



Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control

Alexander Pollatsek^{a,*}, Erik D. Reichle^b, Keith Rayner^c

^a Department of Psychology, University of Massachusetts, Amherst, MA 01003, USA

^b University of Pittsburgh

^c University of Massachusetts

Accepted 10 June 2005

Available online 9 November 2005

Abstract

This paper is simultaneously a test and refinement of the E-Z Reader model and an exploration of the interrelationship between visual and language processing and eye-movements in reading. Our modeling indicates that the assumption that words in text are processed serially by skilled readers is a viable and attractive hypothesis, as it accounts not only for “normal” reading data, but also for the pattern of decrements that occur when text is withheld from view during certain periods of time, such as in the boundary paradigm and the disappearing text paradigm. Our analyses also indicate (a) that lexical processing during reading is essentially continuous and (b) that the trigger for eye movements has to be a different (and prior) event to the trigger for shifts of covert spatial attention. In addition, the parameter values in our model are in accordance with what is known about such processes and should be taken as serious hypotheses for how long these processes last. Although a parallel model may be able to duplicate the predictions of the E-Z Reader model, we think that it is likely to be far less parsimonious and not nearly as good a heuristic device for using eye movements to understand language processing in reading.
© 2005 Elsevier Inc. All rights reserved.

Keywords: Eye movements; Reading; Modeling

* Corresponding author.

E-mail address: pollatsek@psych.umass.edu (A. Pollatsek).

1. Introduction

The E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998; Pollatsek, Reichle, & Rayner, 2003; Rayner, Reichle, & Pollatsek, 2005; Reichle, Rayner, & Pollatsek, 1999, 2003), a quantitative model of skilled reading, has been fairly influential in helping to formulate questions about how cognitive processes control the movement of the eyes in skilled reading, and by extension, how they might do so in other perceptual/cognitive tasks. Our modeling efforts have also inspired the development of competing models (*SWIFT*: Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005; *EMMA*: Salvucci, 2001; *Glenmore*: Reilly & Radach, 2003; *Competition/Interaction* model: Yang & McConkie, 2001; *SERIF*: McDonald, Carpenter, & Shillcock, 2005) and several experiments that have been designed as tests of the model (e.g., see Inhoff, Eiter, & Radach, 2005; Rayner, Ashby, Pollatsek, & Reichle, 2004; Reichle, Tokowicz, Liu, & Perfetti, 2005; Reingold & Rayner, 2005). In our opinion, the model has come out quite well in these tests. One of the reasons that we think the model has had this wide an impact is that it is fairly simple, which makes it a good heuristic device for formulating general questions about the relationship between cognition and the eye movement system. As a result, we think that the model serves as a good framework for thinking about reading—even if one does not accept all of the assumptions—and that further explorations in terms of the model are a useful way to pursue more general questions about cognition and eye movements.

Initially, our goal was to create a model of reading that was *sufficient* to explain how cognitive processes *could* control eye movements in reading. That is, before the E-Z Reader model, there was (and still remains) some skepticism about whether cognitive processes such as word identification could have an on-line influence on eye movements in reading given the time constraints involved in the task of silent reading (O'Regan, 1990, 1992; Reilly & O'Regan, 1998; Suppes, 1990, 1994; Yang & McConkie, 2001, 2004). This skepticism stems from the fact that the typical fixation time in silent reading is about 200–250 ms and the time to identify a word is often thought to be at least 150 ms. Because even the simplest eye movements have a latency of 150–175 ms (Rayner, Slowiaczek, Clifton, & Bertera, 1983), it may appear that there is not enough time for cognition to have an influence on individual eye movements. However, there is now a large body of data that indicate, for example, that there is a lawful relationship between fixation time on a word and its frequency in the language (Altarriba, Kroll, Sholl, & Rayner, 1996; Henderson & Ferreira, 1990; Hyönä & Olson, 1995; Inhoff & Rayner, 1986; Just & Carpenter, 1980; Kennison & Clifton, 1995; Rayner, 1977; Rayner, Ashby, et al., 2004; Rayner & Duffy, 1986; Rayner & Raney, 1996; Sereno, 1992; Vitu, McConkie, Kerr, & O'Regan, 2001). Thus, our goal was to come up with the simplest plausible model that we thought would account for the fact that word frequency and other variables relating to the speed of word encoding have immediate effects on eye movements.

We hasten to point out that our model was not intended as a complete model of reading, as there are many “higher order” processes that influence reading (such as syntactic processing and the construction of a discourse representation) that we did not attempt to model (for a review of the processes that have been shown to influence

eye movements during reading, see Rayner, 1998). However, there is some sense in which one could consider our model as a “module” that accounts for the forward progression of the eyes through the text (and thus accounts for a considerable amount of the data), with the higher-order processes largely influencing reading by stopping this forward progress when there are failures of comprehension (see Rayner, Warren, Juhasz, & Liversedge, 2004).

The goal of the present article is to examine the assumptions of the model more carefully. There are two reasons for doing this. The first is that, although the model was sufficient to explain a large set of phenomena, we realized that some important details of the model were not fully specified. In particular, a key assumption of the model is that there are three stages in the process of word identification, but several key properties of these stages were not made explicit. First, as will be expanded on shortly, we posited an early visual stage of processing prior to the processes specific to word recognition; however, we made differing assumptions about that stage in different versions of our model (cf. Rayner, Pollatsek, & Reichle, 2003; Reichle, Rayner, & Pollatsek, 2003). We now realize that many of the assumptions about this stage are “invisible” in most normal reading situations; however, there are some kinds of experiments (eye-contingent display-change experiments, most notably) that are excellent tests of assumptions about early visual processing in reading. Second, we posited two stages of word recognition after this early visual stage that helped to explain two phenomena in reading—“spillover effects” and the modulation of the amount of intake of information from the parafovea by the difficulty of processing the foveal stimulus. However, as these phenomena might be explained in other ways, we sought other tests of the necessity of positing separate stages of word encoding to understand eye movements in reading. (We will explain and discuss these phenomena more fully later.) In sum, one major goal in this paper is to be more precise about what needs to be assumed about the stages of word encoding in our model.

A second major goal of the paper is to explore one of the key assumptions of the model: that attention is allocated serially—one word at a time. This exploration has led us to extend the model to account for a wider range of phenomena in the reading literature, most notably (a) “boundary” experiments (e.g., Rayner, 1975) in which the availability of the parafoveal information from a word not yet fixated is manipulated and (b) experiments (Liversedge et al., 2004; Rayner, Liversedge, White, & Vergilino-Perez, 2003) in which selected text disappears after a fixed time from the beginning of a fixation. These two issues are related, in that the modeling efforts reported here have forced us to be more precise about which aspects of visual processing are affected by attention. As a result of these modeling efforts, we think we are in a better position to examine whether our assumption of serial processing of words is more or less reasonable than the gradient of attention that is assumed in some of the competing quantitative models of reading (Engbert et al., 2002, 2005; Reilly & Radach, 2003). More generally, we thought that these explorations would be more than just an exercise of working out the details of our model; the simplicity of our assumptions makes our model a good general framework for asking questions about how cognition and visual processing control the eye movement system in an ecologically valid and important task—reading.

2. Background: A brief description of the E-Z Reader model

Although the E-Z Reader model is conceptually reasonably simple, it is probably easier to explain it if we start by outlining an even simpler qualitative model proposed by Morrison (1984) that was in part the inspiration for the E-Z Reader model. Morrison proposed what appears to be the simplest plausible link between word identification and eye movement control: when a word is identified, an internal attention mechanism shifts the focus of processing to the next word in the text and, simultaneously with the attention shift, an eye movement is programmed to the next word. The E-Z Reader model incorporates this key idea of Morrison's model: attention is allocated serially from one word to another, so that the focus of attention at any instant is always on a single word.

This mechanism explains how the encoding of words in the text guides the progress of the eyes through the text. Moreover, it also makes it plausible that cognitive processes can control eye movements given the severe time constraints mentioned earlier (i.e., because of the latency of motor movements, there is not much time on a fixation to do word encoding). That is, because Morrison's model posits that attention shifts to the word immediately to the right of the fixated word when the fixated word is encoded, parafoveal processing of this word begins when this attention shift occurs and will continue for an interval of time approximately equal to the saccadic latency (150–175 ms; Becker & Jürgens, 1979; McPeck, Skavenski, & Nakayama, 2000; Rayner et al., 1983). Thus, according to Morrison's model, each word receives about 175 ms of processing before it is fixated, which can explain how processing of a word can be completed reasonably quickly after it is fixated and consequently be able to direct an eye movement to the subsequent word with a latency of 200–300 ms.

The above account, however, predicts that words would always be fixated in turn and hence would never be skipped. To explain skipping, Morrison posited the cancellation of earlier eye movement programs by later eye movement programs. This *cancellation* hypothesis was inspired by Becker and Jürgens (1979), who employed a simple visual task. Participants were asked to fixate a simple stimulus and then move their eyes to it when it moved to a new location, and the key trials were those in which the object moved twice with a small delay between the two moves. If the delay between the two moves was sufficiently long, people made saccadic eye movements to the two new locations in turn, but if the delay was short enough, people made only one saccade towards the second location. Becker and Jürgens concluded that if a program to move to the second location was initiated soon enough after the program to move to the first location, then the first program would be cancelled; otherwise, both programs were executed in turn. They also showed that there were sometimes fixations to intermediate locations for intermediate delays.¹

Morrison incorporated a saccade-cancellation mechanism in his model to explain skipping, and his mechanism was implemented in the E-Z Reader model by assum-

¹ In some prior, unpublished modeling work, we implemented these “compromise” saccades in the E-Z Reader model and found that they had virtually no effect on the simulations. As a result, we will not discuss compromise saccades in the present article.

ing that eye movements are programmed in two stages: a preliminary *labile stage*, and a subsequent *non-labile stage*. If another eye-movement program is initiated during the labile stage, the first program will be cancelled and only the second one will be executed. On the other hand, if the second program is initiated during the non-labile stage of the first one, both will be executed, with only minimal interaction between the two programs.² These assumptions are sufficient to explain skipping; if the eyes are on $word_n$ (the fixated word), then $word_{n+1}$ (the word to the right of the fixated word) will be skipped if it is encoded while the program to move the eyes to $word_{n+1}$ is still in its labile stage (with the eye movement program to $word_{n+2}$ canceling the program to fixate $word_{n+1}$). Moreover, this is likely to occur when $word_{n+1}$ is easy to process, such as when it is a short word, a high frequency word, and/or predictable from prior context. This qualitative prediction obviously mirrors the observed data, where words that are easy to process are in fact more likely to be skipped (Ehrlich & Rayner, 1981; Inhoff, 1984; O'Regan, 1979; Rayner, Sereno, & Raney, 1996; Rayner & Well, 1996).

