage of word identification. However, it seems clear there must be significant activation for orthographic or phonological codes (or both) prior to anything like full semantic activation. An alternative conceptualization of the familiarity check, therefore, is that it corresponds to the availability of a word's form (i.e., its spelling and/or sound), whereas the completion of lexical access corresponds with the availability of a word's meaning. (Yet another possibility is suggested by **Osaka**: that the first stage of word identification is mediated via the phonological loop component of working memory.)

Of course, these different ways of thinking about the two stages of word identification that are postulated in our model are not mutually exclusive. For example, a recently developed instance-based model of word identification (Reichle & Perfetti 2003) is based on MINERVA 2 (Hintzman 1984) and consequently inherits the dual-process distinction between fast familiarity and slow retrieval operations. Because the grapheme-to-phoneme mappings tend to be more consistent than the arbitrary mappings between spellings and meanings, however, word form information is also generally available more rapidly than semantic information.

Hence, our decision to use "L_1" and "L_2" rather than "familiarity check" and "completion of lexical access" reflects our desire to leave open the possibility that either (or both) of these conceptualizations of word identification may play important roles in eye-movement control. That is, the decision about when to initiate a saccadic program may reflect the overall "sense" that a word is familiar and hence likely to be soon identified, or it may reflect the availability of a word's form, or it may reflect both. We think that an interesting area of future study is to try to understand whether there is a consistent difference in how variables affect gaze duration (largely controlled by L_1 in our model) and how they affect "spillover" (largely controlled by L_2 in our model). If the pattern differs markedly, it would indicate that there may be two fairly discrete stages of word processing. If the pattern in the two types of effects is fairly similar, then it would suggest that L_1 (and the signal to move the eyes) may merely be a result of a lower threshold of word identification than L_2.

In this regard, we posited one distinction between L_1 and L_2: that predictability (as measured by cloze task probabilities) affects L_1 less than L_2. Our justification for this assumption was that higher-level linguistic processing (the effects of which are estimated by our predictability measure) should influence the earlier stage of lexical processing less because it is more data-driven. As **Raney** indicates, there is evidence consistent with this notion: Two recent studies from his lab suggest that higher-level linguistic variables can have isolated effects on L_2 (Raney et al. 1996; 2000). We contend that these results, in conjunction with our preliminary simulations exploring the loci of morpheme effects (Pollatsek et al. 2003) and lexical ambiguity effects (described in our target article), demonstrate the usefulness of our model and support Raney's conjecture that E-Z Reader "provides a transparent (i.e., definable) architecture for explaining how high-level language processes influence eye movements."

R4. Parafoveal-on-foveal effects

A number of the comments deal with so-called parafoveal-on-foveal effects, which is not surprising, given that it is currently a heated topic of debate. A "parafoveal-on-foveal effect" refers to a situation in which one would observe some effect on the fixation time on the fixated word (word$_n$) produced by a parafoveal word that has not yet been fixated (usually the word to the right of the fixated word, i.e., word$_{n+1}$). **Inhoff & Shindler**, **Kennedy**, **Murray**, and **Vitu** all discuss this issue and claim that such a phenomenon is quite damaging to the E-Z Reader model, which assumes that encoding of words is strictly serial. Before discussing the data in some detail, we think two theoretical points need to be made. The first is that the serial assumption about word encoding in E-Z Reader does not preclude the possibility that processing word$_{n+1}$ can affect the fixation time on word$_n$. In fact, a later section of our response (sect. R7) will deal with one such effect: the effect of skipping word$_{n+1}$ on the fixation time on word$_n$. Second, because our model predicts occasional saccadic errors, it also predicts that the eyes will occasionally land on word$_n$ in cases where the intended saccade target is word$_{n+1}$. Such targeting errors allow for some parafoveal-on-foveal effects, although these will be admittedly small, as the percent of such mistargeted saccades will be small.

Now let us consider whether there is substantial evidence that there are parafoveal-on-foveal effects that would require a substantial change in our modeling approach. First, let us consider whether there are consistent effects on the fixation times on word$_n$ (i.e., the fixated word) attributable to the processing of the meaning of the word to the right of fixation (word$_{n+1}$). If so, one would expect that the frequency of word$_{n+1}$ would have an effect, but it does not (Carpenter & Just 1983; Henderson & Ferreira 1993; Rayner et al. 1998c). There have been other studies that have manipulated various properties of word$_{n+1}$, but it is far from clear that these effects can be reliably obtained (see Rayner et al. 2003c). For example, White and Liversedge (in press) found no such effect. Furthermore, Hyönä and Bertram (in press) reported five different studies in which they looked for such effects and found rather inconsistent evidence for them. In their experiments (which used long Finnish words), they varied the frequency of the first constituent or the frequency of the whole word. In one experiment (which varied the frequency of the first constituent), they found that the gaze durations were slightly longer when the first constituent of the word to the right of fixation was low frequency. Yet, in another experiment with exactly the same manipulation, they found the opposite effect. There were also no effects on gaze duration in the experiments in which whole word frequency was manipulated, but they did find an effect on the final fixation before moving to the target word. Likewise, the other measures that were examined (skipping rate and probability of a refixation) yielded variable and inconsistent effects. Clearly, beauty is in the eye of the beholder: Hyönä and Bertram argued that parafoveal-on-foveal effects can occur under some circumstances, but, given that most of their measures yielded no effect, a skeptic could argue that such effects are highly unreliable and seldom if ever occur.

To their credit, Kennedy et al. (2002) attempted to account for discrepancies in parafoveal-on-foveal studies in a systematic manner. However, one of the key experiments has been somewhat difficult to replicate (even by the lab that first reported the effect; see **Murray**'s commentary), and the fact that similar manipulations can yield either facilitation or inhibition is somewhat troubling. Another problem is that much of the evidence for parafoveal-on-foveal effects comes from tasks that may approximate reading, but aren't reading. Instead, subjects engage in what seem to be closer to visual-search or pattern-matching tasks than reading. Because it is known that frequency effects disappear when the task changes from reading to visual search (Rayner & Fischer 1996; Rayner & Raney 1996), one has to be concerned about whether these parafoveal-on-foveal results (even if they are reliable) can generalize from these nonreading tasks to reading. A final problem is that most of the experiments that have claimed to show such parafoveal-on-foveal effects in reading have manipulated some orthographic aspect of the text (e.g., used highly unusual and/or orthographically illegal letter combinations at the beginning of word$_{n+1}$). In some cases, these manipulations are highly visible and may disrupt the normal reading process (e.g., the reader might conclude that part of the task is to notice these strange letter combinations). We do acknowledge that there are a few studies in which subjects were reading for meaning and semantic parafoveal-on-foveal effects were observed. For example, Inhoff et al. (2002) reported such an effect. However, other similar studies from the same lab (Inhoff et al. 2000a; Starr & Inhoff, in press) did not yield these semantic effects; instead, the only effects were attributable to orthography.

Given these various points, we continue to feel that the validity of parafoveal-on-foveal effects remains questionable, and despite **Murray**'s protestations, we continue to think that proponents of such effects need to demonstrate that they are relevant to the task of normal reading (i.e., reading text for meaning). We do agree, however, that, if large consistent effects of the *meaning* of word$_{n+1}$ on the fixation duration on word$_n$ could be demonstrated, then such a demonstration would be problematic for the E-Z Reader model. In contrast, we are not convinced that accounting for effects of the orthography of word$_{n+1}$ on the fixation time on word$_n$ would be so damaging to the model. Related to that issue, we discuss below (sect. R8) the effect of orthographic properties at the beginning of the word that is to the right of fixation with respect to landing positions, and also argue that the visual processing stage that has been incorporated into the most recent version of E-Z Reader could account for such effects in a manner similar to how the attention gradient in SWIFT (Engbert et al. 2002) would handle such effects. (However, for reasons of parsimony, we are reluctant to do so unless the data forces such a revision.)

R5. Do cognitive processes drive the eyes in reading?

At some level, we would have thought that the answer to this question has already been conclusively answered in the affirmative. Rayner (1998) identified well over 100 published articles that have shown some influence of cognitive processing on eye fixations and eye movements. Indeed, all of those researchers who use eye movement data to study some aspect of language processing implicitly assume that the link between eye movements and cognition is quite tight. Furthermore, as we noted in the target article, the majority of the viable existing models of eye-movement control assume such a link. Yet, **Vitu** and **McConkie & Yang** still question this basic assumption.

Vitu's statement that "E-Z Reader may not be a serious candidate" as a realistic model of eye-movement control in reading is a non sequitur. Of course it is a serious candidate (and many other commentators are quite willing to acknowledge this point, even those who favor other models). We believe that, as with some of the other commentators, much of Vitu's argument rests on points that are disputable. She stresses that cognitive activities are not the engine driving saccades in reading; she further suggests that the landing position in the word is more important than frequency. This claim is quite puzzling and reverts to research on the optimal landing position effect. In a series of experiments with words in isolation, O'Regan and colleagues (see, for example, O'Regan et al. 1984) demonstrated that: (a) the probability of refixating a word increases dramatically as the initial landing position moves further from the optimal landing position; and (b) there is a processing cost in gaze duration of about 20 msec for each letter that the eyes deviate from this optimal letter position. From these data, O'Regan (1992b) argued that oculomotor effects related to where a given word is fixated is the most important determinant of how long the eyes remain in place. However, Vitu's argument is quite puzzling because her own work (see Vitu et al. 1990) clearly demonstrated that, during natural reading (i.e., when reading text, as opposed to isolated words), the refixation effect remained, but that the processing cost was greatly attenuated or disappeared (see also Rayner et al. 1996). That is, the gaze durations on words were relatively flat across different landing positions during reading.

To support her claim that cognitive processes don't guide eye movements in reading, **Vitu** now cites a study by Vitu et al. (2001) which reported that single fixations were longer when the eyes landed in the middle of a word than when they landed near the ends of the word, giving an inverted V-shape function. Vitu et al. correctly noted that our model does not predict this effect; however, it is also inconsistent with O'Regan's (1988) Strategy-Tactics model (which inspired much of the work on landing-position effects in the first place). Vitu also notes that this result is inconsistent with results reported by Rayner et al. (1996) that showed relatively flat functions due to landing position. However, she doesn't mention two important points: First, both Rayner et al. (1996) and Vitu et al. (2001) found an effect of word frequency that was independent of where the eyes landed on the word in single fixation cases; that is, single fixations on low-frequency words were consistently longer than on high-frequency words. Second, the inverted V-shape function reported by Vitu et al. (2001) is inconsistent with data reported by Vitu et al. (1990). Specifically, Figure 6b of Vitu et al. (1990) shows gaze durations on target words separately for cases when two fixations were made on the target word and cases when only one fixation was made. The latter curve (like that reported by Rayner et al. 1996) is relatively flat across different fixation positions in the word. (**Hyönä & Bertram** also claim to have replicated the Vitu et al. 2001 finding of an inverted V-shaped function; however, they refer to first fixations, and the point under debate involves single fixations.) So, at this point, there is clearly some inconsistency in Vitu's own data, and it seems reasonable to conclude that there isn't a systematic effect of landing position on fixation times.

McConkie & Yang argue that there is "counterevidence" to the notion that saccades are triggered by a cognitive event. We would like to know what counterevidence they have in mind other than their own work. In their experiments (Yang & McConkie, 2001; in press), text is replaced by some other pattern (strings of X's, random letters, other text, etc.) on selected fixations. Based on this paradigm, they claim that cognitive processes have an effect only very late in fixations because only very long fixations are influenced by this manipulation. They further make the interesting proposal that cognitive events only inhibit saccades, not trigger them.

In contrast to this evidence, there is an overwhelming amount of data from many labs indicating that word frequency is an important determinant of the gaze duration on a word (see Rayner 1998); obviously, the decision of when to move on to the next word is an important element of eye-movement control in reading. Moreover, contrary to **McConkie & Yang**'s claim that effects like word frequency work only by inhibiting saccades (which would mean that only the longer "tail" of the fixation-duration distribution would be affected by frequency), Rayner (1995) demonstrated that the effect of word frequency was to shift the distribution, with its shape remaining largely the same for high- and low-frequency words.

Vitu and **Brysbaert & Drieghe** attempt to undercut the claim that word frequency is an important determinant of eye control in reading by suggesting that word-frequency effects are largely word-length effects in disguise (because word length and word frequency are highly correlated in text; see also the section on the determinants of skipping, R7). Their argument rests largely on correlational analyses using a corpus of text in which neither frequency nor word length is controlled experimentally (Brysbaert & Vitu 1998). In these analyses, word length is found to predict a higher percentage of the variance than a "cognitive load" variable that reflects a mixture of word frequency and predictability. On this basis, they concluded that word-frequency and predictability effects are actually word length effects in disguise. However, the problem with these correlational analyses is that they seduce people into thinking one can infer causal relationships when in fact there is *no way* to infer how much of the "common variance" predicted by word frequency and length is attributable to one variable or the other.

By contrast, experimental manipulations allow one to do just that (i.e., unconfound these two variables). There are many studies (see Rayner 1998) in which word length is equated and where word frequency produces sizeable effects on gaze duration (and other immediate measures of eye movements). With respect to this, it is worth noting that E-Z Reader not only predicted the fixation duration means for several frequency categories of words (in which frequency and length are confounded), but also predicted the means for key target words in two studies (Rayner et al. 2003a; Schilling et al. 1998) in which word length was equated for words of differing frequency. This indicates that word-frequency effects aren't word-length effects in disguise.

We should also point out that evidence for cognitive control is not limited to effects of word frequency, as the predictability of a word in text also influences the gaze duration on a word (Ehrlich & Rayner 1981; Rayner & Well 1996). In addition, the frequencies of the component morphemes of words (for which the frequency and length of the words are equated) also have significant effects on the gaze durations of words (Andrews et al., in press; Hyönä & Pollatsek 1999; Juhasz et al. 2003; Niswander et al. 2000; Pollatsek et al. 2000). Indeed, in the Hyönä & Pollatsek (1999) and Pollatsek et al. (2000) studies using Finnish compound words, the frequency of each component had almost a 100-msec effect on gaze durations – hardly a negligible effect!

We should quickly add that we are not denying that word length influences the time needed to identify a word, although we do think that these "length" effects are also largely related to cognitive processing. Consider an extreme case of a 15-letter word. It is clear that the length of this word will slow the identification of the word considerably because it will take at least two fixations to encode all of its letters (unless the word is so orthographically unusual that one doesn't have to process most of its letters). The current version of the model does incorporate an assumption that has the potential to explain such an effect: Letters are more difficult to process the further they are from fixation. At present, we admit that this part of the model is rough and doesn't fully handle all the factors that influence the legibility of letters (see the section on word encoding, R3). We also agree with **Hyönä & Bertram** that it is still an open question whether our model can satisfactorily handle all word length effects, such as their data indicating an interaction between word length and the effect that is attributable to the frequency of the components of Finnish compound words (see sect. R3).

Let us now return to the issue of what can be concluded from **McConkie & Yang**'s paradigm (in which the text disappears after a few milliseconds on some fixations). We are quite skeptical that the results obtained in this paradigm reflect normal processing activities during reading, and we suspect that people develop special strategies for dealing with this situation. Consider the following analogy: If you ride your bike down a certain road and a large dog jumps out from behind a bush, then we suspect that the next time (and every other time) you ride down that particular road you may do so much more carefully than you did previously; that is, you would adopt a strategy for dealing with the possibility of the lurking dog. Likewise, we suspect that readers adopt non-normal strategies when reading in Yang and McConkie's text-replacement paradigm, and that this partly explains why their conclusions are so much at odds with all the data discussed above.

This does not necessarily mean that conclusions from any study in which there is a display change can't be used to infer something about the process of normal reading. For example, in most of the moving-window and boundary experiments that we (and others) have conducted, readers are seldom aware of the display changes, and are therefore unlikely to adopt special strategies. Furthermore, in those moving-window paradigms in which readers are aware of the display changes (e.g., when all the letters outside of the moving window are replaced by X's), their behavior is approximately the same as when they are unaware of any such display changes. This leads us to speculate that display changes that are apparent to the reader, but which occur on every fixation (as in studies by Rayner et al. 1981), may be less disruptive than those that occur only occasionally. However, the more general point is that, when there is a display change, some argumentation and/or evidence is needed to provide assurance that the results can be generalized to normal reading.

Finally, we note that a recent experiment by Rayner et al. (2003b) involving a display change early in a fixation seems to provide compelling evidence that the cognitive activities associated with processing a word controls when the eyes move, even when the visual information has disappeared. In this experiment, each word that was fixated disappeared after 60 msec and did not reappear until the reader moved to another word. The situation is thus much like the Rayner et al. (1981) and Ishida and Ikeda (1989) studies in which a visual mask obscured the text after 50–60 msec of each fixation. Rayner et al. (2003b), like Rayner et al. (1981) and Ishida and Ikeda (1989), found that reading proceeded quite normally under this situation (i.e., reading rate in the disappearing-text condition did not differ from a control condition in which text was presented normally). Again, this result supports our argument that regular disruption of the text may adversely change the eye movement behavior of readers much less than periodic disruption of the text (as happens in Yang & McConkie's paradigm). Moreover, Rayner et al. (2003b) found that there was still a frequency effect when the fixated word disappeared. Therefore, although the fixated word was not visible after 60 msec, the reader's eyes remained in place longer when the word was low frequency than when the word was high frequency. (The size of the frequency effect was also identical in the disappearing-text and normal-text conditions.) These results appear to be rather good evidence for cognitive triggering of saccades. Therefore, given this and the other evidence discussed above, we continue to think that the basic idea that cognitive events trigger saccades in reading makes quite a bit of sense.

R6. Assumptions about saccadic programming

As stated in the target article, our model's assumptions regarding saccadic programming were largely motivated by the work of (1) Becker and Jürgens (1979), which distinguished between the early and late stages of saccadic programming and indicated the timing of these stages; (2) Morrison (1984), which incorporated the notion of parallel programming into a model of eye control in reading; (3) O'Regan (1990; 1992b) and Rayner (1979) on landing position effects, which suggested that saccades are directed towards the center of words; and (4) McConkie and his colleagues (McConkie et al. 1988; 1991), which examined the relationship between saccadic accuracy and length, and the way this relationship is affected by the duration of the fixation on the launch site. In implementing the model, we tried to be conservative. For example, we used parameter values that make the time needed to program saccades in our model relatively long (i.e., $M_1 + M_2 = 240$ msec), under the assumption that readers may be slightly slower programming their eyes in the context of natural reading than in the relatively simple tasks (e.g., the double-step paradigm used by Becker & Jürgens 1979) that have been used to determine the minimal saccadic latencies. Likewise, we adopted parameter values as suggested by the literature (e.g., a preferred saccade length of 7 character spaces; McConkie et al. 1991) even though these values would undoubtedly vary across different readers and materials (**De Graef & Germeys** seem to concur with this conjecture). And in response to **Doré-Mazars, Vergilino-Perez & Collins** (**Doré-Mazars et al.**), we have used these same metrics for all saccades – initial fixations into words as well as refixation within words.

Nevertheless, several of the commentators raise issues about the timing of events in our model, and, in particular, the timing of saccades. For example, **Findlay & White**, **Liversedge & White**, **Radach**, **Deubel & Heller** (**Radach et al.**), and **Sereno et al.** point out that, given our model's parameter values, there is seemingly very little time during which to complete the first stage of lexical processing on one word (word$_n$), initiate the labile stage of a saccadic program to move the eyes to word$_{n+1}$, complete the second stage of lexical processing on word$_n$, and then complete the first stage of lexical processing on word$_{n+1}$ so that the labile stage will – with the appropriate frequency – be canceled, causing word$_{n+1}$ to be skipped. Indeed, there is evidence that there isn't much time for cognitive operations to influence fixation durations as, for example, sometimes word-frequency effects go away on first fixation durations when there is no preview (Inhoff & Rayner 1986). Nevertheless, cognitive operations do influence fixation times on a word. Furthermore, the commentators seem to ignore the fact that the parameter values that determine the durations of these processes (i.e., L_1, L_2, M_1, and M_2) define their mean durations, and that the actual durations of each of these processes are quite variable, being distributed as gamma distributions having standard deviations equal to 18% of their means.

Osaka also objects to our assumption that the boundaries of an upcoming (parafoveal) word are used to direct the eyes to that word in our model. We did not wish to suggest that other variables are not important determinants of where the eyes move (as we stated in the introduction), and it may well be the case that word boundary information is not as important in some languages (such as Japanese) as it is in English. Instead, we simply wanted to provide an explanation that is reasonable given what we know about the rapid availability of word boundary information in the parafovea (McConkie & Rayner 1975; Morris et al. 1990; Pollatsek & Rayner 1982). Interestingly, though, **Inhoff & Shindler** report an experiment that was designed to test this assumption, and the results of this experiment do provide evidence that is consistent with the idea that word boundaries are important determinants of saccade targets during the reading of English.

Finally, **Beauvillain & Pouget** and **Sereno et al.** note that there is evidence that visual processing is enhanced at those spatial locations that are the intended targets of upcoming saccades (Reynolds et al. 2000; Yeshurun & Carrasco 1998). Our model is not inconsistent with these results; as we stated in the target article (see also sect. R2), we distinguish between two types of attention (based on Treisman 1969) – *input selection,* which is used to orient the oculomotor system during the planning of saccades, and *analyzer selection,* which facilitates lexical processing. The finding that visual processing is enhanced in saccade-target locations is readily handled by E-Z Reader 7 if one simply assumes that shifts of input selection enhance processing in the model's early visual-processing stage.

R7. Determinants of skipping

A number of the commentaries (**Brysbaert & Drieghe**, **Engbert & Kliegl**, **Kliegl & Engbert**, **Radach et al.**) deal with the issue of word skipping. They raise two general points. First, Brysbaert & Drieghe (see also **Vitu**) argue that the most important variable affecting word skipping is word length, and they further suggest that we don't adequately deal with word length in the E-Z Reader model. Second, Engbert & Kliegl, **Kliegl & Engbert**, and Radach et al. argue that a key prediction of E-Z Reader regarding skipping (i.e., that there should be some cost associated with skipping a word) is inconsistent with the empirical data. We will address these points in turn.

With respect to the issue of word length, we agree with **Brysbaert & Drieghe** that word length is a powerful variable with respect to word skipping. Indeed, an earlier paper by Rayner and McConkie (1976) showed quite clearly that as word length increased, the probability of fixating the word increased (and conversely, the probability of skipping it decreased). So, there is no argument there. However, Brysbaert & Drieghe criticize our model on the grounds that we don't explicitly include word length as a separate parameter in the model. As we indicated above, though, word length does have its own effect in the model because, the longer a word, the slower the processing of the component letters (a result of acuity limitations). Analogous to what we argued earlier (sect. R5) it isn't quite fair to focus on the 60% skipping rate difference between very high-frequency and very low-frequency words because this is clearly contrasting words of different lengths. When we look at words that are matched in length but vary in frequency, there are clear skipping effects (Rayner et al. 2003a; Schilling et al. 1998).

As with our comments about gaze duration in the prior section, we would like to point out that a few of the commentators (**Brysbaert & Drieghe**, **Feng)** seem to be under the impression that the only data that have been simulated are the global properties of the sentences used by Schilling et al. (1998). This is not correct. Not only was the model compared with the full set of words in the Schilling et al. sentences, comparisons were also made for specific target words (which varied on frequency) in the Schilling et al. sentences. Furthermore, in some very recent work (Rayner et al. 2003a), E-Z Reader was compared with human data in a study in which target words varied in frequency and predictability but were matched on word length (with lengths varying from 4 to 8 letters). Although the model had to be modified to account for the fixation times, both the original model and the modified model could account for the skipping data. Moreover, as noted in section R5, most of the evidence purporting to show that word frequency has little influence on skipping comes from correlational analyses in which it is impossible to separate the influence of the two variables.

With respect to the issue of fixation costs as a result of skipping, E-Z Reader does indeed predict that, when readers skip a word, the fixation on the word preceding the skip and/or the fixation on the word following the skip should be inflated compared with the fixations on those same words in the absence of skipping. Indeed, there wasn't originally any controversy (that we were aware of) with this issue because we (Pollatsek et al. 1986; Reichle et al. 1998) first reported these effects. More recently, though, Radach and Heller (2000) and Kliegl and Engbert (2003) reported small or nonexistent effects on the fixation prior to a skip (though in their commentary, **Kliegl & Engbert** note that they did find an increased fixation following a skip – as predicted by E-Z Reader). However, we (Rayner et al. 2003a) recently again found a 26-msec effect on the fixation prior to a skip, and Pynte et al. (in press) likewise reported a 43-msec effect on the fixation prior to a skip.

Kliegl & Engbert do have a point: that the size of the skipping-cost effect obtained in the empirical data is typically smaller than that predicted by E-Z Reader. However, there is a problem with negative findings on this point that is related to our earlier discussion about what can and cannot be inferred from correlational analyses. That is, all of these analyses (i.e., both the findings agreeing with E-Z Reader's predictions and those that do not) are correlational, in the sense that the experimenter does *not* control whether the reader skips or doesn't skip a given word. As a result, these data are gathered over different locations in the text, different readers, and readers in different states of alertness. This means that, in the ecology of natural reading, places where lexical processing is likely to proceed most rapidly (because the text is easier and/or the reader more efficient) will produce more skipping and shorter fixation times. This in turn means that, if there is no inherent causal relationship between skipping and lengthening of the preceding fixation, then one would expect that fixations that precede skips would be shorter than fixations that do not precede skips, and this would actually work against the model's prediction by producing the opposite relationship between the two variables. As a result, one would not expect consistent results on this issue because the model's predictions apply to the "ideal case," which has not been produced experimentally (and may not be possible to produce).

R8. Landing position effects

A number of the commentaries deal with landing-position effects. We have already addressed **Vitu**'s points above (see sect. R5). However, **Hyönä & Bertram**, **Liversedge & White**, and **Underwood** also raise this issue in their commentaries. And, although Underwood and Liversedge & White note that we relegate discussion of these findings to a note in our target article, we did not mean to suggest that this issue in not important; indeed, we have two comments that we would like to make regarding this issue.

First, we acknowledge that there is now quite a bit of data suggesting that the initial landing position in a word can be influenced by the orthographic characteristics of a parafoveal word. However, it should also be acknowledged that the sizes of these effects are often rather small. For example, in White and Liversedge (in press), the effect amounted to about one-fourth of a character space (i.e., the landing position when the next word begins with an orthographically illegal letter string is about one-fourth of a character space to the left of the landing position that is observed in the control condition). Frankly, some people find such small effects to be of only minor interest (although we do not agree with such an assessment, if the effects are reliable).

The second, and more critical, comment is simply that, in the most recent version of E-Z Reader, we envisioned that the low-level visual-processing stage we posited would be able to account for such effects. Although we did not implement the specific assumptions necessary to do this, we suspect that doing so would be a relatively trivial endeavor. Nevertheless, certain aspects of how this would be done need to be carefully considered. For example, although **Liversedge & White** focus on our suggestion that our model's early visual-processing stage might be sensitive to word shape considerations, we don't necessarily think that orthographical legality translates into word shape. Rather, the letters themselves would have to be processed at an early filtering stage.

R9. Determinants of refixations

Doré-Mazars et al., **Hyönä & Bertram**, and **Pacht** all raise questions about the model's ability to deal with refixations. We agree that this is perhaps the weakest part of the E-Z Reader model. This is clearly an aspect of reading that is difficult to model, as it's clear that many variables influence the pattern of refixations. As we noted earlier, there is considerable evidence that low-level factors, such as the location of the initial fixation, influence both the initial fixation duration on a word and whether the word is refixated. Higher-level factors, such as word frequency, also influence whether a word is refixated (Rayner & Pollatsek 1987). Our model attempts to explain how these factors influence refixation by positing automatic refixations that can be canceled by subsequent lexical processing (in a manner that is directly analogous to how the cancellation of previously programmed saccades produces skipping).

We think that the current version of the model gives a reasonably plausible account of some of the refixation data. What has become clear, though, is that word encoding during reading is considerably more complex than the relatively simple notion that two stages of word identification are the only processes that influence eye movements and (more generally) the speed with which the reader progresses through the text. As we pointed out in section R3, it is now clear that parts of words (e.g., their morphemes) influence the speed of cognitive processing during reading, as do other aspects of the word, such as the number of "neighbors" that it has and/or the frequency of these neighbors. It therefore seems clear that these factors (and others) need to be taken into account in explaining refixations. Therefore, it doesn't seem worthwhile to discuss in any detail whether we currently have an optimal account of how the first two factors (i.e., the initial fixation location and word frequency) influence refixations, because the data we have been trying to model are undoubtedly also influenced by many other factors (e.g., a word's neighbors) that our model simply ignores. Our attempt to explain refixations should definitely be viewed as work in progress.

R10. Higher-level processing and regressions

As we noted earlier, some commentators maintain that there is too much emphasis on cognitive events influencing eye movements in the E-Z Reader model, whereas others argue that not enough emphasis is put on such events influencing eye movements. **Liversedge & White, Murray**, **Perea & Carreiras**, and **Raney** all seem to be arguing for the latter. Indeed, Liversedge & White and Murray each point out that when readers read syntactically ambiguous sentences (so-called "garden path" sentences) that the fixation times are often inflated when the word that disambiguates the intended meaning of the sentence is first encountered. We can't disagree with this claim because the first such data were collected in our laboratory (Frazier & Rayner 1982; Rayner et al. 1983a). Furthermore, the data that have been collected in our lab (Schustack et al. 1987) clearly show that post-lexical integration processes can also affect the fixation times on a target word, and the data that Raney describes also seem consistent with higher-level processing influencing fixation times.