Although the E-Z Reader model was inspired by Morrison's model, it departed from it in three ways. The first is that it attempted to be more explicit about the details of many of the processes that control eye movements during reading. In particular, Morrison did not assume that word encoding involved stages, and in particular, whether certain phases of word encoding were differentially affected either by the eccentricity of a word or by a saccadic eye movement. In contrast, as indicated above, the E-Z Reader model posits three stages of word processing (that we will explicate below). However, there were several key details that were not made explicit in earlier versions of the model. The first part of the present article will focus on further testing of assumptions about these stages of word processing in the hopes of arriving at a better theory of how visual and cognitive processing interrelate during reading.

The second way that the E-Z Reader model departed from Morrison's model is that it incorporates assumptions that account for why words are often refixated. (Morrison's model can only explain refixations as errors in saccadic programming.) For example, in the original version of E-Z Reader (Reichle et al., 1998), it was assumed that when a word is fixated, a default program is initiated to refixate that word, but that this "automatic" program can be cancelled by the same mechanism that explains skipping. Thus, for words that are easy to process, this refixation program will be cancelled more often than for words that are harder to process (a result that is consistent with the data). However, subsequent versions of the E-Z Reader model have made somewhat different assumptions about refixations. For instance,

² As explained elsewhere (see Reichle et al., 2003, Fig. 5), this is a simplification because the labile stage of saccadic programming actually consists of two sub-stages: a general *system-preparation* stage that readies the oculomotor system to begin programming a saccade, followed by a *location-to-distance-transformation* stage that converts the spatial target of the saccade to the amount of muscle force that is necessary to move the eyes the appropriate distance. There is some saving (in terms of saccadic programming time) if a second saccade is initiated while the labile stage of the first program is still in its system-preparation stage.

in the most recent version of the model (Reichle et al., 2003), the probability of programming a refixation saccade is modulated by word length. Admittedly, this is one of the least well worked out parts of the model, and we will not focus on this in the body of the current article; however, we think the refixation assumptions in the current version of our model are an improvement over prior assumptions. (For a discussion of why this issue is important, see Pollatsek et al., 2003.)

The third, and perhaps most fundamental, way that E-Z Reader differs from Morrison's model is that it decouples the attention shift to the next word from the initiation of the eye-movement program to the next word. (This is related to assumptions about stages of processing in word identification.) In the model, we posited that the eye-movement program is triggered by the completion an early stage of the word identification process, which we called the *familiarity check* (or L_1), and that the completion of a later stage of word processing, (full) *lexical access* (L_2), was the trigger for the shifting of attention. There were two major reasons for departing from the elegant simplicity of Morrison's model. The first is that there is evidence (Henderson & Ferreira, 1990; Inhoff, Pollatsek, Posner, & Rayner, 1989; Kennison & Clifton, 1995; Rayner, 1986; White, Rayner, & Liversedge, 2005) that when the fixated word is more difficult to identify, there is less information extracted from the upcoming word in the parafovea before it is fixated. We will return to the details of how this was demonstrated below, but for now, the essential point is that this phenomenon cannot be explained by a one-stage model such as Morrison's model. This is because Morrison's model assumes that readers shift attention and program the saccade to $word_{n+1}$ simultaneously. Thus, readers begin processing $word_{n+1}$ as soon as they have finished processing $word_n$ (however long the processing of $word_n$ takes) and then move their eyes to $word_{n+1}$ after the eye movement that was programmed is executed. Because this saccadic latency (i.e., the time needed to program and initiate a saccade) is just a function of the oculomotor system and is unrelated to cognitive processing, $word_{n+1}$ will receive approximately 175 ms of parafoveal processing (the duration of the saccadic latency; Becker & Jürgens, 1979; McPeck et al., 2000; Rayner et al., 1983), regardless of the difficulty of $word_n$.

Two key assumption of the E-Z Reader model were: (a) that the time to complete the familiarity check stage on $word_n$ is a function of its difficulty (e.g., its frequency in the language and predictability from the prior text); (b) the difference between the time to complete the familiarity check stage and the time to complete full lexical access of $word_n$ is also a function of its difficulty. The former assumption allows one to predict that fixation time on a word is a function of its difficulty, and the latter assumption allows one to predict the phenomenon discussed in the prior paragraph. This is because for more difficult words, the attention shift is later relative to both the initiation and execution of the eye movement program, and thus there is less time to process $word_{n+1}$ before it is fixated. This is illustrated in Fig. 1, which shows the mean times needed to complete the two stages of word identification (i.e., the familiarity check and full lexical access) relative to the saccadic latency for the full range of word frequencies. Notice that as a word becomes more difficult to identify (e.g., lower in frequency), not only do the times required to complete the familiarity check and lexical access increase, but the relative difference in times between these two pro-

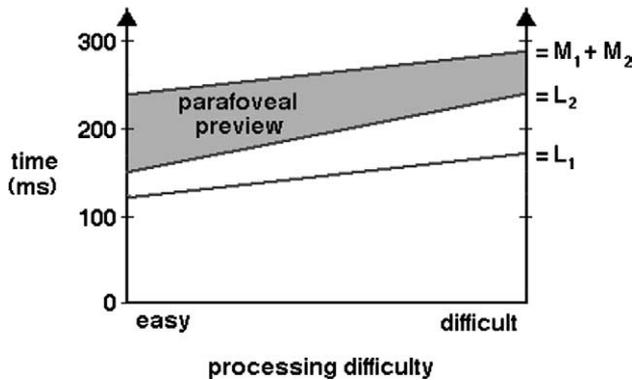


Fig. 1. The relationship between the two stages of word identification (i.e., the familiarity check, L_1 , and the completion of lexical access, L_2), and the saccadic latency ($=M_1 + M_2$), and how this produces parafoveal preview that is modulated by word processing difficulty (e.g., word frequency).

cesses also increases. This means (see Fig. 1) that the more difficult a word is to process, the less time there is between when lexical access completes and when the eyes move, and thus the less time for parafoveal processing of the next word (as illustrated by the height of the shaded area in Fig. 1).

The second phenomenon that Morrison's model cannot explain is spillover. There are several measures of spillover, but perhaps the most common is the effect of the difficulty of $word_n$ on the fixation duration on $word_{n+1}$. There are many reports of such effects (Rayner & Duffy, 1986; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989), and they can not be explained by Morrison's model. The reason for this is similar to the reason that Morrison's model can not explain the phenomenon in the above paragraph: All of the processing that is done on a word is completed in a single stage, with no delayed effects. In contrast, the assumption in our model that the difference between the two stages of word identification is a function of the difficulty of the word nicely explains spillover effects (at least qualitatively): the more difficult $word_n$ is, the less time remains to process $word_{n+1}$ before it is fixated (as illustrated in Fig. 1), so that the difficulty of $word_n$ spills over, inflating the fixation time on $word_{n+1}$.

E-Z Reader thus explains these two, perhaps related, phenomena quite well, but at a cost: the model posits three stages of word identification (an early stage of visual processing, following by two stages of lexical processing). In contrast, although Morrison's model probably tacitly assumes the division between a visual stage of processing and later lexical processing, it has only one lexical stage whose completion simultaneously (a) causes attention to immediately shift to the next word in the text and (b) produces a program to fixate that word. The question then arises as to whether the sacrifice in parsimony that came with positing the two stages of word identification is worth it, given that these phenomena may have other explanations. (For example, it could be that spillover effects are largely due to post-lexical processing.) Thus, another major focus of the current article is to examine the two-stage model of lexical processing in E-Z Reader more fully and to determine whether there are other, perhaps deeper, reasons for why positing separate stages for triggering a

saccade program and an attention shift was necessary. In addition, however, we wanted to explore the full implications of positing that the lag between the two stages of word identification is also a function of word difficulty [see Eq. (3) below].

A fruitful arena for testing some of these questions turns out to be experiments such as those using the boundary paradigm (Rayner, 1975), in which there are restrictions on the amount of useful text that is visible at a particular time. Since there have been no prior attempts to model such experiments, extending our modeling efforts to this domain would also be a good test of the adequacy of the E-Z Reader model and its basic assumptions, such as the assumption that lexical processing only occurs on one word at any given instant. We now describe the E-Z Reader model in greater detail.

3. The stages of processing in E-Z Reader: Early visual processing vs. lexical processing

As indicated above, in E-Z Reader, we have made a distinction between two stages of lexical processing: an early one that is the signal that drives the eyes, and a later one that is the signal that moves attention to the next word. However, we also posited a “visual” stage (prior to these lexical stages) that has qualitatively different properties. This distinction was inspired by the classic distinction between early “featural” processing and later “configural” processing (see Pollatsek & Digman, 1977; Treisman & Gelade, 1980). To some extent, this distinction parallels the distinction in reading between the signal of *where* to move the eyes and *when* to move the eyes (Becker & Jürgens, 1979; Rayner & McConkie, 1976). That is, there seems to be abundant evidence that the computation of where to move the eyes next when reading English is importantly determined by lower-level featural information, such as where the boundaries between words are (i.e., where there are gaps), and furthermore, these targeting decisions can be made over a region that is substantially bigger than the area from which people are extracting useful lexical information. For example, the data from moving window experiments indicates that little useful letter information can be extracted more than 12 characters from fixation (see Rayner, 1998, for a summary). In contrast, there is evidence that return sweeps—eye movements to the beginning of the next line of text—are programmed reasonably accurately (typically, they usually only undershoot by 6–8 characters), as are some large regressive eye movements to areas of text that caused difficulty (Frazier & Rayner, 1982; Kennedy, Brooks, Flynn, & Prophet, 2003; Meseguer, Carreiras, & Clifton, 2002).

These phenomena suggest a substantial disconnect between how these early *visual* processes work and how later *lexical* processes work. In particular, in all versions of the E-Z Reader model, it was assumed that serial attentional assumptions only applied to the lexical stages of processing. That is, the assumptions were: (a) that visual information is acquired in parallel across the visual field in these early stages (although the quality of the visual information may differ depending on the distance from the fixation point), whereas; (b) lexical processing is guided by a spatial attentional system that processes one word at a time. In most of our modeling, this spatial attention is assumed to be distributed in parallel across the entire word; however, in

the case of long polymorphemic words, it is plausible that attention may not be focused on the entire word, but only on a morpheme (Hyönä, Bertram, & Pollatsek, 2004; Pollatsek et al., 2003).