Let us therefore make three points relative to this issue. First, these higher-level effects sometimes show up immediately, but often, they are delayed (Frisson & Pickering 1999; Garrod & Terras 2000; Pickering & Frisson 2001) and in fact can sometimes be quite elusive in the eye movement record. Second, a recent modification of the E-Z Reader model (Rayner et al. 2003a) incorporates postlexical integration activities as a factor influencing fixation time, and this assumption was relatively easy to add into the model. Finally, the kinds of effects that arise from syntactic processing would be much more difficult to incorporate into the model since, as we argued above and in the target article, one would need to have a fully implemented model of parsing (including reanalysis procedures) and discourse processing, and we do not think a viable model of these operations is available yet.

This last point brings us to **Perea & Carreiras**. They clearly would like to have a fully implemented model of eye-movement control that includes both parsing and reanalysis procedures. We agree that this would be a worthy – but very difficult – goal. Perea & Carreiras then go on to distinguish between local (i.e., within-word) and global regressions. They do not really define what the latter are, but since they talk about long-range regressions, we assume they are referring to the type of regressions that occur when comprehension processes break down. Indeed, when readers are "garden-pathed" by syntactically ambiguous sentences, they are sometimes fairly accurate in moving their eyes back to that part of the text where parsing would have initially gone awry (see also **Kennedy**'s commentary). Frazier and Rayner (1982) first demonstrated this, though Perea & Carreiras refer to a study by Mesenguer et al. (2002). They are mistaken, however, when they say that SWIFT (Engbert et al. 2002) can account for such global regressions. There is no parsing mechanism in SWIFT, and any regressions generated by the model are attributed to lexical processing. That said, SWIFT would probably overpredict the number of regressions, because the rate of regressions it currently predicts does not include those regressions that would be a result of comprehension problems. We reiterate that while it is a worthy goal to have a fully implemented model of parsing and discourse processing, such a model is not currently available and/or easily incorporated into the context of a model of eye-movement control.

R11. Generalizing, comparing, and evaluating models

The issue of how to evaluate theories and, more recently, computational models, has not been resolved (see Hintzman 1990; Jacobs & Grainger 1994; and the **Hanes & McCollum** and **Reilly & Radach** commentaries). The arguments about what constitutes a good model are complex, and are often philosophical in nature (Uttal 1990), so instead of rehashing those arguments here, we will instead simply say we agree with **Hyönä & Bertram**'s assessment that models should: (1) account for empirical data, (2) be conceptually transparent, (3) be psychologically and neurally plausible, and (4) generate novel predictions. We shall use these criteria to address the criticisms that were raised in the commentaries in the discussion that follows.

Feng disapproves of our method of fitting E-Z Reader to the Schilling et al. (1998) data. We believe his specific criticisms are unwarranted because his analysis is fundamentally flawed. In fitting our model to the data, for example, we normalized the differences between the observed and predicted fixation durations and the observed and predicted fixation probabilities so that both measures could be compared on a common scale. Although we did this using standard deviations rather than standard errors, we did not use the normalized scores to make statistical inferences but instead used them as an index to compare the relative performance of successive versions of our model. Feng also objects that, by fitting our model to aggregate data, our model unfairly takes advantage of multicolinearity, and hence accounts for the data only slightly better than simpler, statistical methods (e.g., principal components analysis using log frequency as a factor). Again, this criticism ignores the fact that our goal was not simply to account for as much of the variance in the data as possible; rather, our goal was to specify the cognitive operations that occur during reading and how these operations give rise to the data. The criticism also ignores the fact that our simulations were not limited to predicting means; our model accounts for other aspects of the data, including their distributions and a variety of other phenomena (e.g., spillover effects), a few of which are affected by dependencies between adjacent words (e.g., fixation costs due to skipping). Finally, this criticism overlooks the fact that our model was designed to account for the data while respecting a variety of constraints (e.g., visual acuity limitations, what is known about the time course and nature of saccadic programming, etc.).

Murray also voiced a concern that our model had not been tested on data sets other than those that had been used in developing the model (i.e., the sentences of Schilling et al. 1998). We hope this concern is lessened by the fact that **Sparrow**, **Miellet & Coello** (**Sparrow et al.**) have been reasonably successful using E-Z Reader to simulate the eye movements of native French readers who were reading their native tongue. Although these simulations did not fit the observed data perfectly, they were done using the same parameter values that were used to fit the Schilling et al. (1998) sentences, which provides an extremely conservative test of the model's adequacy because many of the model's parameters (e.g., β_1, β_2, and Ψ) would be expected to take on different values across subjects and languages. (This point regarding parameters is equally relevant to **Osaka**'s concern regarding the model's capacity to simulate the eye movements of native Japanese readers.) Two other facts are also noteworthy about the work of Sparrow et al. First, the model was fit to a passage of text rather than a set of unrelated sentences, which again shows the generality of our model. Second, the model was also fit to the data of dyslexic readers, which supports **Fischer**'s contention that the model may be useful for understanding individual differences in reading ability and the source(s) of reading impairment.

We should also mention (in response to **Murray**) that the model has been used with other materials and in other contexts. For example, as noted above, the model has been used with sentences containing word targets whose frequencies and predictabilities were orthogonally manipulated (Rayner et al. 2003a). It has also been used to interpret the patterns of eye movements observed when native readers of Finnish encounter compound words (Pollatsek et al. 2003; for a brief description of this work, see **Hyönä & Bertram**). And, as mentioned in our target article, the model has been used to simulate eye fixations on lexically ambiguous words (Duffy et al. 1988). In all three of these cases, the model was not used simply to fit the data but was instead used to evaluate the adequacy of other theoretical assumptions. Again, we cite these examples to show how E-Z Reader can be used as a framework for understanding phenomena that the model was not explicitly designed to explain.

With respect to the second issue, conceptual transparency, we again point out that the main advantage to our minimalist approach is its simplicity. In contrast to many of the models reviewed in the target article (e.g., Glenmore; Reader), E-Z Reader (1) can be described by a few principles, (2) can be represented by a simple box-and-arrow diagram, and (3) can be implemented using a small number of equations and parameters. We also reject **Reilly & Radach**'s notion that transparency is not important, and that its absence can be offset through the enhanced "biological realism" that comes from being implemented as a connectionist network. (**Raney**'s commentary strongly suggests that he also shares our opinion.)

We maintain that the biology of reading can be described at many different levels, and that, at the level of interacting neural systems, the best descriptions are functional ones. Other, more complicated descriptions simply confuse matters. To cite one example: **Shillcock et al.** suggest that our model is inadequate because it ignores the fact that printed words are identified by a visual processing system that is divided between the two cerebral hemispheres. This objection is – in our opinion – simply not warranted. As Shillcock et al. indicate, the fact that the visual field is divided into two hemi-fields is largely irrelevant in normal reading except in that it suggests why the eyes are directed towards the centers of words. Implementing this architectural constraint in our model would therefore do little to further our understanding of eye guidance during reading, and would probably obscure other, more important relationships. For similar reasons, we contend that – although E-Z Reader could be implemented as a connectionist model – doing so would make the model more difficult to understand and would probably do little to further our understanding of eye movements during reading.

With respect to the third criterion, plausibility, it is important to note that E-Z Reader was developed to be a theory of eye-movement behavior during reading. Consequently, we have placed a premium on what has been learned about eye movements during actual reading (for a review, see Rayner 1998) in designing the model. For example, the time needed to identify words (Rayner & Pollatsek 1989) and the timing and accuracy of saccades (Becker & Jürgens 1979; McConkie et al. 1988) have always constrained our model. Additional assumptions have been added only when they were needed to make the model more physiologically plausible (e.g., visual acuity limitations; Rayner & Morrison 1981). And, although previous versions of our model have been agnostic with respect to neural implementation, E-Z Reader 7 represents an attempt to map the functional components of our model onto the neural structures that have been implicated in reading. Although this mapping is admittedly a gross oversimplification (see the commentaries of **Beauvillain & Pouget**, **Sereno et al.**, and **Shillcock et al.**), we believe that it is precise enough to be falsifiable. Moreover, the mapping seems to have some theoretical utility. **Fischer,** for example, suggests that it may "serve as a model of dyslexia," and this conjecture is supported by the preliminary simulation results that were reported by **Sparrow et al.**

The last – and we would argue the most important – criterion for evaluating models is their capacity to generate novel predictions and guide new research. With respect to this issue, **Hyönä & Bertram** ask whether the model is "capable of producing novel predictions?" or whether it instead is "restricted to predicting only the effects that it is designed to simulate?" We would argue that our model can – and already has – generated new predictions. To cite one example acknowledged by Hyönä & Bertram: E-Z Reader predicted that the gaze durations on the words following skips should be inflated because of less parafoveal processing from the more distant launch site word. This finding was confirmed in the Schilling et al. (1998) data (see Note 1 of **Kliegl & Engbert**'s commentary).

Several of the commentaries also contain new predictions. **Engbert & Kliegl**'s analysis of word skipping in our model and SWIFT (Engbert et al. 2002) indicates that they make very different predictions regarding the causes of skipping: In E-Z Reader, skipping depends on the outcome of the "horse race" between the first stage of lexical processing and saccadic programming, whereas in SWIFT, skipping depends only on lexical processing. Similarly, **Hane & McCollum**'s analysis suggests that the basic distinction between L_1 and L_2 in our model has strong implications for the underlying neural architecture of the systems that guide eye movements during reading. Although neither prediction is specified in detail, we are confident they will ultimately guide new research. Finally, **Reingold**'s analysis of the factors that should selectively influence L_1 (e.g., stimulus quality; Borowsky & Besner 1993) has led to a prediction that can be tested: Visual degradation of word$_n$ should slow L_1 and hence increase the fixation duration on that word, but it should not affect the spillover effect observed on word$_{n+1}$. Although this prediction is based on both a specific interpretation of the distinction between L_1 and L_2 (that the former corresponds to orthographic processing; see our discussion in sect. R3) and the assumption that visual degradation will only affect very early processing (e.g., that it won't lead to misidentification of letters), its confirmation would complement the results described by **Raney** and thereby show that the two stages of lexical processing that are postulated in our model can be dissociated.

Finally, we would like to reply to some of the commentaries that have advocated that models of reading should be consistent with both basic eye-movement and neurophysiological data (e.g., **Beauvillain & Pouget**, **De Graef & Germeys**, **Osaka**, **Sereno et al.**, and **Shillcock et al.**). On some level, one cannot disagree with these prescriptions. However, we do think the specific points about how our model is found wanting point to dangers of applying such strictures. First, with respect to the claims about inconsistency with "basic" eye-movement data, virtually all of the claims rely on a hidden assumption, termed the "subtractive" method by Sternberg (1966; 1969; 1975). This assumption is, essentially, that in an apparently simple task, such as maintaining fixation and then moving one's eyes to the sudden onset of a stimulus, one is studying a processing module that is common to eye-movement control in general, and that a task like reading merely engages other, later processes that operate "on top of" the basic processes exposed by these simpler tasks. That is, the assumption is that the eye-movement parts of these nonreading tasks and reading are identical. While this assumption is, of course, logically possible, it is not a given and needs to be defended – at least by a serious task analysis of both tasks. Contrary to **Radach et al.**, for example, we are skeptical that the task of reading, which involves no sudden onsets of stimuli, no requirement to maintain fixation artificially, and no stimuli that require any serious cognitive analysis, is a candidate for literal extrapolations of findings from experiments of this kind, especially with respect to the exact timing that various cognitive processes require (e.g., attention shifts).

The point we would like to make about the relationship between a model of reading and neurophysiological data involves both a similar point and a different one. First, there is a similarity to the point made about task analysis; that is, in most neurophysiological research (e.g., ERP experiments), one typically examines word identification in isolation and then extrapolates from this to the reading of text. Clearly, this should be done with some caution. Perhaps more important, however, is the point that one can't take such neurophysiological findings as being written on stone tablets. For example, **Radach et al.** asserted that ERP data indicating that two curves do not significantly deviate before 130 msec can be used to infer that lexical analysis does not begin until then. This clearly assumes the technique is flawless and that this is a direct "window" into all the relevant neurological structures that underlie word identification – an assumption we find quite hard to take seriously.

More generally, we think such comments seriously misgauge the state of the art in both cognitive psychology and cognitive neuroscience. That is, they seem to be based on an assumption that we really understand all of the "atoms" of the cognitive process and have carefully measured all of their relevant properties, and that this information is now ready to be put together into something like a "periodic table." Instead, we would like to suggest that we are still groping around in the dark on many basic issues – for example, after 40–50 years of research we still do not fully understand how people recognize words or objects – and that many of the suggestions that we standardize the way we model (e.g., **Reilly & Radach**) are quite misguided. In fact, we would like to invite those who think otherwise (e.g., **Sereno et al.**) to develop a fully implemented model of eye-movement control that is both neurophysiologically based and neurophysiologically plausible.

R12. Conclusion

In closing, we would like first to restate that we fully acknowledge that our model, E-Z Reader, is not a complete model of reading. Our minimalist approach to modeling will (by definition) insure that whatever model we end up developing will grossly oversimplify the cognitive processes that determine when and where the eyes move during reading. However, we also believe that this is a modest price to pay for having a model that is transparent enough to be used as a heuristic for guiding research. And, as witnessed by both the number of responses to our target article and much of the content of those responses (especially the predictions that were discussed in sect. R11), the E-Z Reader model has already clearly demonstrated its value in this regard.

We would also like to note that – in our opinion – the bulk of the responses to our target article were valuable in at least two ways. First, they provided us with the opportunity to clarify points in the description of our model that may have been incomplete or otherwise vague. Second, and more important, is that the dialogue further clarified many of the issues that are currently the subject of (sometimes heated) debate within the area of reading research. This process forced us to rethink certain aspects of our model, and, in so doing, has already led us to modify E-Z Reader 7. For example, as outlined at the outset of this response, **Finday & White**'s analysis of the time course of lexical processing forced us to re-implement the early stage of visual processing in our model so that visual processing needs to be initiated with each new fixation. (As noted earlier, this modification did not adversely affect the model's performance, and had the desired effect of reducing the overall size of the parafoveal preview benefit to approximately half the value we reported in the target article.) Clearly, we expect that E-Z Reader will continue to develop as our understanding of the determinants of eye movement control in reading continues to grow.

References

Letters "a" and "r" appearing before authors' initials refer to target article and response, respectively.

Altarriba, J., Kambe, G., Pollatsek, A. & Rayner, K. (2001) Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish-English bilinguals. *Perception and Psychophysics* 63:875–890. [aEDR]

Altarriba, J., Kroll, J. F., Sholl, A. & Rayner, K. (1996) The influence of lexical and conceptual constraints on reading mixed language sentences: Evidence from eye fixation and naming times. *Memory and Cognition* 24:477–92. [aEDR]

Anderson, J. R. (1983) *The architecture of cognition.* Harvard University Press. [aEDR]

Anderson, J. R. & Lebiere, C. (1998) *The atomic components of thought.* Erlbaum. [aEDR]

Andrews, S. (1997) The role of orthographic similarity in lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin and Review* 4:439–61. [SA]

Andrews, S. & Heathcote, A. (2001) Distinguishing common and task-specific processes in lexical retrieval: A matter of some moment? *Journal of Experimental Psychology: Learning Memory and Cognition* 27:514–44. [SA]

Andrews, S., Miller, B. & Rayner, K. (in press) Eye movements and morphological segmentation of compound words: There is a mouse in the mousetrap. *European Journal of Cognitive Psychology.* [SA, aEDR, rKR]

Atkinson, R. C. & Juola, J. F. (1973) Factors influencing speed and accuracy of word recognition. In: *Attention and performance IV,* pp. 583–611, ed. S. Kornblum. Academic Press. [rKR]

(1974) Search and decision processes in recognition memory. In: *Contemporary developments in mathematical psychology: Vol. 1. Learning, memory, and thinking,* ed. D. H. Krantz, R. C. Atkinson, R. D. Luce & P. Suppes. Freeman. [rKR]

Balota, D. A., Pollatsek, A. & Rayner, K. (1985) The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology* 17:364–90. [aEDR, EMR]

Basso, M. A. & Wurtz, R. H. (1997) Modulation of neuronal activity by target uncertainty. *Nature* 389(6646):1966–69. [GWM]

(1998) Modulation of neuronal activity in superior colliculus by changes in target probability. *Journal of Neuroscience* 18(18):7519–34. [GWM]

Bavelier, D., Corina, D., Jezzard, P., Padmanabhan, S., Clark, V. P., Karni, A., Prinster, A., Braun, A., Lalwani, A., Rauschecker, J. P., Tuner, R. & Neville, H. (1997) Sentence reading: A functional MRI study at 4 Tesla. *Journal of Cognitive Neuroscience* 9:664–86. [aEDR]

Beauvillain, C. & Doré, K. (1998) Orthographic codes are used in integrating information from the parafovea by the saccadic computation system. *Vision Research* 38:115–23. [SPL, aEDR, GU]

Beauvillain, C., Doré, K. & Baudouin, V. (1996) The "center of gravity" of words: Evidence for an effect of the word-initial letters. *Vision Research* 36:589–603. [aEDR, GU]

Becker, W. & Jürgens, R. (1979) An analysis of the saccadic system by means of double step stimuli. *Vision Research* 19(9):1967–83. [GWM, KDM, aEDR, rKR, RRa]

Bertram, R. & Hyönä, J. (2002) Word processing and the optimal viewing position. Poster presented at the 3rd International Conference on the Mental Lexicon, Banff, Canada. [JH]

(2003) The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language* 48:615–34. [JH, rKR]

Bertram, R., Hyönä, J. & Pollatsek, A. (2003) Segmentation cues in processing Finnish compound words. (in preparation). [JH]

Biederman, I. (1981) On the semantics of a glance at a scene. In: *Perceptual organization,* pp. 213–63, ed. M. Kubovy & J. R. Pomerantz. Erlbaum. [PDG]

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M. & Prieto, T. (1997) Human brain language areas identified by functional magnetic resonance imaging. *The Journal of Neuroscience* 17:353–62. [aEDR]

Binder, K. S., Pollatsek, A. & Rayner, K. (1999) Extraction of information to the left of the fixated word in reading. *Journal of Experimental Psychology: Human Perception and Performance* 25:1162–72. [aEDR]

Binder, K. S. & Rayner, K. (1998) Contextual strength does not modulate the subordinate bias effect: Evidence from eye fixations and self-paced reading. *Psychonomic Bulletin and Review* 5:271–76. [aEDR]

Biscaldi, M., Fischer, B. & Hartnegg, K. (2000) Voluntary saccade control in dyslexia. *Perception* 29:509–21. [BF]

Blanchard, H. E., McConkie, G. W., Zola, D. & Wolverton, G. S. (1984) Time course of visual information utilization during fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance* 10:75–89. [GWM]

Blanchard, H. E., Pollatsek, A. & Rayner, K. (1989) The acquisition of parafoveal word information in reading. *Perception and Psychophysics* 46:85–94. [aEDR]

Bookheimer, S. Y., Zeffiro, T. A., Blaxton, T., Malow, B. A., Gaillard, W. D., Sato, S., Kufta, C., Fedio, P. & Theodore, W. H. (1997) A direct comparison of PET activation and electrocortical stimulation mapping for language localization. *Neurology* 48:1056–65. [aEDR]

Borowsky, R. & Besner, D. (1993) Visual word recognition: A multi-stage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:815–40. [EMR, rKR]

Bouma, H. & deVoogd, A. H. (1974) On the control of eye saccades in reading. *Vision Research* 14:272–84. [aEDR]

Boutsen, L., Lamberts, K. & Verfaillie, K. (1998) Recognition times of different views of 56 depth-rotated objects: A note concerning Verfaillie and Boutsen (1995). *Perception and Psychophysics* 60:900–907. [PDG]

Briihl, D. & Inhoff, A. W. (1995) Integrating information across fixations during reading: The use of orthographic bodies and of exterior letters. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21:55–67. [AWI]

Broadbent, D. E. (1971) *Decision and stress.* Academic Press. [CB]

Brown, P. (1991) DEREK: The direct encoding routine for evoking knowledge. *Basic processing in reading: Visual word recognition,* ed. D. Besner & G. W. Humphreys. Erlbaum. [aEDR]

Bruce, C. J. & Goldberg, M. E. (1985) Primate frontal eye fields. I. Single neurons discharging before saccades. *Journal of Neurophysiology* 53:603–35. [BF]

Brysbaert, M. & Vitu, F. (1998) Word skipping: Implications for theories of eye movement control in reading. In: *Eye guidance in reading and scene perception,* pp. 125–47, ed. G. Underwood. Elsevier. [MB, FV, aEDR, rKR]

Bullinaria, J. A. (1997) Modeling reading, spelling, and past tense learning with artificial neural networks. *Brain and Language* 59:236–66. [aEDR]

Byrne, M. D. & Anderson, J. R. (1998) Perception and action. In: *The atomic components of thought,* ed. J. R. Anderson & C. Lebiere. Erlbaum. [aEDR]

Caplan, D. (1992) *Language: Structure, processing, and disorders.* MIT Press. [aEDR]

Caplan, D., Alpert, N. & Waters, G. (1999) PET studies of syntactic processing with auditory sentence presentation. *NeuroImage* 9:343–51. [aEDR]

Carpenter, P. A. & Just, M. A. (1983) What your eyes do while your mind is reading. In: *Eye movements in reading: Perceptual and language processes,* pp. 275–305, ed. K. Rayner. Academic Press. [AK, aEDR, rKR]

Carpenter, R. H. S. & Williams, M. L. L. (1995) Neural computation of log likelihood in control of saccadic eye movements. *Nature* 377:59–62. [JMF]

Chelazzi, L., Miller, E. K., Duncan, J. & Desimone, R. (1993) A neural basis for visual search in inferior temporal cortex. *Nature* 363:345–47. [CB]

Churchland, P. S. (1986) *Neurophilosophy: Toward a unified science of the mind-brain.* MIT Press. [aEDR]

Churchland, P. S. & Sejnowski, T. J. (1992*) The computational brain.* MIT Press. [aEDR]

Clark, J. J. & O'Regan, J. K. (1999) Word ambiguity and the optimal viewing position in reading. *Vision Research* 39:843–57. [aEDR]

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

Cohen, J. D., Romero, R. D., Farah, M. J. & Servan-Schreiber, D. (1994) Mechanisms of spatial attention: The relation of macrostructure to microstructure in parietal neglect. *Journal of Cognitive Neuroscience* 6:377–87. [RS]

Colby, C. L., Duhamel, J.-R. & Goldberg, M. E. (1996) Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *Journal of Neurophysiology* 76:2841–52. [PDG]

Coltheart, M. & Rastle, K. (1999) Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance* 20:1197–211. [RS]

Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. (2001) DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review* 108:204–56. [SA]

Corbetta, M. (1998) Frontoparietal cortical networks for directing attention and the eye to visual locations: Identical, independent, or overlapping neural systems? *Proceedings of the National Academy of Sciences USA* 95:831–38. [aEDR]

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. L., Drury, H. A., Linenweber, M. R., Petersen, S. E., Raichle, M. E., Van Essen, D. C. & Shulman, G. L. (1998) A common network of functional areas for attention and eye movements. *Neuron* 21:761–73. [CB]

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, J. M. & Shulman, G. L. (2000) Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience* 3(3):292–97. [CB]

Corbetta, M., Miezin, F. M., Schulman, G. L. & Petersen, S. E. (1993) A PET study of visuospatial attention. *Journal of Neuroscience* 13:1202–26. [aEDR]

Corbetta, M. & Shulman, G. L. (2002) Control of goal-driven and stimulus driven attention in the brain. *Nature Reviews Neuroscience* 3:201–15. [CB]

Davachi, L., Mitchell, J. P. & Wagner, A. D. (2003) Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Science USA* 100:(2):157–62. [rKR]

De Graef, P. (1998) Prefixational object perception in scenes: Objects popping out of schemas. In: *Eye guidance in reading and scene perception,* pp. 315–38, ed. G. Underwood. Elsevier. [PDG]

De Graef, P., Christiaens, D. & d'Ydewalle, G. (1990) Perceptual effects of scene context on object identification. *Psychological Research* 52:317–29. [PDG]

Den Buurman, R., Boersma, T. & Gerrissen, J. F. (1981) Eye movements and the perceptual span in reading. *Reading Research Quarterly* 16:227–35. [aEDR]

Deubel, H., O'Regan, K. & Radach, R. (2000) Attention, information processing and eye movement control. In: *Reading as a perceptual process,* ed. A. Kennedy, R. Radach, D. Heller & J. Pynte. Elsevier. [RRa]

Deubel, H. & Schneider, W. X. (1996) Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research* 36:(1)827–37. [CB, PDG]

Deutsch, A., Frost, R., Peleg, S., Pollatsek, A. & Rayner, K. (in press) Early morphological effects in reading: Evidence from parafoveal preview benefit in Hebrew. *Psychonomic Bulletin and Review.*10:415–22. [aEDR]

Deutsch, A., Frost, R., Pollatsek, A. & Rayner, K. (2000) Early morphological effects in word recognition in Hebrew: Evidence from parafoveal preview benefits. *Language and Cognitive Processes* 15:487–506. [aEDR]

Deutsch, A. & Rayner, K. (1999) Initial fixation location effects in reading Hebrew words. *Language and Cognitive Processes* 14:393–421. [aEDR]

Dopkins, S., Morris, R. K. & Rayner, K. (1992) Lexical ambiguity and eye fixations during reading: A test of competing models of lexical ambiguity resolution. *Journal of Memory and Language* 31:461–77. [aEDR]

Doré-Mazars, K., Collins, T. & Vergilino-Perez, D. (2003) About the effects of word-length and word-frequency on the decision and metrics of refixations in long words. *Psychological Science* (submitted). [KD-M]

Doré-Mazars, K., Pouget, P. & Beauvillain, C. (in press) Attentional selection during the preparation of eye movements. *Psychological Research.* [CB]

Duffy, S. A., Kambe, G. & Rayner, K. (2001) The effect of prior disambiguating context on the comprehension of ambiguous words: Evidence from eye movements. In: *On the consequences of meaning selection: Perspectives of resolving lexical ambiguity,* ed. D. Gorfein. American Psychological Association. [aEDR]

Duffy, S. A., Morris, R. K. & Rayner, K. (1988) Lexical ambiguity and fixation times in reading. *Journal of Memory and Language* 27:429–46. [aEDR, rKR]

Duncan, J., Ward, W. & Shapiro, K. (1994) Direct measurement of attentional dwell time in human vision. *Nature* 369:313–15. [RRa]

Ehrlich, S. F. & Rayner, K. (1981) Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior* 20:641–55. [aEDR, rKR]

Engbert, R. & Kliegl, R. (2001) Mathematical models of eye movements in reading: A possible role for autonomous saccades. *Biological Cybernetics* 85:77–87. [RE, GWM, aEDR, RRe]

(2003) Noise-enhanced performance in reading. *Neurocomputing* 50:473–78. [RE]

Engbert, R., Longtin, A. & Kliegl, R. (2002) A dynamic model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research* 42(5):621–36. [AK, RE, LH, GWM, MP, aEDR, rKR]

(2004) Complexity of eye movements in reading. *International Journal of Bifurcation and Chaos* 14. (in press). [RE]

Engbert, R., & Trukenbrad, H. (2003) Modeling eye movements in visual search. Poster presented at the 12th European Conference on Eye Movements, Dundee, Scotland, 2003. {rkR}

Epelboim, J., Booth, J. R. & Steinman, R. M. (1993) Reading unspaced text: Implications for theories of eye movements. *Vision Research* 34:1735–66. [AK]

(1996) Much ado about nothing: The place of space in text. *Vision Research* 36:465–70. [AK]

Erdmann, B. & Dodge, R. (1898) *Psychologische Untersuchungen uber das Lesen.* N. Niemeyer. [aEDR]

Erlhagen, W. & Schöner, G. (2002) Dynamic field theory of movement preparation. *Psychological Review* 109:545–72. [RE]

Everatt, J., Bradshaw, M. F. & Hibbard, P. B. (1998) Individual differences in reading and eye movement control. In: *Eye guidance in reading and scene perception,* pp. 223–42, ed. G. Underwood. Elsevier. [GU]

Everatt, J. & Underwood, G. (1992) Parafoveal guidance and priming effects during reading: A special case of the mind being ahead of the eyes. *Consciousness and Cognition* 1:186–97. [GU]

Felleman, D. J. & Van Essen, D. C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex* 1:1–47. [aEDR]

Feng, G. (2001) SHARE: A Stochastic Hierarchical Architecture for Reading Eye-movement, Department of Psychology, University of Illinois, Champaign-Urbana. (Unpublished doctoral dissertation). Available at: www.garyfeng.com/dissertation/Feng_2001_dissertation.pdf [GF]

Feng, G., Miller, K. & Shu, H. (2003) Reading development in English and Chinese: Results from an eye-movement study. (submitted). [GF]

Feng, G., Miller, K. F., Zhang, H. & Shu, H. (2001) Rowed to recovery: The use of phonological and orthographic information in reading Chinese and English. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27:1079–100. [GF]