Our view is not that this attentional allocation is hard-wired in the human visual system, but that it is an attentional allocation strategy that is close to optimal for reading and that people learn this fairly early in their reading history.³ That is, in the task of reading, one attempts to decode print to produce a representation of a spoken utterance that is logically sequential in nature. Furthermore, especially in English, word order is crucial in understanding the meaning of many utterances. Thus, a spatial attentional system that operates from left-to-right across the page will automatically reproduce the temporal order of the words in a spoken sequence of English. A model in which words are processed in parallel can not guarantee this, and thus some additional mechanism would be needed to sort out the appropriate word order. Although it is plausible that such a mechanism can be devised, it seems non-optimal to have to rely on one.

On the other hand, there may be other tasks involving words, such as deciding whether two words are physically identical or are synonyms, where reproducing order is not critical and in which other attentional strategies may be used. Certainly, there are real-world situations involving scenes where one has to judge the relation and/or interaction of discrete objects in which a serial processing strategy might be quite non-optimal. However, even in these cases, there might be costs in attempting parallel processing of physically distinct objects, such as mis-assigning features of one object to the other object (Treisman & Schmidt, 1982).

We should also make clear that the assumption of serial allocation of attention does not imply that only one word is attended on a given fixation. On the contrary, as switches of attention almost always occur during a fixation in reading, the default state of affairs is that both the fixated word ($word_n$) and the word to its right ($word_{n+1}$) are attended to at different points in time on a given fixation. (Occasionally, $word_{n+2}$ is also attended as well.) We were less clear in our prior modeling, however, about some of the details of visual processing, especially (a) what the ‘visual processing time’ prior to lexical processing represented, and (b) how visual and lexical processing proceeded when the eyes moved from fixation to fixation. As will be made clear in the following section, some of these details are not particularly important for modeling standard reading situations, but become quite critical when one is modeling situations in which the availability of text is experimentally manipulated.

4. E-Z Reader 9: The model’s core assumptions

We begin by providing a brief overview of the E-Z Reader model and a detailed description of the model’s core assumptions regarding visual processing and the role

³ One piece of evidence for this is that there are asymmetries in the perceptual span in reading even in the first year of reading (Rayner, 1986).

of attention in word identification. Fig. 2 is a schematic diagram of the E-Z Reader model. As already stated (and as Fig. 2 shows), word identification in the E-Z Reader model is completed in three stages: an early visual processing stage (*V*) that is pre-attentive, and later lexical processing having two substages, each of which require attention, and that culminates in the word’s meaning being available for further linguistic processing. Furthermore, there are two “signaling” events that occur during these later, attention-demanding states: the completion of the *familiarity check* (L_1) signals the oculomotor system to program a saccade to the next word, and the *completion of lexical access* (L_2) signals the attention system to shift attention to the next word. Formally, we assume that the signals to the eye movement and attentional systems are transmitted instantaneously. Although this assumption of instantaneity may seem unrealistic, one could conceive that some part of the time that is needed to program the saccade includes whatever time is necessary to transmit the signal from the parts of the brain mediating the familiarity check to the parts of the brain responsible for programming the saccade. Likewise, one could conceive of a portion of the time that is needed to complete lexical access as subsuming the latency of an attention shift.

Although the familiarity check and the completion of lexical access formally appear as sequential processing stages, we do not necessarily conceptualize them

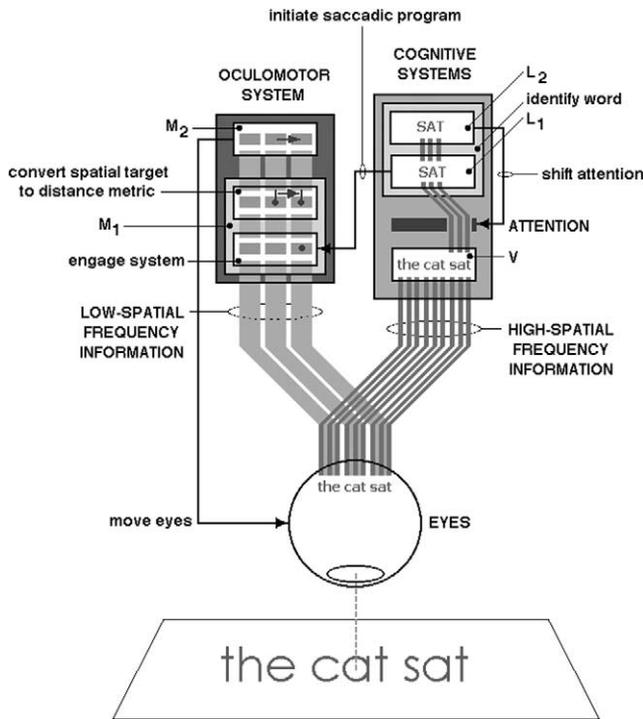


Fig. 2. A schematic diagram of the E-Z Reader model.

that way. That is, we see word identification as being a complex process which involves competition among various lexical entries and which involves letter identification, letter cluster identification, activation of phonological codes and of other types of codes (e.g., morphological) as inputs to this process. For the purposes of this discussion, however, it is sufficient to think of the familiarity check as a point in lexical processing in which the processing system, through experience, knows that it is highly likely that full word identification will be achieved shortly, and that it is quite “safe” to program an eye movement to the next word (i.e., the saccade is unlikely to be premature and thus unlikely to result in a subsequent regression back to the word that is being processed). In contrast, the completion of lexical access is a point when processing of a word is complete enough so that the processing system can shift attention to the next word so that lexical processing of that word can begin with no “crosstalk” from the processing of the prior word.

The time that is required to complete the early visual processing stage, $t(V)$, is assumed to be unaffected by lexical variables and is simply set to a fixed time. Moreover, it is assumed that the visual information that is acquired from this stage of processing is acquired from across the visual field in parallel and at the same rate, but that the quality of the information decreases the further it is from the fixation point. The time we assumed for this initial processing stage, however, varied from simulation to simulation, and we realized that the fits of the model were not sensitive to the time that was assumed for the visual processing stage. This insensitivity bothered us, and a key issue that will be addressed later is whether the model is similarly insensitive in fitting experiments in which there are display changes, and whether the values that best fit these data are reasonable, given what is known about the visual system. Based on recent estimates of the amount of time required to propagate information on the retina to the occipital cortex, we set the value of $t(V)$ equal to 50 ms (i.e., the duration of the “eye-to-mind lag”; Clark, Fan, & Hillard, 1995; Foxe & Simpson, 2002; Mouchetant-Rostaing, Giard, Bentin, Aguera, & Pernier, 2000; Van Rullen & Thorpe, 2001).

A second key aspect of the model is that all of this visual information is assumed to be potentially available for further processing, but the type of processing that is necessary to identify a given word requires that attention be focused on that word. Thus, after a given word has been identified, attention shifts to the next word so that lexical processing of that word can begin using information that is available from the visual processing stage. What we were not clear about is the relationship of this attentional allocation to maintenance of visual and/or lexical information when the visual input disappears. Clarifying this issue is another focus of our present modeling.

The details of how the model assumes subsequent lexical processing to be carried out are as follows. First, upon shifting attention to a given word, there is a certain probability that the word will be “guessed” from its context using little or none of the information that is available from the visual processing stage. For the sake of keeping our model as simple as possible, we assume that this happens with a probability that is equal to a word’s mean cloze-task predictability, so that only words

that are highly constrained by their sentence context (e.g., function words) are likely to be processed in this top–down manner. This assumption is based on the observation that readers often do not fixate highly predictable word locations in experimental paradigms even when the correct letter information is not present, as in a boundary experiment in which a predictable target word is replaced by a preview of equal length, but whose letters are quite different from those of the target word (e.g., Balota, Pollatsek, & Rayner, 1985; cf. Drieghe, Rayner, & Pollatsek, 2005). Such results suggest that the higher-level linguistic processing that is going on in parallel with word identification is sometimes sufficient to “fill in” the gaps in the sentence meaning that is being constructed.

The vast majority of words are not guessed in this manner, however, and the mean time that is needed to complete the familiarity check on a given word, $t(L_1)$, is given by Eq. (1). Here, the base time that is needed to complete the familiarity check on a word (determined by the free parameter α_1 ; see Table 1 for parameter values and Appendix A for a discussion of how they were selected) is reduced as a function of the natural logarithm of that word’s frequency of occurrence in printed text

Table 1
E-Z Reader 9 parameters, their values, and their interpretation

Model component	Parameter	Value	Interpretation
Visual processing	V	50	Eye-to-brain lag (ms)
	ε	1.15	Effect of visual acuity on L_1
Lexical processing	α_1	122	Intercept: mean maximum L_1 processing time (ms)
	α_2	4	Slope: effect of frequency on L_1 processing time (ms)
	α_3	10	Slope: effect of predictability on L_1 processing time (ms)
	Δ	.5	Proportional difference between L_1 and L_2
	σ_γ	.22	Standard deviation of γ distributions (i.e., $\sigma = .22\mu$)
Saccadic programming	M_1	100	Mean labile programming time (ms)
	ξ	.5	Proportion of $t(M_1)$ allocated to “engaging oculomotor system”
	M_2	25	Mean non-labile programming time (ms)
	R	117	Mean refixation “decision” time (ms)
	S	25	Saccade duration (ms)
	Ψ	7	Optimal saccade length (character spaces)
	Ω_1	7.3	Intercept: effect of launch-site fixation duration on systematic range error
	Ω_2	3	Slope: effect of launch-site fixation duration on systematic range error
	η_1	.5	Intercept: saccade random error component (character spaces)
	η_2	.15	Slope: saccade random error component (character spaces)
λ	.09	Increase in refixation probability per character space deviation from word center	

(as tabulated in the Francis & Kučera, 1982, norms⁴ and scaled by the free parameter α_2), and the word's predictability within its sentence context (as determined by separate cloze-task experiments and the free parameter α_3). Thus, words that are more common (frequent) and/or more constrained by their contexts (predictable) take less time to process than do words that are less common and/or less constrained.

$$t(L_1) = \alpha_1 - \alpha_2 \ln(\text{frequency}) - (\alpha_3 \text{ predictability}). \quad (1)$$