Ferreira, F. & Clifton, C., Jr. (1986) The independence of syntactic processing. *Journal of Memory and Language* 25:348–68. [AK]

Findlay, J. M. & Walker, R. (1999) A model of saccadic eye movement generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences* 22:661–721. Also available online: http://psychology.dur.ac.uk/staff/pubs/JMF-1999b.pdf. [CB, JMF, aEDR]

Fischer, B. (1987) The preparation of visually guided saccades. *Reviews of Physiology, Biochemistry, and Pharmacology* 106:1–35. [BF]

Fischer, B., Biscaldi, M. & Gezeck, S. (1997a) On the development of voluntary and reflexive components in human saccade generation. *Brain Research* 754:285–97. [BF]

Fischer, B., Gezeck, S. & Hartnegg, K. (1997b) The analysis of saccadic eye movements from gap and overlap paradigms. *Brain Research Protocols* 2:47–52. [BF]

Fischer, B. & Hartnegg, K. (2000a) Effects of visual training on saccade control in dyslexia. *Perception* 29:531–42. [BF]

(2000b) Stability of gaze control in dyslexia. *Strabismus* 8:119–22. [BF]

Fischer, B., Hartnegg, K. & Mokler, A. (2000) Dynamic visual perception of dyslexic children. *Perception* 29:523–30. [BF]

Fischer, B. & Ramsperger, E. (1984) Human express saccades: Extremely short reaction times of goal directed eye movements. *Experimental Brain Research* 57:191–95. [GWM]

(1986) Human express saccades: Effects of randomization and daily practice. *Experimental Brain Research* 64:569–78. [BF]

Francis, W. N. & Kučera, H. (1982) *Frequency analysis of English usage: Lexicon and grammar.* Houghton Mifflin. [aEDR]

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. [GER, aEDR, rKR]

Frisson, S. & Pickering, M. J. (1999) The processing of metonymy: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 25:1366–83. [rKR]

Gautier, V., O'Regan, J. K. & LaGargasson, J. F. (2000) 'The skipping' revisited in French: Programming saccades to skip the article 'les'. *Vision Research* 40:2517–31. [aEDR]

Garrod, S. & Terras, M. (2000) The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language* 42:526–44. [rKR]

Germeys, F., De Graef, P. & Verfaillie, K. (2002) Transsaccadic identification of saccade target and flanker objects. *Journal of Experimental Psychology: Human Perception and Performance* 28:868–83. [PDG]

Goldberg, M. E. (1994) The cortical control of saccadic eye movements. In: *Eye movements in reading,* ed. J. Ygge & G. Lennerstrand. Pergamon Press. [aEDR]

Goodale, M. A. & Milner, A. D. (1992) Separate visual pathways for perception and action. *Trends in Neurosciences* 15:20–25. [aEDR]

Grainger, J. (2003) Moving eyes and reading words: How can a computational model combine these two? In: *The mind's eye: Cognition and applied aspects of eye movement research,* ed. J. Hyönä, R. Radach & H. Deubel. Elsevier. [LH]

Grainger, J. & Jacobs, A. M. (1996) Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review* 103(3):518–65. [SA, LH]

(1998) On localist connectionism and psychological science. In: *Localist connectionistic approaches to human cognition,* pp. 1–38, ed. J. Grainger & A. M. Jacobs. Erlbaum. [LH]

Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y. & Malach, R. (1998) A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping* 6:316–28. [aEDR]

Guitton, D., Buchtel, H. A. & Douglas, R. M. (1985) Frontal lobe lesions in man cause difficulties in suppressing reflexive glances and in generating goal-directed saccades. *Experimental Brain Research* 58:455–72. [BF]

Hallett, P. (1978) Primary and secondary saccades to goals defined by instructions. *Vision Research* 18:1279–96. [BF]

Hays, W. L. (1988) *Statistics,* 4th edition. Holt, Rinehart & Winston. [GF]

Heinze, H. J., Mangun, G. R., Burchert, W., Hinrichs, H., Scholz, M., Munte, T. F., Gos, A., Scherg, M., Johannes, S., Hundeshagen, H., Gazzaniga, M. S. & Hillyard, S. A. (1994) Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature* 372:543–46. [CB]

Heller, D. & Radach, R. (1999) Eye movements in reading: Are two eyes better than one? In: *Current oculomotor research: Physiological and psychological aspects,* ed. W. Becker, H. Deubel & T. Mergner. Plenum Press. [RS]

Henderson, J. M. & Ferreira, F. (1990) Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16:417–29. [JP, EMR, aEDR, GU]

(1993) Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology* 47:201–21. [AK, aEDR, rKR]

Henderson, J. M., Weeks, P. A., Jr. & Hollingworth, A. (1999) The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance* 25(1):210–28. [CB, PDG]

Herdman, C. M., Chernecki, D. & Norris, D. (1999) Naming case alternated words. *Memory and Cognition* 27:254–66. [EMR]

Hillyard, S. A. & Anllo-Vento, L. (1998) Event-related brain potentials of visual selective attention. *Proceedings of the National Academy of Sciences USA* 95:781–87. [CB]

Hintzman, D. L. (1984) MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments and Computers* 16:96–101. [rKR]

(1986) "Schema abstraction" in a multiple-trace memory model. *Psychological Review* 93:411–28. [rKR]

(1987) Recognition and recall in MINERVA 2: Analysis of the "recognition-failure" paradigm. In: *Modeling cognition,* pp. 215–29, ed. P. Morris. Wiley. [rKR]

(1988) Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review* 95:528–51. [rKR]

(1990) Human learning and memory: Connections and dissociations. *Annual Review of Psychology* 41:109–39. [rKR]

(1991) Why are formal models useful in psychology? In: *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock,* ed. W. E. Hockley & S. Lewandowsky. Erlbaum. [aEDR]

Hoffman, J. E. & Subramaniam, B (1995) The role of visual attention in saccadic eye movements. *Perception and Psychophysics* 57:787–95. [CB, aEDR]

Horwitz, B., Rumsey, J. M. & Donohue, B. C. (1998) Functional connectivity of the angular gyrus in normal reading and dyslexia. *Proceedings of the National Academy of Sciences USA* 95:8939–44. [aEDR]

Huey, E. B. (1908) *The psychology and pedagogy of reading.* Macmillan. [aEDR, rKR]

Humphreys, M. S., Bain, J. D. & Pike, R. (1989) Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review* 96:208–33. [SA]

Hyönä, J. (1995a) An eye movement analysis of topic-shift effect during repeated reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21:1365–73. [SPL]

(1995b) Do irregular letter combinations attract readers' attention? Evidence from fixation locations in words. *Journal of Experimental Psychology: Human Perception and Performance* 2:68–81. [aEDR, GU]

(2000) Processing of Finnish compound words in reading. In: *Reading as a perceptual process,* ed. A. Kennedy, R. Radach, D. Heller, & J. Pynte. Elsevier. [aEDR]

Hyönä, J. & Bertram, R. (in press) Do frequency characteristics of non-fixated words influence the processing of fixated words during reading? *European Journal of Cognitive Psychology.* [rKR]

Hyönä, J., Niemi, P. & Underwood, G. (1989) Reading long words embedded in sentences: Informativeness of word halves affects eye movements. *Journal of Experimental Psychology: Human Perception and Performance* 15:142–52. [aEDR, GU]

Hyönä, J. & Olson, R. K. (1995) Eye movement patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21:1430–40. [aEDR]

Hyönä, J. & Pollatsek, A. (1998) Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance* 24:1612–27. [JP, aEDR]

(1999) The role of component morphemes on eye fixations when reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance* 24:1612–27. [JH, rKR]

Ikeda, M. & Saida, S. (1978) Span of recognition in reading. *Vision Research* 18:83–88. [aEDR]

Inhoff, A. W. (1984) Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior* 23:612–24. [aEDR]

(1989a) Lexical access during eye fixations in reading: Are word access codes used to integrate lexical information across interword fixations? *Journal of Memory and Language* 28:444–61. [AWI]

(1989b) Parafoveal processing of words and saccade computation during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance* 15:544–55. [AWI]

(1990) Integrating information across eye fixations in reading. *Acta Psychologica* 73:281–97. [AWI]

Inhoff, A. W. & Briihl, D. (1991) Semantic processing of unattended text during selective reading: How the eyes see it. *Perception and Psychophysics* 49:289–94. [aEDR]
Inhoff, A. W., Briihl, D. & Schwartz, J. (1996) Compound word naming in reading, on-line naming and delayed naming tasks. *Memory and Cognition* 24:466–76. [GU]
Inhoff, A. W. & Liu, W. (1998) The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance* 24:20–34. [aEDR]
Inhoff, A. W., Pollatsek, A., Posner, M. I. & Rayner, K. (1989) Covert attention and eye movements during reading. *Quarterly Journal of Experimental Psychology* 41A:63–89. [aEDR]
Inhoff, A. W. & Radach, R. (1998) Definition and computation of oculomotor measures in the study of cognitive processes. In: *Eye guidance in reading and scene perception,* ed. G. Underwood. Elsevier. pp. 1–28 cited by [PDG]; pp. 29–53 by [GF, aEDR]
Inhoff, A. W., Radach, R., Eiter, B. & Juhasz, B. (2003) Distinct subsystems for the parafoveal processing of spatial and linguistic information during eye fixations in reading. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 56A:803–28. [AWI]
Inhoff, A. W., Radach, R., Starr, M. S. & Greenberg, S. (2000a) Allocation of visuo-spatial attention and saccade programming during reading. In: *Reading as a perceptual process,* ed. A. Kennedy, R. Radach, D. Heller & J. Pynte. Elsevier. [aEDR, rKR]
Inhoff, A. W. & Rayner, K. (1986) Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics* 40(6):431–39. [KDM, aEDR, rKR]
Inhoff, A. W., Starr, M. & Shindler, K. L. (2000b) Is the processing of words during eye fixations in reading strictly serial? *Perception and Psychophysics* 62(7):1474–84. [CB, AWI, MP, aEDR]
Inhoff, A. W. & Topolski, R. (1992) Lack of semantic activation from unattended text during passage reading. *Bulletin of the Psychonomic Society* 30:365–66. [aEDR]
Irwin, D. E. (1998) Lexical processing during saccadic eye movements. *Cognitive Psychology* 36:1–27. [aEDR]
Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L. & Haxby, J. V. (1999) Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences USA* 96:9379–84. [aEDR]
Ishida, T. & Ikeda, M. (1989) Temporal properties of information extraction in reading studied by a text-mask replacement technique. *Journal of the Optical Society A: Optics and Image Sciences* 6:1624–32. [aEDR, rKR]
Itti, L. & Koch, C. (2000) A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research* 40:1489–506. [CB]
Jacobs, A. M. & Grainger, J. (1994) Models of visual word recognition – sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance* 20:1311–34. [rKR]
Juhasz, B., Starr, M. S., Inhoff, A. W. & Placke, L. (2003) The effects of morphology on the processing of compound words: Evidence from naming, lexical decision, and eye fixations. *British Journal of Psychology* 94:233–44. [aEDR, rKR]
Just, M. A. & Carpenter, P. A. (1980) A theory of reading: From eye fixations to comprehension. *Psychological Review* 87:329–54. [aEDR]
(1987) *The psychology of reading and language comprehension.* Allyn and Bacon. [aEDR]
(1992) A capacity theory of comprehension: Individual differences in working memory. *Psychological Review* 99:122–49. [aEDR]
Kajii, N., Nazir, T. A. & Osaka, N. (2001) Eye movement control in reading unspaced text: The case of the Japanese script. *Vision Research* 41:2503–10. [NO]
Kambe, G. (in press) Parafoveal processing of prefixed words during eye fixations in reading: Evidence against morphological influences on parafoveal preprocessing. *Perception and Psychophysics.* [aEDR]
Kambe, G., Rayner, K. & Duffy, S. A. (2001) Global context effects on processing lexically ambiguous words: Evidence from eye fixations. *Memory and Cognition* 29:363–72. [aEDR]
Kennedy, A. (1998) The influence of parafoveal words on foveal inspection time: Evidence for a processing trade-off. In: *Eye guidance in reading and scene perception,* ed. G. Underwood. Elsevier. [AK, aEDR]
(2000a) Attention allocation in reading. In: *Reading as a perceptual process,* ed. A. Kennedy, R. Radach, D. Heller & J. Pynte. Elsevier. [AK]
(2000b) Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology* 53A:429–55. [AK]
Kennedy, A., Brooks, R., Flynn, L. & Prophet, C. (2003) The reader's spatial code. In: *The mind's eye: Cognitive and applied aspects of eye movement research,* ed. J. Hyönä, R. Radach & H. Deubel. Elsevier. [AK, MP]
Kennedy, A. & Murray, W. S. (1987) Spatial coding and reading: Some comments on Monk (1985). *Quarterly Journal of Experimental Psychology* 39A:649–718. [AK]
Kennedy, A., Murray, W. S. & Boissiere, C. (in press) Parafoveal pragmatics revisited. *European Journal of Cognitive Psychology.* [WSM]
Kennedy, A., Pynte, J. & Ducrot, S. (2002) Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology* 55A(4):1307–38. [CB, AK, aEDR, rKR]
Kennedy, A., Radach, R., Heller, D. & Pynte, J. (2000) *Reading as a perceptual process.* Elsevier. [aEDR]
Kennison, S. M. & Clifton, C. (1995) Determinants of parafoveal perview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21:68–81. [aEDR, EMR]
Kim, Y.-H., Gitelman, D. R., Nobre, A. C., Parrish, T. B., LaBar, K. S. & Mesulam, M. M. (1999) The large-scale neural network for spatial attention displays multifunctional overlap but differential asymmetry. *NeuroImage* 9:269–77. [aEDR]
Kinsbourne, M. (1970) A model for the mechanism of unilateral neglect of space. *Transactions of the American Neurological Association* 95:143–46. [RS]
Klein, C., Fischer, B., Hartnegg, K., Heiss, W. H. & Roth, M. (2000) Optomotor and neuropsychological performance in old age. *Experimental Brain Research* 135:141–54. [BF]
(1977) Hemineglect and hemispheric rivalry. *Advances in Neurology* 18:41–49. [RS]
Kliegl, R. & Engbert, R. (2003) SWIFT explorations: In: *The mind's eye: Cognitive and applied aspects of eye movement research,* pp. 391–411, ed. J. Hyönä, R. Radach & H. Deubel. Elsevier. [RE, aEDR, rKR]
Kliegl, R., Grabner, E., Rolfs, M. & Engbert, R. (in press) Length, frequency, and predictablity effects of word on eye movements in reading. *European Journal of Cognitive Psychology.*
Kliegl, R. M., Olson, R. K. & Davidson, B. J. (1982) Regression analyses as a tool for studying reading processes: Comment on Just and Carpenter's eye fixation theory. *Memory and Cognition* 10:287–96. [GF]
Klitz, T. S., Legge, G. E. & Tjan, B. S. (2000) Saccade planning in reading with central scotomas: Comparison of human and ideal performance. In: *Reading as a perceptual process,* ed. A. Kennedy, R. Radach, D. Heller & J. Pynte. Elsevier. [aEDR]
Kolers, P. A. (1976) Buswell's discoveries. In: *Eye movements and psychological processes,* ed. R. A. Monty & J. W. Senders. Erlbaum. [aEDR]
LaBerge, D. L. (1990) Attention. *Psychological Science* 1:156–62. [aEDR]
Lamme, V. A. F. & Roelfsema, P. R. (2000) The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences* 23:571–79. [aEDR]
Lavigne, F., Vitu, F. & d'Ydewalle, G. (2000) The influence of semantic context on initial eye landing sites in words. *Acta Psychologia* 104:191–204. [aEDR, FV]
Lee, H., Rayner, K. & Pollatsek, A. (1999) The time course of phonological, semantic, and orthographic coding in reading: Evidence from the fast-priming technique. *Psychonomic Bulletin and Review* 6:624–34. [rKR]
Leff, A. P., Scott, S. K., Rothwell, J. C. & Wise, R. J. S. (2001) The planning and guiding of reading saccades: A repetitive transcranial magnetic stimulation study. *Cerebral Cortex* 11:918–23. [aEDR]
Legge, G. E., Klitz, T. S. & Tjan, B. S. (1997) Mr. Chips: An ideal-observer model of reading. *Psychological Review* 104:524–53. [aEDR]
Leigh, R. J. & Zee, D. S. (1999) *The neurology of eye movements.* Oxford University Press. [aEDR]
Lima, S. D. (1987) Morphological analysis in sentence reading. *Journal of Memory and Language* 26:84–99. [aEDR]
Lima, S. D. & Inhoff, A. W. (1985) Lexical access during eye fixations in reading: Effects of word-initial letter sequences. *Journal of Experimental Psychology: Human Perception and Performance* 11:272–85. [AWI,aEDR]
Liversedge, S. P. & Findlay, J. M. (2000) Saccadic eye movements and cognition. *Trends in Cognitive Science* 4:6–14. [aEDR]
Liversedge, S. & Underwood, G. (1998) Foveal processing load and landing position effects in reading. In: *Eye guidance in reading and scene perception,* pp. 201–21, ed. G. Underwood. Elsevier. [GU]
Luna, B., Thulborn, K. R., Strojwas, M. H., McCurtain, B. J., Berman, R. A., Genovese, C. R. & Sweeney, J. A. (1998) Dorsal cortical regions subserving visually guided saccades in humans: An fMRI study. *Cerebral Cortex* 8:40–47. [aEDR]
MacWhinney, B. (1995) *The CHILDES project: Tools for analyzing talk,* 2nd edition. Erlbaum. [RRe]
Mandler, G. (1980) Recognizing: The judgement of previous occurrence. *Psychological Review* 87:252–71. [rKR]
Mangun, G. R., Hopfinger, J., Kussmaul, C. L., Flechter, E. & Heinze, H. J. (1997) Covariations in ERP and PET measures of spatial selective attention in human extrastriate cortex. *Human Brain Mapping* 5:273–79. [CB]
Martinez, A., DiRusso, F., Anllo-Vento, L., Sereno, M. L., Buxton, R. B. & Hillyard, S. A. (2001) Putting spatial attention on the map: Timing and localization of stimulus selection processes in striate and extrastriate visual areas. *Vision Research* 41:1437–57. [CB]