Several points about Eq. (1) warrant discussion. First, we employ Monte-Carlo simulations, so that the actual time predicted for a particular word is a random deviate that is sampled from a γ -distribution with a mean defined by Eq. (1). (The values of $t(L_2)$, $t(M_1)$, and $t(M_2)$ below are similarly drawn from γ -distributions whose mean values are given below.) The second point is that Eq. (1) instantiates what we refer to as the “additive” version of the model (Rayner, Ashby et al., 2004) in that the time needed to identify a word is an additive function of its frequency and predictability.⁵ Finally, the rate at which the familiarity check is completed is modulated by visual acuity, so that a word will take longer to identify as the mean disparity between each of its letters and the center of fixation (i.e., the fovea, or point of maximal visual acuity) increases. This is done by increasing (multiplying) the actual duration of $t(L_1)$ that is sampled from the γ distribution by a processing rate value that is determined by Eq. (2). Here, ε is a free parameter that modulates the effect of visual acuity and N is the number of letters in the word being processed. Thus, a word that is longer and/or further from the center of vision requires more time to process than a word that is shorter and/or closer to the center of vision.

$$\text{rate} = \varepsilon^{(\sum|\text{letter}-\text{fixation}|)/N}. \quad (2)$$

Finally, the mean time that is needed to complete lexical access on a word, $t(L_2)$, is given by Eq. (3). Here, $t(L_1)$ is the mean time that is needed to complete the familiarity check (as specified by Eq. (1), prior to the actual familiarity check time being sampled from a γ -distribution or being modulated by visual acuity) and λ is a free parameter that insures that $t(L_2)$ is some fixed proportion of $t(L_1)$. This later fact is important because the amount of parafoveal processing that is completed on word_{n+1} from word_n will *decrease* as the processing difficulty of word_n *increases*. This allows the model to predict less preview benefit from difficult-to-process words, which is consistent with empirical results (Drieghe et al., 2005; Henderson & Ferreira, 1990; Inhoff et al., 1989; Kennison & Clifton, 1995; White et al., 2005). It also allows the model to predict spillover effects on word_{n+1} because, if word_n is difficult

⁴ We used the Francis & Kučera (1982) norms because they were the most convenient ones to use for American English. However, we do not think our modeling would have come out substantially differently if we had used one of the other currently available corpora.

⁵ Originally, we had posited a model in which word frequency and predictability were combined using a multiplicative version of Eq. (1) (Reichle et al., 1998, 2003). However, in an experiment that examined these two variables factorially (Rayner, Ashby et al., 2004), we found that this assumption was not tenable. Upon reflection, we also realized that the assumption was counterintuitive, as it predicted smaller predictability effects for frequent words.

to process, then there will be less parafoveal processing of $word_{n+1}$, which will in turn increase the fixation duration on $word_{n+1}$. Again, such effects are consistent with empirical results (Rayner & Duffy, 1986). We should point out that, in contrast to $t(L_1)$, the actual value of $t(L_2)$ is not modulated by visual acuity. This distinction is based on our assumption that visual processing has largely finished by the time the familiarity check has ended, and that the completion of lexical access proceeds using more abstract representations. Finally, as already mentioned, the completion of L_2 on one word causes attention to shift to the next word so that L_1 can begin on that word.

$$t(L_2) = \Delta t(L_1). \quad (3)$$

The above summarizes the “front end” of the E-Z Reader model; that is, those parts of the model that deal with visual processing, the allocation of attention, and the identification of words. The remaining parts of the model deal with the programming and execution of eye movements. Because these assumptions are not different from those that we have used in earlier versions of our model, we will provide only a brief description of them here (for a more complete description, see Reichle et al., 2003).

The major concept that is embodied in the model is that eye movement programs can be cancelled if a signal to program another eye movement is initiated early enough. This assumption is based on the work of Becker and Jürgens (1979; see also Leff, Scott, Rothwell, & Wise, 2001; McPeck, Skavenski, & Nakayama, 2000; Moller & Fischer, 1999; Vergilino & Beauvillain, 2000), and is captured by the assumption that saccades are programmed in two stages: a preliminary *labile* stage (M_1) that is subject to cancellation by subsequent saccadic programs, followed by a *non-labile* stage (M_2) that is not subject to cancellation (see Fig. 2). This division allows the model to account for word skipping. For example, consider the situation in which the eyes are on $word_n$ and lexical processing of $word_{n+1}$ has started. If the familiarity check on $word_{n+1}$ finishes before the labile program to move the eyes to $word_{n+1}$ is completed, then the oculomotor system will initiate a second labile program to move the eyes to $word_{n+2}$, thereby canceling the first program and causing $word_{n+1}$ to be skipped. However, if the familiarity check on $word_{n+1}$ completes after the labile program, then the non-labile program will be started, resulting in an obligatory saccade to $word_{n+1}$.

The durations of M_1 and M_2 are random deviates that are sampled from γ -distributions with means set equal a priori to 100 and 25 ms, respectively. The labile stage of programming is further decomposed into two sub-stages: an initial period during which the oculomotor system is engaged and made ready to program a saccade, followed by a period during which the spatial coordinates of the saccade target are converted into a distance metric. Each of these sub-stages was assumed to subsume some proportion of the total duration of M_1 (as specified by the free parameter ζ). Finally, for the sake of simplicity, the durations of saccades were set equal to 25 ms even though the durations will vary somewhat as a function of saccade length. The above durations for M_1 and M_2 are quite consistent with those inferred from a *counter-manding* paradigm (Hanes & Carpenter, 1999) in which, on some trials, people are

given a signal (at varying delays) to not make a saccade (i.e., cancel an already existing eye movement program).

Based on the work of McConkie, O'Regan, and their colleagues (McConkie, Zola, Kerr, Bryant, & Wolff, 1991; McConkie, Kerr, Reddix, & Zola, 1988; O'Regan, 1990, 1992; O'Regan & Lévy-Schoen, 1987) we (see Reichle et al., 1999) assume that all saccades are directed towards the centers of the intended word target (though they typically land short of the center; Rayner, 1979) using the low-spatial frequency information that is available from parafoveal vision to identify word boundaries (this is schematically depicted in Fig. 2). The length of the actual saccade that is executed is equal to the sum of the intended saccade length, systematic error, and random error. The systematic error is given by Eq. (4), and is a function of both the difference between the intended saccade length and the “preferred” saccade length (i.e., the saccade length that tends to neither overshoot nor undershoot its intended target; $\Psi = 7$ character spaces). The systematic error is also a function of the launch-site fixation duration (duration_{n-1}), as modulated by two free parameters, Ω_1 and Ω_2 . Thus, seven-character saccades will tend to be more accurate than either shorter or longer saccades, and saccades following long fixations will tend to be more accurate than saccades following shorter fixations. The random error component of the saccade that is executed is sampled from a Gaussian distribution with $\mu = 0$ and σ given by Eq. (5). The random error thus increases as the intended saccade length increases, as determined by two free parameters: η_1 and η_2 .

$$\text{systematic error} = (\Psi - \text{intended length})[\Omega_1 - \ln(\text{duration}_{n-1})/\Omega_2]. \quad (4)$$

$$\sigma = \eta_1 + \eta_2(\text{intended length}). \quad (5)$$

Our final assumption about saccades concerns refixations. Our assumption here is very similar to what is assumed in the most recent version of the SWIFT model of eye-movement control—that upon initially fixating a given word, a corrective refixation saccade can be programmed and executed to move the eyes to a better viewing location (Engbert et al., 2005). The intuition behind this assumption is that initial fixations landing near the beginnings/endings of words do not allow efficient lexical processing due to the poorer visual acuity of many letters in the word from such locations [see Eq. (2)], and that a rapid corrective saccade will (on average) move the eyes closer to the center of the word so that lexical processing can proceed more rapidly from this better viewing location. As mentioned above, p , the probability of initiating a refixation saccade, is a function of the absolute distance (in character spaces) between the initial landing position on a word and the word's optimal viewing position, as scaled by the free parameter λ [see Eq. (6)]. Note that, because the potential size of the absolute distance increases with word length, longer words are more likely to be the recipients of refixations than shorter words (Vergilino & Beauvillain, 2000). For it to be plausible that such an “intelligent” eye movement could be programmed, it was necessary to assume that the “decision” about whether or not to initiate a corrective refixation saccade could be made only after a delay that is greater than the time that is necessary to get the information about where the

initial fixation is located. This delay was sampled from a γ -distribution with a mean equal to the free parameter (R); because feedback about the initial fixation location is based on visual information, the value of $t(R)$ was restricted so that $t(R) > t(V)$ (i.e., the duration of the eye–brain lag).⁶ Thus, upon fixating a word, a refixation saccade [obeying Eqs. (4) and (5), just like interword saccades] is initiated if: (a) another labile saccade has not already been initiated; and (b) a saccade has not already been initiated to move the eyes to the next word. These two restrictions are sufficient to prevent refixation saccades from canceling saccades that would otherwise cause the eyes to move forward in the text.

$$p = \lambda |\text{optimal viewing position} - \text{initial landing position}|. \quad (6)$$

Unless otherwise noted, all of the simulations that are reported in this article were completed using the model as described above Eqs. (1)–(6). Because the goal in this article is to focus on key theoretical issues related to visual processing, word identification, and attention, a number of simulations are presented in the Appendix A that were necessary to show that the current version of the model whose variants we explore in the current study can handle the standard benchmark phenomena that were explained by earlier versions of the model. (We will refer to the variants that use these estimates of parameters—except for differing assumptions about the time needed for visual preprocessing and the availability of visual information after it disappears from the retina—as *E-Z Reader 9*.) Table 1 indicates the values of parameters that were assumed a priori in these simulations and the values of free parameters that were estimated. However, to emphasize the point that the assumed duration of the visual processing stage is largely irrelevant in modeling standard reading situations, we present the results of two simulations, one assuming a value of $t(V) = 50$ ms and one assuming a value of $t(V) = 90$ ms. As can be seen in Table 2, the modeling is largely indifferent to the value selected, both in terms of goodness of fit and in terms of the values of the other free parameters.

5. Visual processing in E-Z reader: How does vision interface with cognition?

The issue of how visual information is processed and combined on successive fixations is clearly a central one in visual perception, as our eyes move 4–5 times per second when we read or view any static scene; yet we manage to perceive a stable and reasonably coherent scene. Although there are many complexities to this issue that are beyond the scope of the current exploration, it seems necessary to provide a reasonably broad perspective to frame our modeling questions appropriately.

⁶ The inequality should be on values of each sampled random deviate, but we imposed the inequality on the mean of the distribution of $t(R)$ for modeling convenience. However, since the distribution mean of $t(R)$ is 117 ms and $t(V)$ is a constant 50 ms, the probability that each value of $t(R)$ is greater than the time for $t(V)$ is close to 1.