Maunsell, J. H. R. & Newsome, W. T. (1987) Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience* 10:363–401. [aEDR]
Mayall, K. A., Humphreys, G. W., Mechelli, A., Olson, A. & Price, C. J. (2001) The effects of case mixing on word recognition: Evidence from a PET study. *Journal of Cognitive Neuroscience* 13:844–53. [EMR]
Mayall, K. A., Humphreys, G. W. & Olson, A. (1997) Disruption to word or letter processing? The origins of case mixing effects. *Journal of Experimental Psychology: Learning, Memory and Cognition* 23:1275–86. [EMR]
McClelland, J. L. (1986) The programmable blackboard model of reading. In: *Parallel distributed: Explorations in the microstructure of cognition, vol. 2: Psychological and biological models,* ed. J. L McClelland, D. E. Rumelhart & the PDP Research Group. MIT Press. [aEDR]
McClelland, J. L. & Rumelhart, D. E. (1981) An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review* 88:375–407. [CB, GER, aEDR]
McClelland, J. L., Rumelhart, D. E. & the PDP Research Group (1986) *Parallel distributed: Explorations in the microstructure of cognition, vol. 2: Psychological and biological models.* MIT Press. [aEDR]
McConkie, G. W. (1979) On the role and control of eye movements in reading. In: *Processing of visible language: Vol. 1,* pp. 37–48, ed. P. A. Kolers, M. E. Wrolstad & H. Bouma. Plenum Press. [AK, aEDR]
McConkie, G. W. & Dyre, B. P. (2000) Eye fixation durations in reading: Models of frequency distributions. In: *Reading as a perceptual process,* ed. A. Kennedy, R. Radach, D. Heller & J. Pynte. Elsevier. [GF, aEDR]
McConkie, G. W., Kerr, P. W. & Dyre, B. P. (1994) What are "normal" eye movements during reading: Toward a mathematical description. In: *Eye movements in reading,* ed. J. Ygge & G. Lennerstrand. Pergamon Press. [aEDR]
McConkie, G. W., Kerr, P. W., Reddix, M. D. & Zola, D. (1988) Eye movement control during reading: I. The location of initial eye fixations on words. *Vision Research* 28:1107–18. [aEDR, rKR]
McConkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D. & Jacobs, A. M. (1989) Eye movement control during reading: II. Frequency of refixating a word. *Perception and Psychophysics* 46:245–53. [KDM, JP, aEDR]
McConkie, G. W. & Rayner, K. (1975) The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics* 17:578–86. [NO, aEDR, rKR]
(1976) Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society* 8:365–68. [aEDR]
McConkie, G. W. & Yang, S-N. (2003) How cognition affects eye movements during reading. In: *The mind's eye: Cognitive and applied aspects of eye movement research,* ed. J. Hyönä, R. Radach & H. Deubel. Elsevier. [GWM]
McConkie, G. W. & Zola, D. (1987) Visual attention during eye fixations while reading. In: *Attention and performance XII: The psychology of reading,* pp. 385–401, ed. M. Coltheart. Erlbaum. [AK]
(1979) Is visual information integrated across successive fixations in reading? *Perception and Psychophysics* 25:211–24. [aEDR]
(1984) Eye movement control during reading: The effect of words units. In: *Cognition and motor processes,* ed. W. Prinz & A. F. Sanders. Springer-Verlag. [aEDR]
McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R. & Wolff, P. M. (1991) Children's eye movements during reading. In: *Vision and visual dyslexia,* ed. J. F. Stein. Macmillan Press. [aEDR, rKR]
McPeek, R. M., Skavenski, A. A. & Nakayama, K. (2000) Concurrent processing of saccades in visual search. *Vision Research* 40:2499–516. [aEDR]
Meseguer, E., Carreiras, M. & Clifton, C. J. (2002) Overt reanalysis strategies and eye movements during reading on mild garden sentences. *Memory and Cognition* 30:551–61. [MP, rKR]
Mesulam, M. M. (1990) Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology* 28:597–613. [aEDR]
(1998) From sensation to cognition. *Brain* 121:1013–52. [aEDR]
Molker, A. & Fischer, B. (1999) The recognition and correction of involuntary prosaccades in an antisaccade task. *Experimental Brain Research* 125:551–16. [aEDR]
Monaghan, P., Shillcock, R. & McDonald, S. (2003) Hemispheric asymmetries in the split-fovea model of semantic processing. *Brain and Language* 86:83–98. [RS]
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. [GER]
Morris, R. K., Rayner, K. & Pollatsek, A. (1990) Eye movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance* 16:268–81. [aEDR, rKR]
Morrison, R. E. (1984) Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance* 10:667–82. [RE, DH, GWM, NO, RRa, aEDR, rKR]
Morrison, R. E. & Rayner, K. (1981) Saccade size in reading depends upon character spaces and not visual angle. *Perception and Psychophysics* 30:395–96. [aEDR]
Motter, B. C. & Mountcastle, V. B. (1980) Active directed gaze controls the excitability of the light sensitive neurons of the inferior parietal lobule in the waking monkey. *Neuroscience Abstracts* 6, No. 230:14673. [BF]
Müller, R.-A., Rothermel, R. D., Behen, M. E., Muzik, O., Mangner, T. J. & Chugani, H. T. (1997) Receptive and expressive language activations for sentences: A PET study. *NeuroReport* 8:3767–70. [aEDR]
Munoz, D. P. & Wurtz, R. H. (1992) Role of the rostral superior colliculus in active visual fixation and execution of express saccades. *Journal of Neurophysiology* 67:1000–1002. [BF]
Murray, W. S. (1998) Parafoveal pragmatics. In: *Eye guidance in reading and scene perception,* pp. 181–99, ed. G. Underwood. Elsevier. [WSM, aEDR]
Murray, W. S. & Clayes, E. (1998) Plausibility and syntactic complexity. Paper presented at Architectures and Mechanisms for Language Processing Conference, Freiburg, September 1998. [WSM]
Murray, W. S., Dobbin, C., Murray, A. M. & Clayes, E. (1999) Plausibility and sentence interpretation. Paper presented at Architectures and Mechanisms for Language Processing Conference, Edinburgh, September 1999. [WSM]
Murray, W. S. & Kennedy, A. (1988) Spatial coding and the processing of anaphor by good and poor readers: Evidence from eye movement analyses. *Quarterly Journal of Experimental Psychology* 40A:693–718. [AK]
Murray, W. S. & Liversedge, S. P. (1994) Referential context effects on syntactic processing. In: *Perspectives on sentence processing,* pp. 359–88, ed. C. Clifton, L. Frazier & K. Rayner. Erlbaum. [SPL, WSM]
Murray, W. S. & Rowan, M. (1998) Early, mandatory, pragmatic processing. *Journal of Psycholinguistic Research (CUNY Special Issue)* 27:1–22. [WSM]
Newell, A. (1990) *Unified theories of cognition.* Harvard University Press. [aEDR]
Ngo, T. T., Miller, S. M., Liu, G. B. & Pettigrew, J. D. (2000) Binocular rivalry and perceptual coherence. *Current Biology* 10:134–36. [RS]
Niswander, E., Pollatsek, A. & Rayner, K. (2000) The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes* 15:389–420. [aEDR, rKR]
Nobre, A. C., Gitelman, D. R., Dias, E. C. & Mesulam, M. M. (2000) Covert visual spatial orienting and saccades: Overlapping neural Systems *NeuroImage* 11:210–16. [CB]
Oliva, A. & Torralba, A. (2001) Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision* 42:145–75. [PDG]
Olson, R. K., Kliegl, R. & Davidson, B. J. (1983) Dyslexic and normal readers' eye movements. *Journal of Experimental Psychology: Human Perception and Performance* 9:816–25. [BF]
O'Regan, J. K. (1979) Eye guidance in reading: Evidence for linguistic control hypothesis. *Perception and Psychophysics* 25:501–509. [aEDR]
(1980) The control of saccade size and fixation duration in reading: The limits of linguistic control. *Perception and Psychophysics* 28:112–17. [aEDR]
(1983) Elementary perceptual and eye movement control processes in reading. In: *Eye movements in reading: Perceptual and language processes,* ed. K. Rayner. Academic Press. [aEDR]
(1990) Eye movements and reading. In: *Eye movements and their role in visual and cognitive processes,* pp. 395–453, ed. E. Kowler. Elsevier. [aEDR, rKR, RRe]
(1992a) Eye movements and reading. In: *Eye movements and visual cognition: Scene perception and reading,* pp. 333–54, ed. K. Rayner. Springer-Verlag. [JP]
(1992b) Optimal viewing position in words and the strategy-tactics theory of eye movements in reading. In: *Eye movements and visual cognition: Scene perception and reading,* ed. K. Rayner. Springer-Verlag. [aEDR, rKR]
O'Regan, J. K. & Lévy-Schoen, A. (1987) Eye movement strategy and tactics in word recognition and reading. In: *Attention and performance XII: The psychology of reading,* ed. M. Coltheart. Erlbaum. [aEDR]
O'Regan, J. K., Lévy-Schoen, A., Pynte, J. & Brugaillère, B. (1984) Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance* 10(2):250–57. [KDM, aEDR, rKR]
Osaka, M., Nishizaki, Y., Komori, M. & Osaka, N. (2002) Effect of focus in verbal working memory: Critical role of focus word in reading. *Memory and Cognition* 30:562–71. [NO]
Osaka, M., Osaka, N., Kondo, H., Morishita, M., Fukuyama, H., Aso, T. & Shibasaki, H. (2003) Neural basis of individual differences in working memory: An fMRI study. *NeuroImage* 18:789–97. [NO]
Osaka, N. (1992) Size of saccade and fixation duration of eye movements during reading: Psychophysics of Japanese text processing. *Journal of the Optical Society of America* A 9:5–13. [NO, aEDR]

Osaka, N. & Oda, K. (1994) Moving window for reading experiments. *Behavior Research Methods, Instruments and Computers* 26:49–53. [NO]

Osaka, N. & Osaka, M. (2002) Individual differences in working memory during reading with and without parafoveal information: A moving window study. *American Journal of Psychology* 115:501–13. [NO]

O'Shea, R. & Corballis, P. M. (2001) Binocular rivalry between complex stimuli in split-brain observers. *Brain and Mind* 2:151–60. [RS]

Ottes, F. P., van Gisbergen, J. A. M. & Eggermont, J. J. (1984) Metrics of saccade responses to visual double stimuli: Two different modes. *Vision Research* 24:1169–79. [RR]

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

Pacht, J., Schroyens, W. & d'Ydewalle, G. (1999) Inhibitory neighborhood effects in visual word recognition. Paper presented at the Tenth European Conference on Eye Movements, September, Utrecht, The Netherlands. [JP]

Pavlidis, G. T. (1981) Do eye movements hold the key to dyslexia? *Neuropsychologia* 19:57–64. [BF]

Perea, M. & Pollatsek, A. (1998) The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance* 24:767–79. [MP, aEDR]

Perfetti, C. A. & Bell, L. (1991) Phonemic activation during the first 40 ms of word identification: Evidence from backwards masking and priming. *Journal of Memory and Language* 30:473–85. [rKR]

Peterson, S. E., Fox, P. T., Posner, M. I., Mintun, M. & Raichle, M. E. (1989) Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature* 331:585–89. [aEDR]

Peterson, S. E., Fox, P. T., Snyder, A. Z. & Raichle, M. E. (1990) Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science* 249:1041–44. [aEDR]

Pickering, M. J. & Frisson, S. (2001) Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27:556–73. [rKR]

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

Poldrack, R. A., Desmond, J. E., Glover, G. H. & Gabrieli, J. D. E. (1998) The neural basis of visual skill learning: An fMRI study of mirror reading. *Cerebral Cortex* 8:1–10. [aEDR]

Pollatsek, A., Bolozky, S., Well, A. D. & Rayner, K. (1981) Asymmetries in the perceptual span for Israeli readers. *Brain and Language* 14:174–80. [aEDR]

Pollatsek, A., Hyönä, J. & Bertram, R. (2000) The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance* 26:820–33. [SA, JH, aEDR, rKR]

Pollatsek, A., Lesch, M., Morris, R. K. & Rayner, K. (1992) Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance* 18:148–62. [aEDR, rKR]

Pollatsek, A., Perea, M. & Binder, K. S. (1999a) The effects of neighborhood size in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance* 25:1142–58. [SA, MP, aEDR]

Pollatsek, A., Raney, G. E., LaGasse, L. & Rayner, K. (1993) The use of information below fixation in reading and in visual search. *Canadian Journal of Experimental Psychology* 47:179–200. [aEDR]

Pollatsek, A. & Rayner, K. (1982) Eye movement control in reading: The role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance* 8:817–33. [aEDR, rKR]

(1990) Eye movements and lexical access in reading. In: *Comprehension processes in reading,* pp. 143–63, ed. D. A. Balota, G. B. Flores d'Arcais & K. Rayner. Erlbaum. [JP, aEDR]

(1999) Is covert attention really unnecessary? *Behavioral and Brain Sciences* 22:695–96. [AK, aEDR]

Pollatsek, A., Rayner, K. & Balota, D. A. (1986) Inferences about eye movement control from the perceptual span in reading. *Perception and Psychophysics* 40:123–30. [aEDR, rKR]

Pollatsek, A., Rayner, K., Fischer, M. H. & Reichle, E. D. (1999b) Attention and eye movements in reading. In: *Reading and dyslexia: Visual and attentional processes,* ed. J. Everatt. Routledge. [aEDR]

Pollatsek, A., Rayner, R. & Lee, H. W. (2000) Phonological coding in word perception and reading. In: *Reading as a perceptual process,* pp. 399–426, ed. A. Kennedy, R. Radach, D. Heller & J. Pynte. Elsevier. [LH, rKR]

Pollatsek, A., Reichle, E. D. & Rayner, K. (2003) Modeling eye movements in reading: Extensions of the E-Z Reader model. In: *The mind's eyes: Cognitive and applied aspects of oculomotor research,* pp. 361–90, ed. J. Hyönä, R. Radach & H. Deubel. Elsevier. [aEDR, rKR]

Posner, M. I. (1980) Orienting of attention. *Quarterly Journal of Experimental Psychology* 32:3–25. [aEDR]

Posner, M. I., Aldullaev, Y. G., McCandliss, B. D. & Sereno, S. C. (1999a) Anatomy, circuitry and plasticity of word reading. In: *Reading and dyslexia: Visual and attentional processes,* ed. J. Everatt. Routledge. [aEDR]

(1999b) Neuroanatomy, circuitry and plasticity of word reading. *NeuroReport* 10:R12–R23. [aEDR]

Posner, M. I. & Raichle, M. I. (1997) *Images of mind.* Scientific American Library. [aEDR]

Price, C. J. (1997) Functional anatomy of reading. In: *Human brain function,* ed. R. S. J. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan & J. C. Mazziotta. Academic Press. [aEDR]

Pugh, K. R., Mencl, W. E., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Constable, R. T., Skudlarski, P., Marchione, K. E., Jenner, A. R., Fletcher, J. M., Liberman, A. M., Shankweiler, D. P., Katz, L., Lacadie, C. & Gore, J. C. (2000) The angular gyrus in developmental dyslexia: Task-specific differences in functional connectivity with posterior cortex. *Psychological Science* 11:51–56. [aEDR]

Pulvermüller, F., Assadollahi, R. & Elbert, T. (2001) Neuromagnetic evidence for early semantic access in word recognition. *European Journal of Neuroscience* 13:201–205. [SCS]

Pynte, J. (1996) Lexical control of within-word eye movements. *Journal of Experimental Psychology: Human Perception and Performance* 22:958–69. [JP]

Pynte, J., Kennedy, A. & Ducrot, S. (in press) The influence of parafoveal typographical errors on eye movements in reading. *European Journal of Cognitive Psychology.* [rKR]

Pynte, J., Kennedy, A., Murray, W. S. & Courrieu, P. (1988) The effects of spatialisation on the processing of ambiguous pronominal reference. In: *Eye movement research: Physiological and psychological aspects,* pp. 214–25, ed. G. Luer, U. Lass & J. Shallo-Hoffman. Hogrefe. [AK]

Raaijmakers, J. G. W. & Shiffrin, R. M. (1981) Search of associative memory. *Psychological Review* 88:93–134. [SA]

Radach, R. & Heller, D. (2000) Relations between spatial and temporal aspects of eye movement control. In: *Reading as a perceptual process,* ed. A. Kennedy, R. Radach, D. Heller & J. Pynte. Elsevier. [RRa, aEDR, rKR]

Radach, R., Inhoff, A. & Heller, D. (in press) Orthographic regularity gradually modulates saccade amplitudes in reading. *European Journal of Cognitive Psychology.* [SPL]

Radach, R. & Kempe, V. (1993) An individual analysis of initial fixation positions in reading. In: *Perception and cognition: Advances in eye movement research,* ed. G. d'Ydewalle & J. Van Rensbergen. North Holland. [aEDR]

Radach, R. & McConkie, G. W. (1998) Determinants of fixation positions in words during reading. In: *Eye guidance in reading and scene perception,* pp. 77–100, ed. G. Underwood. Elsevier. [AK, aEDR]

Rafal, R. & Robertson, L. (1995) The neurology of visual attention. In: *The cognitive neurosciences,* ed. M. S. Gazzaniga. MIT Press. [aEDR]

Raney, G. E., Atilano, R. & Gomez, L. (1996) Language representation and bilingual reading. Paper presented at the meeting of the Psychonomic Society, Chicago, IL, November 1996. [GER, rKR]

Raney, G. E. & Rayner, K. (1995) Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology* 49:151–72. [aEDR]

Raney, G. E., Therriault, D. & Minkoff, S. (2000) Repetition effects from paraphrased text: Evidence for an integrative model of text representation. *Discourse Processes* 28:61–81. [GER, rKR]

Rayner, K. (1975) The perceptual span and peripheral cues in reading. *Cognitive Psychology* 7:65–81. [JP, aEDR]

(1977) Visual attention in reading: Eye movements reflect cognitive processes. *Memory and Cognition* 4:443–48. [aEDR]

(1978) Eye movements in reading and information processing. *Psychological Bulletin* 85:618–60. [aEDR]

(1979) Eye guidance in reading: Fixation locations within words. *Perception* 8:21–30. [aEDR, rKR]

(1986) Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology* 41:211–36. [aEDR]

(1995) Eye movements and cognitive processes in reading, visual search, and scene perception. In: *Eye movement research: Mechanisms, processes and applications,* pp. 3–22, ed. J. M. Findlay, R. Walker & R. W. Kentridge. North Holland. [rKR]

(1998) Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124:372–422. [JP, aEDR, rKR]

Rayner, K., Ashby, J., Pollatsek, A. & Reichle, E. D. (2003a) The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. (submitted). [rKR]

Rayner, K., Balota, D. A. & Pollatsek, A. (1986) Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology* 40:473–83. [aEDR]

Rayner, K. & Bertera, J. H. (1979) Reading without a fovea. *Science* 206:468–69. [aEDR]

Rayner, K., Binder, K. S., Ashby, J. & Pollatsek, A. (2001) Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research* 41:943–54. [aEDR, LS]

Rayner, K., Carlson, M. & Frazier, L. (1983a) 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* 22:358–74. [SPL, rKR]

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

(1988) Online comprehension processes and eye movements in reading. In: *Reading research: Advances in theory and practice,* ed. M. Daneman, G. E. MacKinnon & T. G. Waller. Academic Press. [aEDR]

Rayner, K. & Fischer, M. H. (1996) Mindless reading revisited: Eye movements during reading and scanning are different. *Perception and Psychophysics* 58:734–47. [aEDR, rKR]

Rayner, K., Fischer, M. H. & Pollatsek, A. (1998a) Unspaced text interferes with both word identification and eye movement control. *Vision Research* 38:1129–144. [aEDR, rKR]

Rayner, K. & Frazier, L. (1989) Selection mechanisms in reading lexically ambiguous word. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:779–90. [aEDR]

Rayner, K., Inhoff, A. W., Morrison, R., Slowiczek, M. L. & Bertera, J. H. (1981) Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance* 7:167–79. [aEDR, rKR]

Rayner, K., Kambe, G. & Duffy, S. A. (2000) The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology* 53A:1061–80. [aEDR]

Rayner, K., Liversedge, S. P., White, S. J. & Vergilino-Perez, D. (2003b) Reading disappearing text: Cognitive control of eye movements. *Psychological Science* 14:385–89. [rKR]

Rayner, K. & McConkie, G. W. (1976) What guides a reader's eye movements. *Vision Research* 16:829–37. [aEDR, rKR]

Rayner, K. & Morris, R. K. (1992) Eye movement control in reading: Evidence against semantic preprocessing. *Journal of Experimental Psychology: Human Perception and Performance* 18:163–72. [aEDR, GU]

Rayner, K. & Morrison, R. E. (1981) Eye movements and identifying words in parafoveal vision. *Bulletin of the Psychonomic Society* 17:135–38. [aEDR, rKR]

Rayner, K. & Pollatsek, A. (1981) Eye movement control during reading: Evidence for direct control. *Quarterly Journal of Experimental Psychology* 33A:351–73. [aEDR]

(1987) Eye movements in reading: A tutorial review. In: *Attention and performance XII,* pp. 327–62, ed. M. Coltheart. Erlbaum. [rKR]

(1989) *The psychology of reading.* Prentice Hall. [aEDR, rKR]

(1996) Reading unspaced text is not easy: Comments on the implications of Epelboim et al.'s (1994) study for models of eye movement control in reading. *Vision Research* 36:461–65. [AK]

Rayner, K., Pollatsek, A. & Binder, K. S. (1998b) Phonological codes and eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24:476–97. [rKR]

Rayner, K. & Raney, G. E. (1996) Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin and Review* 3:238–44. [aEDR, rKR]

Rayner, K., Reichle, E. D. & Pollatsek, A. (1998c) Eye movement control in reading: An overview and model. In: *Eye guidance in reading and scene perception,* pp. 243–68, ed. G. Underwood. Elsevier. [MB, aEDR, rKR]

(2000) Eye movement control in reading: Updating the E-Z Reader model to account for initial fixation locations and refixations. In: *Reading as a perceptual process,* A. Kennedy, R. Radach, D. Heller, & J. Pynte. Elsevier. [aEDR]

Rayner, K. & Sereno, S. C. (1994) Eye movements in reading: Psycholinguistic studies. In: *Handbook of psycholinguistics,* ed. M. Gernsbacher. Academic Press. [aEDR]

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. [rKR, SCS]

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R. & Clifton, C. (1989) Eye movements and online language comprehension processes. *Language and Cognitive Processes* 4 [Special issue]: 21–49. [aEDR]

Rayner, K., Sereno, S. C. & Raney, G. E. (1996) Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance* 22(5):1188–200. [JH, aEDR, rKR, FV]

Rayner, K., Slowiaczek, M. L., Clifton, C. & Bertera, J. H. (1983b) Latency of sequential eye movements: Implications for reading. *Journal of Experimental Psychology: Human Perception and Performance* 9:912–22. [SPL, aEDR]

Rayner, K. & Well, A. D. (1996) Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin and Review* 3:504–509. [aEDR, rKR, LS]

Rayner, K., Well, A. D. & Pollatsek, A. (1980) Asymmetry of the effective visual field in reading. *Perception and Psychophysics* 27:537–44. [aEDR]

Rayner, K., Well, A. D., Pollatsek, A. & Bertera, J. H. (1982) The availability of useful information to the right of fixation in reading. *Perception and Psychophysics* 31:537–50. [PDG, aEDR]

Rayner, K., White, S. J., Kambe, G., Miller, B. & Liversedge, S. P. (2003c) On the processing of meaning from parafoveal vision during eye fixations in reading. In: *The mind's eye: Cognitive and applied aspects of eye movements,* ed. J. Hyona, R. Radach & H. Deubel. Elsevier. [WSM, aEDR, rKR]

Reichle, E. D & Perfetti, C. A. (2003) Morphology in word identification: A word-experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading.* 7:219–38. [rKR]

Reichle, E. D., Pollatsek, A., Fisher, D. L. & Rayner, K. (1998) Toward a model of eye movement control in reading. *Psychological Review* 105:125–57. [SA, RE, GF, aEDR, rKR, EMR]

Reichle, E. D. & Rayner, K. (2001) Cognitive processes and models of reading. In: *Models of the visual system,* ed. G. Hung & K. Ciufferda. Plenum. [aEDR]

Reichle, E. D., Rayner, K. & Pollatsek, A. (1999) Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research* 39:4403–411. [GF, JP, aEDR, rKR, EMR]

Reilly, R. (1993) A connectionist framework for modeling eye movement control in reading. In: *Perception and cognition: Advances in eye movement research,* ed. G. d'Ydewalle & Van Rensbergen. North Holland. [aEDR]

Reilly, R. & O'Regan, J. K. (1998) Eye movement control in reading: A simulation of some word-targeting strategies. *Vision Research* 38:303–17. [aEDR, RRe]

Reilly, R. & Radach, R. (2003) Foundations of an interactive activation model of eye movement control in reading. In: *The mind's eye: Cognitive and applied aspects of eye movements,* ed. J. Hyönä, R. Radach & H. Deubel. Elsevier. [JH, RRa, LH, aEDR, rKR, RRe]

Reingold, E. M. & Stampe, D. M. (1999) Saccadic inhibition in complex visual tasks. In: *Current oculomotor research: Physiological and psychological aspects,* pp. 249–55, ed. W. Becker & T. Mergner. Plenum. [GWM]

Reynolds J. H., Pasternak, T. & Desimone, R. (2000) Attention increases sensitivity of V4 neurons. *Neuron* 26:703–14. [rKR, SCS]

Robinson, D. L., Goldberg, M. E. & Stanton, G. B. (1978) Parietal association cortex in the primate: Sensory mechanisms and behavioral modulations. *Journal of Neurophysiology* 41:910–32. [BF]