5.1. *The integration of information across saccades*

The first point that needs to be made is that the bulk of the evidence indicates that there is little or no ‘visual’ information in reading that is combined across successive fixations. Perhaps the clearest demonstration of this came from experiments employing eye-contingent display change techniques in reading. In one (McConkie & Zola, 1979), participants read text in which the case of letters alternated (e.g., “IiKt tHiS”) and, during every saccade, the case of every letter changed (e.g., so that “IiKt tHiS” became “LiKe ThIs”). The central finding was that reading was no different in this condition than in one in which the text was in alternating case but in which there were no display changes. Although many upper and lower case letters are visually similar, most are quite different in size (e.g., “o” vs. “O”) and some are quite different in form (e.g., “g” vs. “G”). Thus, if the information that is combined across successive fixations is largely based on visual information, then one would expect appreciable cost in the condition in which the letters change with each saccade. A similar finding was obtained by Rayner, McConkie, and Zola (1980) in an experiment that used a variant of the *boundary* technique (which will be described in more detail in the next section). In some conditions in this experiment, a single target word changed case during the saccade immediately prior to the fixation on the target word. Again, the change of case had no effect on processing time: naming times on the target word were no longer in the case change condition than when the case did not change. Thus, the ‘visual information’ (i.e., lower-level featural information about the stimulus) that is extracted from a word when it is in the parafovea has virtually no effect on the speed of encoding that word when it is later fixated. (Consistent with this result, there was also little evidence for the retention of visual information across fixations when line drawings of objects were named; Pollatsek, Rayner, & Collins, 1984.)

We should quickly add, however, that there is abundant evidence that information extracted from the parafovea does influence the processing of a word when it is later fixated. If one contrasts a condition in which all the letters of the parafoveal preview are different from the target to a *full preview* condition in which there is no display change, there is a large difference in the *gaze duration* (the total fixation time on a word on the “first pass” through the text) on the target word (usually about 40–50 ms; see Hyönä et al., 2004; for a recent meta-analysis). Moreover, one does not need an identical preview to achieve benefit over a condition in which all the letters of the word change during the saccade; for example, a large benefit accrues even if only the first two or three letters remain unchanged (Balota & Rayner, 1983; Rayner et al., 1980), and there is also evidence that overlap of phonological codes has a benefit above and beyond overlap of orthographic codes (Chace, Rayner, & Well, 2005; Henderson, Dixon, Petersen, Twilley, & Ferreira, 1995; Mielliet & Sparrow, 2004; Pollatsek, Lesch, Morris, & Rayner, 1992). There is an extensive literature on the details of parafoveal preview benefit, which is beyond the scope of this discussion (see Rayner, 1998, for a summary), but the essential point is that information is retained from one fixation to the next, but at a more abstract level than that of visual feature information. This information may be at various levels of orthography (e.g.,

letter clusters as well as letters) and various levels of phonology (e.g., phoneme and syllable).

In the E-Z Reader model, this distinction between low-level visual information and the more abstract orthographic, phonological, lexical (and perhaps morphemic) information is captured in the distinction between the early visual processing stage and the later lexical processing stages. Moreover, consistent with the above data, there is the distinction in the E-Z Reader model that visual information is not integrated across fixations whereas information relevant to the lexical processing stages is. Another way to phrase this is that the visual processing stage needs to start anew on each fixation in the model, but the later stages do not.

6. An in-depth examination of the visual pre-processing assumptions

In our prior simulations, we did not think carefully enough about our assumptions about what happens to visual and lexical processing when useful visual information is no longer present at the retina, most notably during and immediately after a saccade. In the baseline simulations in the current study using the E-Z Reader 9 model (see Appendix A for the details), we assumed, as indicated earlier, that the visual processing stage takes a fixed amount of time ($t(V) = 50$ ms) at the beginning of each fixation before higher-order encoding processes can operate on the new visual information acquired from the fixation. In contrast, we assumed that the later stages of word identification (which give rise to both the familiarity check and full lexical access signals, as discussed above) do not decay and continue where they left off prior to the saccade. In addition, we assumed that the higher-order processes continued through the saccade and during the visual processing stage of the next fixation (working on the visual information obtained from the prior fixation). This raises three obvious questions. The first is what the justification is for assuming that processing of the old visual material continues for this long; the second is why the simulations assuming 50 and 90 ms for the duration of the $t(V)$ stage came out approximately the same (see Table 2; also, cf. Rayner, Liversedge, & White, 2005; Reichle et al., 2003); and the third is what is the best way to test for the “best” duration of $t(V)$ and to test for the “best” assumption about how long lexical processing continues after visual information disappears.

First consider the question of how long the visual information remains available for later cognitive processing once a fixation ends. What is plausible? The answer to this question is likely to depend, at least partially, on how one conceptualizes the processing time assumed for the visual processing stage. The simplest conceptualization is that the duration of the stage largely reflects an “eye–brain” lag. (Let’s assume for the moment that this value is 90 ms.) If we adopt this conception of the visual processing stage, it seems unreasonable to assume that the information in it immediately becomes unavailable to later stages of processing the instant that the eye begins to move. That is, if it takes 90 ms for the visual information to reach the appropriate lexical processing centers at the beginning of the fixation, it should take

Table 2

Observed (Schilling et al., 1998) and predicted (E-Z Reader 9, using $V = 50$ vs. 90 ms) mean fixation duration and fixation probability measures for five frequency classes of words

		Word frequency (# Occurrences per million)				
		1-10	11–100	101–1000	1001–10,000	10,001+
Gaze durations	Observed	293	272	256	234	214
	E-Z Reader, $V = 50$	295	275	241	220	216
	E-Z Reader, $V = 90$	291	269	235	223	220
First fixation durations	Observed	248	234	228	223	208
	E-Z Reader, $V = 50$	251	239	225	217	214
	E-Z Reader, $V = 90$	252	237	221	220	218
Single fixation durations	Observed	265	249	243	235	216
	E-Z Reader, $V = 50$	259	244	226	217	215
	E-Z Reader, $V = 90$	260	243	222	220	219
Probability of making one fixation	Observed	.68	.70	.68	.44	.32
	E-Z Reader, $V = 50$.75	.74	.75	.52	.42
	E-Z Reader, $V = 90$.75	.74	.75	.50	.41
Probability of making two fixations	Observed	.20	.16	.10	.02	.01
	E-Z Reader, $V = 50$.15	.13	.06	.01	0
	E-Z Reader, $V = 90$.14	.12	.05	.01	0
Probability of skipping	Observed	.10	.13	.22	.55	.67
	E-Z Reader, $V = 50$.09	.12	.19	.47	.58
	E-Z Reader, $V = 90$.10	.14	.20	.49	.59

Note. The overall goodness-of-fit metrics for the two versions of E-Z Reader 9 are: (a) using $V = 50$ ms, $\text{RMSD} = .153$; (b) using $V = 90$ ms, $\text{RMSD} = .171$. The best-fitting parameter values for the second simulation (with $V = 90$ ms) were: $\alpha_1 = 110$; $\alpha_2 = 4$; $\alpha_3 = 14$; $\lambda = .07$; and $R = 109$. All of the other parameter values were the same for both simulations (see Table 1).

roughly the same 90 ms for the “news” that it was no longer present on the retina to reach these brain areas as well. Thus, there would be a gap of at most the saccade duration (on average, about 25 ms) between when the visual information from the prior saccade would become unavailable and when the visual information from the current saccade would become available. Moreover, it is not clear that the 25 ms in which the visual information coming in during the saccadic eye movement reaches the brain would wipe out the availability of the visual information acquired on the prior fixation. That is, this information coming in during the saccade is a “smear” on the retina and is likely to have few visual features that would have a strong masking effect on the old information, and, as we will discuss later, having visual information disappear with little or no masking following it usually has little or no effect on its functional persistence.

This plausibility argument led us to assume that the visually processed information from the prior fixation was available to be worked on until the visually processed information from the current fixation replaced it, and we used this assumption in the simulations in the Appendix A. The above discussion hopefully also helps in answering the second question posed above: why the simulations for normal reading that assume that $t(V)$ lasts 90 ms are so similar to the simulations that assume that $t(V)$ lasts 50 ms (see Table 2). That is, if one assumes that lexical

processing on the visual information extracted from the prior fixation continues without interruption during the saccade and until the visual material from the new fixation displaces it, lexical processing is continuous regardless of what the delay time is assumed for the duration of $t(V)$. The delay time assumed would not be completely irrelevant, because making the visual processing time longer would cause there to be a somewhat longer time at the beginning of a fixation when poorer visual information about a word from the parafovea is still being processed rather than the new information from the fovea. However, this only produces minor effects on lexical processing that can be compensated for in the model by speeding up lexical processing parameters a bit.

Thus, the assumption of how long the visual processing stage lasts is largely irrelevant when one is modeling normal reading if one assumes that lexical processing of information from prior fixations can continue until the new information displaces it. In contrast, situations when there is no useful visual information in the parafovea can provide an important test of the processing assumptions about early visual processing because, in these situations, visual, and hence lexical, processing of a word can not begin until the word is fixated. Thus, these situations provide three key tests of the E-Z Reader model: (a) whether any version of the E-Z Reader model can account for situations in which no useful information can be extracted from the parafovea; (b) whether a successful fit of these data depends on the assumptions one makes about the duration of the visual processing stage; and (c) whether there is a critical test of how long lexical processing continues through a saccade and the ensuing fixation.

There are two important paradigms that have been employed to examine the use of parafoveal information in reading: the *moving window* paradigm (McConkie & Rayner, 1975; Rayner & Bertera, 1979; Rayner, Well, Pollatsek, & Bertera, 1982) and the *boundary* paradigm (Rayner, 1975). Here, we will focus on the boundary paradigm because it is conceptually the simpler situation to model and understand, and thus the more relevant test case. The essence of the paradigm is that there is an eye-contingent display change involving a key target word (which is usually a noun). Typically, a sentence is presented and the stimulus in the target word location is initially a *preview* stimulus that is something other than the target word (see Fig. 3). However, when the eyes cross an invisible *boundary* that is typically one or two characters prior to the beginning of the target word location, the preview stimulus is

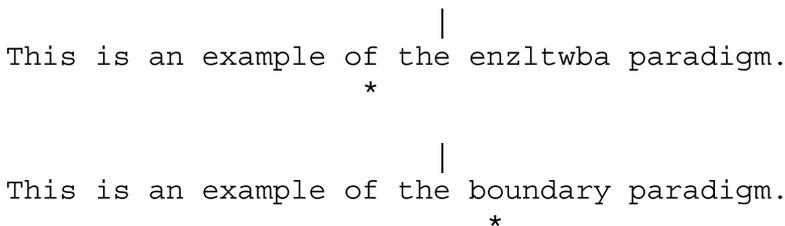


Fig. 3. A schematic diagram showing the sequence of events in the boundary paradigm.

replaced by the target word. Due to saccadic suppression, participants are rarely aware of either this change or of the identity of the preview.⁷

A major motivation for developing the boundary paradigm was to understand the kinds of information that are acquired from the parafovea that facilitate the processing of a word when it is fixated. As a result, the chief focus in most boundary experiments has been on manipulating the visual, orthographic, or phonological similarity of the preview to the target. A discussion of all of these conditions, however, would be considerably beyond our current modeling efforts, as dealing with them would necessitate having a detailed model of word identification. Instead, we will merely consider the difference between two extreme conditions: (a) *full preview* (when there is effectively no display change and the target word is visible in the parafovea) and (b) *no preview* such that there is no useful letter information in the target word location until it is fixated (e.g., when the preview is random letters or X's). (In all of the experiments, the preview has the same number of letters as the target; this is necessary to indicate how long the target word will be and insures in most conditions that the display change is not noticed.)

As indicated above, a recent review of the data on preview benefit (Hyönä et al., 2004) indicates that the difference in gaze duration between the full preview condition and a no preview condition is usually about 40–50 ms. We quickly realized that it would be a challenge to explain why this effect was as small as it was. First consider that we had been assuming a 90 ms visual processing time in E-Z Reader 7, and that processing of the word from the prior fixation in the full preview condition continued until the new information replaced it. This means that processing in the full preview condition has at least a 90 ms head start over the no preview condition! But the head start is even greater than that because the parafoveal information in the full preview condition is likely being processed during the saccade and some of the prior fixation as well. This indicates that assuming 90 ms as the visual processing time in the model will probably not work and that the 50 ms value would be better. But are there any values for the visual processing time in the E-Z Reader model that will provide a good fit to boundary data? And if so, are these optimal values sensible (i.e., consistent with what we know about early visual processing)?

In addition, when we thought further about modeling the boundary paradigm when doing the basic simulations in the Appendix A, we realized that it put significant constraints on the values of other parameters in the model. In particular, we had assumed unrealistically long motor programming times in earlier versions of the model, largely to demonstrate that one did not have to posit unrealistically short motor programming times to explain how word identification can affect eye movements in reading. However, we realized we had overdone it, as our prior estimates would have predicted unreasonably long eye-movement latencies to the visual onset of a simple stimulus. As a result, in all the modeling in the present paper, we assumed

⁷ There is a version of the boundary paradigm (Rayner, McConkie, & Ehrlich, 1978) in which the target word is displayed in isolation and naming or lexical decision latencies are used to assess performance. Although the data from this paradigm are similar to those from the reading paradigm discussed in the text, we will not deal with it in our discussion.

Table 3

Mean predicted first-fixation durations (FFD, in ms) and gaze durations (GD) in a simulated boundary paradigm using E-Z Reader with $t(V) = 50$ vs. 90 ms

Condition	$t(V) = 50$ ms		$t(V) = 90$ ms	
	FFD	GD	FFD	GD
No preview	283	337	298	348
Preview	244	291	247	289
Difference	39	46	51	59

that the total motor programming time was 125 ms, which when added to a visual processing time of 50 ms is consistent with the observed 175 ms saccadic latencies to simple visual onsets.

In our first simulations of display-change phenomena, we tested what an optimal visual processing time would be to explain the typical size of the (maximum) preview benefit: the difference in the gaze duration on a target word between when there is a full preview (i.e., there is no display change) and when there is no preview (i.e., there is a preview of a letter string in the parafovea with the same number of letters as the target word, but in which there is no useful letter information). In these simulations, in the no-preview condition, lexical processing of the target words in the Schilling, Rayner, and Chumbley (1998) corpus was not allowed to start until after the model's "eyes" land on or to the right of the blank space immediately to the left of the target words. In the preview condition, normal parafoveal processing of the target words was allowed. The duration of $t(V)$ was set equal to 50 ms in one simulation and 90 ms in the other. The simulations (see Table 3) indicate that assuming that the duration of $t(V)$ was 50 ms is at least close to optimal in that it simultaneously fits the normal reading data (see Table 2) and predicts a value for preview benefit that is similar to that typically observed in boundary experiments—40–50 ms (see Hyönä et al., 2004).

As a result of these first simulations, we think we have determined parameter values for the model that are reasonable and that successfully account for the 40–50 ms difference between full preview and no preview conditions while still being able to explain the corpus of reading data that we had successfully simulated before. Moreover, given the constraints that this paradigm imposes, these parameter values are far from arbitrary and should be taken seriously as estimates for how long the various mental processes take. In particular, we think that the convergence of visual processing time imposed on the model to explain the boundary experiments and that implied by the physiological data is quite striking (Clark et al., 1995; Foxe & Simpson, 2002; Mouchetant-Rostaing et al., 2000; Van Rullen & Thorpe, 2001).

6.1. When does cognitive processing of information on a fixation stop?

As discussed above, a key question is whether lexical processing in normal reading continues between the time the saccade is initiated (which displaces whatever information is on the retina) and the time that the visual information on the next fixation

Table 4

A comparison of how well E-Z Reader and Morrison's model account for the aggregate data of the Schilling et al. (1998) sentence corpus

		Word frequency (# Occurrences per million)				
		1–10	10–100	100–1000	1000–10,000	10,000+
Gaze durations	Observed	293	272	256	234	214
	E-Z Reader 9	297	275	242	221	216
	Morrison's model	292	275	238	227	222
First fixation durations	Observed	248	234	228	223	208
	E-Z Reader 9	252	239	225	217	214
	Morrison's model	250	239	223	223	220
Single fixation durations	Observed	265	249	243	235	216
	E-Z Reader 9	260	244	226	217	214
	Morrison's model	255	243	222	223	220
Probability of making one fixation	Observed	.68	.70	.68	.44	.32
	E-Z Reader 9	.74	.75	.75	.52	.42
	Morrison's model	.73	.73	.74	.50	.40
Probability of making two fixations	Observed	.20	.16	.10	.02	.01
	E-Z Reader 9	.16	.13	.06	.01	0
	Morrison's model	.15	.12	.06	.01	0
Probability of skipping	Observed	.10	.13	.22	.55	.67
	E-Z Reader 9	.09	.12	.19	.47	.58
	Morrison's model	.11	.15	.20	.50	.60

Note. The following parameter values were used with Morrison's model: $\alpha_1 = 200$; $\alpha_2 = 5$; $\alpha_3 = 21$; $\lambda = .09$; $R = 115$; and $\varepsilon = 1.03$. All other parameter values were fixed equal to those used by E-Z Reader 9 (see Table 1). The overall goodness-of-fit metric for the models are: E-Z Reader 9 = .153, and Morrison's model = .148.

reaches the higher cognitive systems that are involved in lexical processing. (We will discuss more artificial situations where the text disappears during a fixation below.) An alternative possibility is that lexical processing simply stops whenever a saccade begins and only re-starts after the visual information from the next fixation location has been encoded sufficiently for the orthographic and/or phonological processing to be able to continue. (We should emphasize that the visual information from the two fixation locations is not identical; among other things, the visual information extracted when the word is fixated is of higher quality than the visual information that is extracted from the parafovea; Rayner & Morrison, 1981).

Although this alternative assumption about how the deeper processes work may seem reasonable, we thought that it was problematic for a couple of reasons. First, introspectively, it does not seem like there are any blank periods when one is not processing visual information. Second, as indicated earlier, if the model assumes that it takes the visual information a certain amount of time to reach the brain before processing can start, then it seems inconsistent for the model to posit that the news that the visual information has disappeared from the retina reaches the brain instantly. This is why we started with the assumption in the prior section that the deeper lexical processes continue to work on the visual codes from the previous fixation until new

and better visual information becomes available following the eye–mind lag. However, thinking of $t(V)$ as the time from the beginning of a fixation for adequate information to arrive at the brain is only one way to conceptualize such a visual stage. More importantly, we thought it was best not to rely on plausibility arguments, and instead that it was important to test the assumption that lexical processing of the information from the prior fixation continues through the saccade and until the information from the new fixation displaces it. For the sake of simplicity, we will focus below on two assumptions about the relation between the visual processing stage and later lexical processing that can be viewed as extreme cases. The first (which we refer to as the “discontinuous-processing” variant of our model in the exposition below) is that lexical processing stops immediately when the saccade begins and remains suspended until the visual information from the new fixation location reaches the brain and is available for further lexical processing. The second, “standard” version of our model is the one that has been assumed so far (i.e., used in the simulations reported in the Appendix A): when a saccade is made, lexical processing continues for as long as it takes the visual information from the next fixation to go from the eye to the brain (i.e., the visual processing time on the subsequent fixation), and then continues—using the visual information that is available from the new viewing location.

6.1.1. *Two models of the vision–cognition interface*

Eqs. (1)–(3) in the prior section summarize the basic assumptions about how long the various stages of visual and lexical processing take and which variables they are affected by. As indicated above, however, the equations do not specify exactly how the initial visual processing stage relates to the later lexical stages, nor do they specify how this interface works in the context of ongoing saccades and new viewing locations. We now describe two alternative models of the vision–cognition interface in some detail.

6.1.2. *The “discontinuous-processing” model*

This first way of conceptualizing the interface between visual processing and word identification was instantiated by E-Z Reader 7 (Rayner et al., 2003). In the discontinuous-processing variant of the model, the visual processing stage begins as soon as a new fixation begins, but lexical processing cannot begin until after visual processing has been completed (i.e., until after adequate visual information has reached the higher perceptual centers). Furthermore, this model assumes that when the next saccade is initiated, lexical processing is halted (i.e., no further progress on lexical processing takes place) until the visual processing stage is completed from the subsequent viewing location. If the fixated word is identified on this second fixation, then attention will shift to the next word and lexical processing of the next word will begin. Thus, lexical information typically accrues when a word is seen parafoveally, but this accrual stops for a while and then resumes some time after the word is fixated. Additional pauses in lexical processing can also occur if a word is refixated. Finally, it is worth noting that both models implicitly assume that visual processing is completed in parallel across the visual field.