Rousselet, G. A., Fabre-Thorpe, M. & Thorpe, S. J. (2002) Parallel processing in high-level categorization of natural images. *Nature Neuroscience* 5:629–30. [SCS]

Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986a) Learning internal representations by error propagation. In: *Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1: Foundations,* ed. D. E. Rumelhart, J. L. McClelland & the PDP Research Group. MIT Press. [aEDR]

Rumelhart, D. E., McClelland, J. L. & the PDP Research Group. (1986b) *Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1: Foundations.* MIT Press. [aEDR]

Sagi, D. & Julesz, B. (1985) "Where" and "what" in vision. *Science* 228:1217–19. [aEDR]

Salvucci, D. D. (2000) An interactive model-based environment for eye movement protocol visualization and analysis. In: *Proceedings of the Eye Tracking Research and Applications Symposium.* ACM Press. [aEDR]

(2001) An integrated model of eye movements and visual encoding. *Cognitive Systems Research* 1:201–20. [aEDR]

Salvucci, D. D. & Anderson, J. R. (1998) Tracing eye movement protocols with cognitive process models. *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society.* Erlbaum. [aEDR]

(2001)Automated eye movement protocol analysis. *Human Computer Interaction* 16:39–86. [aEDR]

Schall, J. D. & Hanes, K. G. (1993) Neural basis of target selection in frontal eye field during visual search. *Nature* 366:467–69. [CB]

Schiller, P. H., Sandell, J. H. & Maunsell, J. H. (1987) The effect of frontal eye field and superior colliculus lesions on saccadic latencies in the rhesus monkey. *Journal of Neurophysiology* 57:1033–49. [BF]

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

Schlosser, M. J., Aoyagi, N., Fulbright, R. K., Gore, J. C. & McCarthy, G. (1998) Functional MRI studies of auditory comprehension. *Human Brain Mapping* 6:1–13. [aEDR]

Schmidt, D. C. & Fayad, M. (1997) Object-oriented application frameworks. *Communications of the Association for Computing Machinery* 40:32–38. [RRe]

Schneider, W. X. & Deubel, H. (2002) Selection-for-perception and selection-for-spatial-motor-action are coupled by visual attention: A review of recent findings and new evidence for stimulus driven saccade control. In: *Attention and performance XIX: Common mechanisms in perception and action,* pp. 609–27, ed. W. Prinz & B. Hommel. Oxford University Press. [RRa]

Schreuder, R. & Baayen, R. H. (1995) Modelling morphological processing. In: *Morphological aspects of language processing,* pp. 131–54, ed. L. B. Feldman. Erlbaum. [JH]

Schroyens, W., Vitu, F., Brysbaert, M. & d'Ydewalle, G. (1999) Eye movement control during reading: Foveal load and parafoveal processing. *Quarterly Journal of Experimental Psychology* 52A(4):1021–46. [aEDR, FV]

Schustack, M. W., Ehrlich, S. F. & Rayner, K. (1987) The complexity of contextual facilitation in reading: Local and global influences. *Journal of Memory and Language* 26:322–40. [aEDR, rKR]

Seidenberg, M. S. (1989) Visual word recognition and pronunciation: A computational model and its implications. In: *Lexical representation and process,* ed. W. Marslen-Wilson. MIT Press. [aEDR]

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

Sereno, A. B. (1996) Parsing cognitive processes: Psychopathological and neurophysiological constraints. In: *Psychopathology: The evolving science of mental disorder,* pp. 407–32, ed. S. Matthysse, D. L. Levy, J. Kagan & F. M. Benes. Cambridge University Press. [SCS]

Sereno, S. C. (1992) Early lexical effects when fixating a word in reading. In: *Eye movements and visual cognition: Scene perception and reading,* ed. K. Rayner. Springer-Verlag. [aEDR]

(1995) Resolution of lexical ambiguity: Evidence from an eye movement-priming paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21:582–95. [GER, aEDR, SCS]

Sereno, S. C., Brewer, C. C. & O'Donnell, P. J. (2003) Context effects in word recognition: Evidence for early interactive processing. *Psychological Science* 14:328–33. [SCS]

Sereno, S. C., Pacht, J. M. & Rayner, K. (1992) The effect of meaning frequency on processing lexically ambiguous words: Evidence from eye fixations. *Psychological Science* 3:296–300. [aEDR, SCS]

Sereno, S. C. & Rayner, K. (1992) Fast priming during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance* 18:173–84. [SCS]

(2000) Spelling-sound regularity effects on eye fixations in reading. *Perception and Psychophysics* 62:402–409. [SCS]

Sereno, S. C., Rayner, K. & Posner, M. I. (1998) Establishing a time-line of processing during reading: Evidence from eye movements and event-related potentials. *NeuroReport* 9:2195–200. [RRa, aEDR, SCS]

Shillcock, R., Ellison, T. M. & Monaghan, P. (2000) Eye-fixation behavior, lexical storage, and visual word recognition in a split-processing model. *Psychological Review* 107:824–51. [aEDR, RS]

Shioiri, S. & Ikeda, M. (1989) Useful resolution for picture perception as a function of eccentricity. *Perception* 18:347–61. [PDG]

Sieroff, E., Pollatsek, A. & Posner, M. I. (1988) Recognition of visual letter strings following injury to the posterior visual spatial attention system. *Cognitive Neuropsychology* 5:427–49. [rKR]

Somers, D. C., Dale, A. M., Seiffert, A. E. & Tootell, R. B. H. (1999) Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences USA* 96:1663–68. [aEDR]

Sommer, M. & Schiller, P. (1992) Express saccades elicited during natural fixations of visual search. *Neuroscience Abstracts* 18, No. 587(9):1398. [BF]

Sparks, D. L. & Mays, L. (1990) Signal transformations required for the generation of saccadic eye movements. *Annual Review of Neuroscience* 13:309–36. [aEDR]

Starr, M. S. & Inhoff, A. W. (in press) Attention allocation to the right and left of a fixated work: Use of orthographic information from multiple words during reading. *European Journal of Cognitive Psychology.* [AWI, aEDR, rKR]

Starr. M. S. & Rayner, K. (2001) Eye movements during reading: Some current controversies. *Trends in Cognitive Science* 5:156–63. [MP, aEDR]

Stein, J. F. & Fowler, M. S. (1993) Unstable binocular control in dyslexic children. *Journal of Research in Reading* 16(1):30–45. [BF]

Stelmach, L. B., Campsall, J. M. & Herdman, C. M. (1997) Attentional and ocular movements. *Journal of Experimental Psychology: Human Perception and Performance* 23:823–44. [aEDR]

Sternberg, S. (1966) High-speed scanning in human memory. *Science* 153:652–54. [rKR]

(1969) The discovery of processing stages: Extensions of Donder's method. *Acta Psychologica* 30:276–315. [rKR]

(1975) Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology* 27:1–32. [rKR]

Stromswold, K., Caplin, D., Alpert, N. & Rauch, S. (1996) Localization of syntactic comprehension by Positron Emission Tomography. *Brain and Language* 52:452–73. [aEDR]

Suppes, P. (1990) Eye movement models for arithmetic and reading performance. In: *Eye movements and their role in visual and cognitive processes,* ed. E. Kowler. Elsevier. [aEDR]

(1994) Stochastic models of reading. In: *Eye movements in reading,* ed. J. Ygge & G. Lennerstrand. Pergamon Press. [aEDR]

Suppes, P., Cohen, M., Laddaga, R., Anliker, J. & Floyd, H. (1982) Research on eye movements in arithmetic performance. In: *Cognition and eye movements,* ed. R. Groner & P. Fraisse. North-Holland. [aEDR]

(1983) A procedural theory of eye movements in doing arithmetic. *Journal of Mathematical Psychology* 27:341–69. [aEDR]

Sweeney, J. A., Mintun, M. A., Kwee, S., Wiseman, M. B., Brown, D. L., Rosenberg, D. R. & Carl, J. R. (1996) Positron emission tomography study of voluntary saccadic eye movements and spatial working memory. *Journal of Neurophysiology* 75:454–68. [aEDR]

Tanaka, K. (1996) Inferotemporal cortex and object vision. *Annual Review of Neuroscience* 19:109–39. [aEDR]

Theios, J. & Amrhein, P. C. (1989) Theoretical analysis of the cognitive processing of lexical and pictorial stimuli: Reading, naming, and visual and conceptual comparisons. *Psychological Review* 96:5–24. [PDG]

Thibadeau, R., Just, M. A. & Carpenter, P. A. (1982) A model of the time course and content of reading. *Cognitive Science* 6:157–203. [aEDR]

Traxler, M. J. & Pickering, M. J. (1996) Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language* 35:454–75. [WSM]

Treisman, A. M. (1969) Strategies and models of selective attention. *Psychological Review* 76:242–99. [aEDR, rKR]

Treisman, A. M. & Gelade, G. (1980) A feature-integration theory of attention. *Cognitive Psychology* 12:97–136. [RRa]

Underwood, G., Bloomfield, R. & Clews, S. (1989) Information influences the pattern of eye fixations during sentence comprehension. *Perception* 17:267–78. [GU]

Underwood, G., Clews, S. & Everatt, J. (1990) How do readers know where to look next? Local information distributions influence eye fixations. *Quarterly Journal of Experimental Psychology* 42A:39–65. [aEDR, GU]

Underwood, N. R. & McConkie, G. W. (1985) Perceptual span for letter distinctions during reading. *Reading Research Quarterly* 20:153–62. [aEDR]

Ungerleider, L. G., Courtney, S. M. & Haxby, J. V. (1998) A neural system for human visual working memory. *Proceedings of the National Academy of Science* 95:883–90. [aEDR]

Ungerleider, L. G. & Haxby, J. V. (1994) "What" and "where" in the human brain. *Current Opinion in Neurobiology* 4:157–65. [aEDR]

Ungerleider, L. G. & Mishkin, M. (1982) Two cortical visual systems. In: *Analysis of visual behavior,* ed. D. J. Ingle, M. A. Goodale & R. J. W. Mansfield. MIT Press. [aEDR]

Uttal, W. R. (1990) On some two-way barriers between models and mechanisms. *Perception and Psychophysics* 48:188–203. [rKR]

Van Essen, D. C. & De Yoe, E. A. (1995) Concurrent processing in the primate visual cortex. In: *The cognitive neurosciences,* ed. M. S. Gazzaniga. MIT Press. [aEDR]

Vecera, S. P. & Farah, M. J. (1997) Is visual image segmentation a bottom-up or an interactive process? *Perception and Psychophysics* 59:1280–96. [PDG]

Vergilino, D. & Beauvillain, C. (2000) The planning of refixation saccades in reading. *Vision Research* 40(25):3527–38. [KDM, JP, aEDR]

Vergilino-Perez, D. & Beauvillain, C. (in press) The ability of the saccadic system to change motor plans in scanning letter strings. *Psychonomic Bullletin and Review.* [KDM]

Vitu, F. (1991a) The existence of a center of gravity effect during reading. *Vision Research* 31(7/8):1289–313. [FV]

Vitu, F. (1991b) The influence of parafoveal processing and linguistic context on the optimal landing position effect. *Perception and Psychophysics* 50:58–75. [aEDR]

Vitu, F., Brysbaert, M. & Lancelin, D. (in press) A test of parafoveal-on-foveal effects with pairs of orthographically related words. *European Journal of Cognitive Psychology.* [FV]

Vitu, F. & McConkie, G. W. (2000) Regressive saccades and word perception in adult reading. In: *Reading as a perceptual process,* pp. 301–26, ed. A. Kennedy, R. Radach, D. Heller & J. Pynte. Elsevier. [FV]

Vitu, F., McConkie, G. W., Kerr, P. & O'Regan, J. K. (2001) Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision Research* 41:3513–33. [JH, aEDR, rKR, FV]

Vitu, F., McConkie, G. W. & Zola, D. (1998) About regressive saccades in reading and their relation to word identification. In: *Eye guidance in reading and scene perception,* ed. G. Underwood. Elsevier. [FV]

Vitu, F., O'Regan, J. K., Inhoff, A. W. & Topolski, R. (1995) Mindless reading: Eye-movement characteristics are similar in scanning letter strings and reading texts. *Perception and Psychophysics* 57(3):352–64. [GWM, FV]

Vonk, W., Radach, R. & van Rijn, H. (2000) Eye guidance and the saliency of word beginnings in reading text. In: *Reading as a perceptual process,* ed. A. Kennedy, R. Radach, D. Heller & J. Pynte. North Holland/Elsevier. [aEDR, GU]

Walker, R., Kentridge, R. W. & Findlay, J. M. (1995) Independent contributions of the orienting of attention, fixation offset and bilateral stimulation on human saccadic latencies. *Experimental Brain Research* 103:294–310. [JMF]

Watt, S. M. & Murray, W. S. (1996) Prosodic form and parsing commitments. *Journal of Psycholinguistic Research (CUNY Special Issue)* 25:291–318. [AK]

Wenban-Smith, M. G. & Findlay, J. M. (1991) Express saccades: Is there a separate population in humans? *Experimental Brain Research* 87:218–22. [JMF]

White, S. J. & Liversedge, S. P. (in press) Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology.* [SPL, rKR]

(2003) Orthographic regularity influences the eyes' landing position in reading. (submitted). [SPL]

Wiley, J. & Rayner, K. (2000) Effects of titles on the processing of text and lexically ambiguous words: Evidence from eye movements. *Memory and Cognition* 28:1011–21. [aEDR, GER]

Woldorff, M. G., Fox, P. T., Matzke, M., Lancaster, J. L., Veeraswamy, S., Zamarripa, F., Seabolt, M., Glass, T., Gao, J. H., Martin, C. C. & Jerabek, P. (1997) Retinotopic organization of the early visual-spatial attention effects as revealed by PET and ERPs. *Human Brain Mapping* 5:280–86. [CB]

Wolfe, J. M. & Bennett, S. C. (1996) Preattentive object files: Shapeless bundles of basic features. *Vision Research* 37:25–43. [aEDR]

Wolverton, G. S. & Zola, D. (1983) The temporal characteristics of visual information extraction during reading. In: *Eye movements in reading: Perceptual and language processes,* ed. K. Rayner. Academic Press. [aEDR]

Yang, S.-N. & McConkie, G. W. (2001) Eye movements during reading: A theory of saccade initiation time. *Vision Research* 41:3567–85. [aEDR, rKR]

(in press) Saccade generation during reading: Are words necessary? *European Journal of Cognitive Psychology.* [GWM, rKR]

Yeshurun, Y. & Carrasco, M. (1998) Attention improves or impairs visual performance by enhancing spatial resolution. *Nature* 396:72–75. [SCS, rKR]

Yonelinas, A. P. (2002) The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language* 46:441–517. [rKR]

Zola, D. (1984) Redundancy and word perception during reading. *Perception and Psychophysics* 36:277–84. [aEDR]