6.1.3. The “standard” model

This second conceptualization of the vision–cognition interface was discussed in an earlier version of E-Z Reader (Reichle et al., 2003). The primary difference between the two models is that, in the standard model, lexical processing does not have to stop when a saccade begins. That is, our standard model assumes that the visual feature information that the lexical processing system is using does not instantaneously disappear with the movement of the eyes to a new viewing location. Instead, the visual information *that is attended to* is maintained and can be used by later processes until new visual information displaces it. One could conceptualize this as the information being maintained in a short-term visual memory “buffer.” Thus, the core assumption of our standard model is that, after lexical processing begins, it continues using whatever attended information happens to be available from visual processing (Irwin, 1998). Because the standard model assumes that lexical processing uses whatever visual information happens to be available in its buffer and because visual processing is assumed to occur in parallel across the entire visual field, lexical processing is only significantly affected by the delay in the visual processing stage in cases where the eyes move so far that the visual information seen on the first fixation was of no use for the lexical processing of the word fixated after the saccade, making it necessary to wait for visual processing to finish before starting lexical processing on visual information from the new viewing location. Such instances might occur after long regressions and/or return sweeps, and when sentences are first displayed during eye-tracking experiments.

To summarize, the standard model assumes that lexical processing of text is usually continuous. On the first fixation on the text, lexical processing has to wait for visual processing to complete before it can start. However, after that, there are no “breaks” except when there is no useful information in the parafovea. That is, at the end of most saccades, lexical processing is assumed to continue to operate using the visual information that was accrued on the prior fixation (and that was maintained in some sort of buffer memory), and that this “old” visual information disappears only when the visual information from the new fixation bumps it out of this buffer memory.

6.1.4. Simulations

We compared the discontinuous-processing and standard variants of the model on two different criteria. The first was to see whether they were equally successful in modeling the standard reading data. The second was to see what the models predict about preview benefit. For the first test, the discontinuous-processing model was fit to the same Schilling et al. (1998) sentence corpus that was used to evaluate our standard model (see Appendix A). The result of this simulation indicated that the discontinuous-processing version did as good of a job predicting the observed means for all of the standard word-based measures (e.g., gaze duration, probabilities of making a refixation, etc.) as the standard model, and that the fits were comparable to those of our previously published best fits—an RMSD of 0.149. (Our standard model produced an RMSD of .153; see Table 2). However, although the models predicted the mean fixation durations equally well, they did not perform equally well in

predicting the fixation duration distributions. Fig. 4 shows the observed first-fixation and gaze-duration distributions for the five frequency classes of words in the Schilling et al. (1998) corpus, along with the distributions predicted by the discontinuous-processing model (cf. Fig. 6, Appendix A). As can be seen, this variant of our model seems to be problematic because it predicts that the distribution of fixation durations is somewhat bi-modal, with a second smaller peak at about 100 ms. This is more obvious for the high-frequency words. This result is clearly at odds with the observed data. The reason for this discrepancy is clear: on some fixations, attention shifts to a new word and it is identified before the saccade to the next word can be executed. That is, the lexical processing of $word_n$ (while fixating $word_{n-1}$) is fast enough to beat the saccade program to $word_n$, but not fast enough for it to initiate a saccade to $word_{n+1}$ that would cancel the program to fixate $word_n$. These fixations on $word_n$ would therefore be quite short. In contrast, if $word_n$ is not processed fully before the saccade to it, then lexical processing has to wait until the visual processing stage on the new fixation finishes before lexical processing can be renewed. We think it is quite unlikely that this problem with the discontinuous-processing model is due to some other aspect of the model (i.e., besides the visual processing assumption being incorrect). We will discuss that issue later.

The second test of whether lexical processing continues through the saccade and into the next fixation is whether the models correctly predict preview benefit. This test was informative; the discontinuous-processing version of E-Z Reader predicted a preview benefit effect that is far too small: 20 ms for the gaze durations and 19 ms for the first-fixation durations. The reason for this is quite transparent. That is, it posits that the amount of head start time for lexical processing that a word gets from parafoveal information is the interval between when attention shifts to $word_{n+1}$ and when the saccade begins moving the eyes to $word_{n+1}$, which is the saccadic programming time minus the L_2 processing time for $word_n$ (see Fig. 1). Because saccadic programming times average 125 ms and L_2 times are typically about 60–70 ms, this only allows 35–45 ms of processing time on the word before the word is processed from the fovea. However, because the rate of lexical processing in the parafovea is substantially slower than the rate of foveal processing, this means that the predicted size of the preview benefit is going to be too small. Of course, one might try to remedy this problem by assuming that there is virtually no drop-off in the rate of lexical processing as eccentricity increases [see Eq. (2)]. Besides being implausible, this solution raises other problems that will be discussed in a later section.

6.1.5. Conclusions

Our simulations are consistent with one's introspection about visual processing: when reading or viewing a scene, there does not seem to be any down time when the system is not working. Of course, we have not explored all possible models of how early visual processing and later lexical processing might interface across fixations; however, our simulations indicate that the discontinuous-processing version of E-Z Reader is unlikely to be even a good approximation of the vision–cognition interface and that this interface is probably close to what is posited in our standard model. As an aside, it is also interesting that changing the assumption about whether

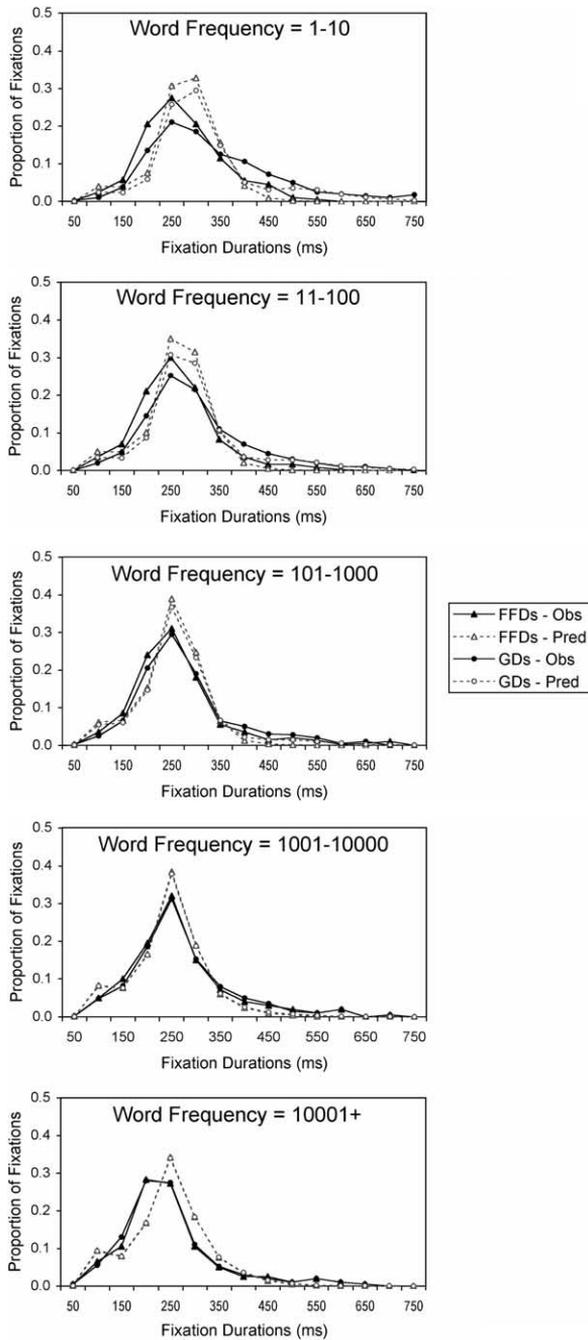


Fig. 4. Observed and predicted first-fixation and gaze duration distributions for the five frequency classes of words in the Schilling et al. (1998) corpus, and the distribution predicted by E-Z Reader under the assumption that lexical processing halts during saccades.

there is down time in lexical processing has fairly dramatic consequences for what one observes in normal reading, whereas changing the assumption about the duration of the visual processing stage only appears to have significant consequences in display-change situations.

It is worth mentioning, however, that the assumptions of the discontinuous-processing model might be viewed as being a bit too extreme. An alternative to these assumptions is the following: given the assumption that lexical processing continues until the brain receives a signal that the information in the visual field has changed, one could argue that (a) lexical processing of the parafoveal information continues until the new visual information reaches the brain (i.e., when visual information resulting from the saccade onset reaches the brain), and that (b) lexical processing starts up again when the foveal information reaches the brain. This would imply a down time equal to the saccade duration (approximately 25 ms). Such a model, however, would be quite hard to distinguish from the standard, continuous-processing, version of our model.

Another possible objection to the continuous-processing assumption is that there needs to be some adjustment period before the lexical system can start working on the visual information from the new fixation location. This could either be because the quality of the visual information is different across successive fixations, or because the attention system needs to “find” the word that is being attended in its new retinal location. These could be modeled by including a down time at some point during each fixation, and this would be formally similar (in terms of the model) to assuming a brief down time equal to the duration of the saccade. However, the extremely good fits of the discontinuous-processing model suggest that, if there is such down time in processing, its duration is likely to be quite brief. We should also add that it is not necessary to envisage this down time as being due to nonsensical retinal information coming in during the saccade. The extensive work on backward and forward masking indicates that the type of information that follows or precedes a stimulus has a large influence on how well the information masks the stimulus. In particular, the literature on backward masking suggests that a stimulus that has similar features and is retinally (or spatially) near another stimulus masks effectively, whereas a stimulus lacking distinct features (e.g., a patch of light) is usually ineffective unless the mask is overwhelmingly brighter than the target stimulus (Alpern, 1953; Eriksen & Lappin, 1964). Thus, it is plausible that the retinal stimulation occurring during the saccade is coded by the visual cortex largely as motion (blurs), and that this information does little to mask the stimulus seen on the prior fixation. The information from the prior fixation may decay slightly over the period when the visual information from the saccade is coming to visual cortex, but the saccade duration is brief enough that this decay might not be too noticeable. (It is also worth pointing out that there are data showing that the visual information from the fixation after the saccade serves as an excellent mask of the visual information coming in during the saccade, which is why we generally lack awareness of this information; Campbell & Wurtz, 1978.)

In sum, our modeling exercise indicates that the visual processing assumptions of our standard model are at least a good approximation of reality. As discussed above,

there are some reasons to expect some disruption of processing due either to the noise coming in during the saccade, or to having to cope with the differing quality or retinal locations of the stimulus across the two fixations. If so, then it appears that this disruption in lexical processing is slight and needs a more sensitive test than our global fit of the fixation duration data to detect it.

7. Disappearing text: The role of visual attention in preserving visual information

The prior section concentrated on modeling the effects of “normal” disappearance of visual information in reading; that is, the disappearance of the visual information at the end of a fixation. But what happens to processing when more explicit breaks in the stimulus occur, either due to the stimulus disappearing during a fixation or being replaced by a pattern mask? In some cases, the answer is virtually nothing. For example, participants’ reading rates are virtually the same as in normal reading when the fixated word disappears or is masked after 50–60 ms (Ishida & Ikeda, 1989; Liversedge et al., 2004; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981; Rayner et al., 2005; Rayner et al., 2003). This indicates that tying an argument about the use of visual information too closely to one’s introspection about visual persistence is hazardous. That is, one might have concluded from the prior section that visual information remains available for further processing even after it disappears, but only if and *only if* one is not aware of its disappearance. However, the Liversedge et al. (2004) and Rayner et al. (2003) data indicate that some type of information remains usable even after the visual information has disappeared phenomenally.

These data on disappearing text also show that lexical processing does not necessarily cease after the visual information disappears. But how do we best understand this? There appear to be two logical possibilities within our framework of a qualitative distinction between early visual processing and later lexical processing. The first is that visual processing automatically continues throughout a fixation, despite the text disappearing or being replaced by a mask. The second is that visual processing ceases when the stimulus disappears⁸ or is masked, but that by 50–60 ms after the beginning of the fixation, enough of the relevant information has been transferred to the lexical system to make its continued presence in the visual system largely irrelevant. The former explanation seems fairly implausible because there are quite appreciable disruptions in performance when a mask replaces the text at delays of less than 50 ms (Rayner et al., 1981), so that one does not want to say that the masking or disappearance of the stimulus is irrelevant to whether visual processing continues. Thus, the latter explanation—that it takes some critical amount of time for the visual information to be transferred to some more durable form of representation that can be processed by the lexical system—seems preferable. (Deciding whether

⁸ When the text merely disappears, there is, of course, both some phosphor persistence on the screen as well as iconic persistence (Sperling, 1960) in the visual system. However, Rayner et al. (2003) did employ a shutter test in which a shutter opens simultaneously with the offset of the text, and this indicated that no useful persistence remained from the words.

this more permanent storage is “in” the lexical system or is a separate system seems fruitless until one has a more complete theory of what a lexical processing system is).

This argument implies that lexical processing continues almost unabated when visual stimulation is withdrawn within a fixation if it has been visible for 50–60 ms. This indicates that our conclusion (in the prior section) that visual processing continues during the entire fixation and during the saccade that follows it may have been premature, although these data do not contradict it. However, the essence of our conclusion—that lexical processing continues as if the visual information were still present during the saccade and until the visual information from the next saccade reaches the brain—still seems quite plausible.

This last statement raises an important question: is *all* of the relevant visual information that impinges upon the retina during a fixation *automatically* encapsulated into a more permanent representation (that persists during the fixation) if the information is only visible for the first 50–60 ms, or does attention play a role? Specifically, is attention to a word necessary for the information to be encapsulated? This question is relevant to the E-Z Reader model because attention is assumed to move sequentially from $word_n$ to $word_{n+1}$, and usually moves during the period when $word_n$ is fixated. Fortunately, this question is testable because, if the text is masked or disappears after 50–60 ms, then the information about its disappearance should reach the lexical processing system 50 ms later, or 100–110 ms after the beginning of the fixation. If only the attended information is “fixed” so that it can be processed after the visual information has disappeared, then one would predict a substantial difference between situations when: (a) the fixated word ($word_n$) disappears after 50–60 ms, and (b) both $word_n$ and the next word ($word_{n+1}$) disappear after 50–60 ms. That is, at 100–110 ms after the beginning of a fixation (which is when the impact of the disappearance of the text would reach the lexical processing system) the probability is virtually equal to one that the fixated word is either: (a) the current focus of attention or (b) has already been fully processed up through the L_2 stage. (The only exception would be cases where the currently fixated word was wrongly fixated due to saccadic error.) Thus, under the assumption that only the attended visual information is “fixed,” there should be little or no disruption on processing from the disappearance of the fixated word after 50 ms. In contrast, there is an appreciable probability (given the assumptions and parameters of the E-Z Reader model) that the fixated word is still being processed 100–110 ms after the beginning of a fixation, and thus that attention has not shifted to $word_{n+1}$ by that time. As a result, if attention is needed to encapsulate the visual information, then there should be noticeable disruption in reading if $word_{n+1}$ disappears after 50–60 ms (regardless of whether $word_n$ disappears).

Thus, one can contrast two versions of the E-Z Reader model: in an “attention-irrelevant” variant of the model, the attentive processes it posits are irrelevant to maintaining the visual information that is needed for lexical processing (i.e., only the amount of time that the word is visible at the beginning of fixation and the masking quality of whatever follows influence the amount of visual information that is maintained). In the other “standard” variant of the model (i.e., the model used to complete the simulations in the Appendix A), the attentive process that is controlling

which word is selected for lexical processing is crucial for maintaining the visual information for lexical processing. The former version predicts no disruption in performance caused by disappearance of the text as long as it is visible for 50–60 ms at the beginning of the fixation, whereas the latter version predicts disruption, but mainly when material to the right of the fixated word disappears.

Here is a qualitative argument for the latter prediction. If one assumes that the attended material is encapsulated 50–60 ms after it arrives in the relevant cortical processing areas, then there should be little disruption when only the fixated word disappears because the probability is essentially equal to one that the fixated word is either attended to or has already been processed by that time. (The reason for the hedge ‘essentially’ is that there is some probability that $word_n$ was wrongly fixated, and that $word_{n-1}$ was actually the word initially being attended to during the fixation; in these cases, attention might not shift to $word_n$ until after the text had already disappeared.) In contrast, on an appreciable fraction of the trials when $word_n$ is fixated, attention does not shift to $word_{n+1}$ before the text disappears. On these trials, there should be no useful preview benefit, and the disappearance of $word_{n+1}$ should not affect the fixation measures on $word_n$ because $word_n$ was encoded normally. However, the disappearance should significantly increase fixation duration on $word_{n+1}$. (Again, we have to hedge a bit here, as there can be mis-targeting of saccades.)

Fortunately, there are data relevant to this issue of what happens with disappearing text (Rayner et al., 2005). In conditions where only the fixated word ($word_n$) disappeared after 60 ms, reading was virtually unaffected, as was the word frequency effect (as predicted by both versions of the E-Z Reader model above). However, when both $word_n$ and $word_{n+1}$ disappeared after 60 ms, there was a significant cost—about a 30% decrement in the overall reading rate. Qualitatively, this is incompatible with the attention-irrelevant variant of our model (above), but is compatible with the standard version of E-Z Reader. However, to actually explain the results, one has to explain how the encapsulation process works. For the sake of simplicity, we will assume that the encapsulation process is part of what it means to complete the familiarity check on a word, and that the conversion of visual codes into orthographic and/or phonological codes provides a more robust representational format for the word that is being processed. Thus, in the simulations that are reported below, we will assume that the encapsulation process requires no additional time, and that it begins on a word as soon as the reader shifts attention to that word.⁹

Before getting into the details of the modeling, we need to make an important point about the data being simulated. The preservation of the word frequency effect on first fixation duration and gaze duration is strong evidence that lexical processing of the fixated word is the primary determinant of the fixation time on that word, and that, at least in some cases, lower-level visual aspects of the situation, such as the

⁹ We realize that our assumption that the completion of the encapsulation process is virtually instantaneous is a bit unrealistic. However, we have also not assumed any iconic persistence after the text disappears, and adding such an assumption would give a bigger window for the encapsulation process to complete. It is also possible that this encapsulation process is speeded by the signal that the information has disappeared.

sudden offsets of visual information, play a minor role. That is, if the offset of a fixated word was the trigger for a substantial percent of the eye movements from a word, then this would imply that the frequency effect would either disappear or be substantially attenuated. The preservation of the frequency effect in the face of these display changes reinforces the central assumption of the E-Z Reader model that some aspect of the processing of the fixated word is the primary signal for moving the eyes forward. It is not clear whether or not a model that assumes parallel processing of multiple words and a “clock” to drive eye movements (e.g., SWIFT; Engbert et al., 2002, 2005) would make the same predictions because the primary determinant of when the eyes move is not the identification of a specific word.

We should also make it clear that we are not claiming that the eye movement system is oblivious to low-level factors such as the sudden onsets and offsets of visual stimuli. Indeed, some of the data of Rayner et al. (2005) reflect these low-level factors. In their experiments, display changes occurred on all words (rather than a selected set of target words) so as to minimize the surprise due to these low-level visual changes. However, they were unlikely to eliminate reactions to these visual changes completely. Another feature of their experiments was that, once a region of text disappeared, it did not reappear until after the reader had fixated a word other than the one that was fixated when the text disappeared. Either of these aspects of the Rayner et al. (2005) experiments may have induced special strategies in the readers that would not be that interesting to model because (a) dealing with this kind of text disappearance is not part of normal reading and (b) it would involve adding several largely ad hoc assumptions. As a result, our modeling efforts below are directed towards simulating the general pattern of the results from these experiments. To that end, we did not attempt to fit their data, but instead attempted to simulate the effects of these disappearing-text conditions on the Schilling et al. (1998) corpus, without making any assumptions that readers adopt special strategies to deal with the disappearance or that the eye movements of readers are affected by the transients that accompany the offset of the text and/or the onset of a mask. These simulations should thus be viewed as giving a “zero-order” approximation to the pattern of data.

7.1. Simulations

A series of simulations were completed to determine if E-Z Reader 9 can in fact account for the finding that removing both the fixated word and the word to right of fixation shortly after fixation is more detrimental to reading than simply removing the fixated word. We ran six simulations corresponding to a subset of the conditions reported by Rayner et al. (2005). These conditions involved three types of display changes (no display change vs. $word_n$ disappears vs. $word_n$ and $word_{n+1}$ both disappear) crossed with two target-word ($word_n$) frequencies (1 vs. 105 per million, as tabulated by the CELEX database; Baayen, Piepenbrock, & Gulikers, 1995), resulting in a 3×2 factorial design. We were thus able to look at how the disappearance of $word_n$ and/or $word_{n+1}$ affects the fixation durations on these words, and how this manipulation interacts with the frequency of $word_n$.

One key issue had to be considered, however, before we could complete our simulations: What do readers do when attention is directed towards a region in which useful letter information has disappeared? The question is critical in the context of our model because word identification is the engine that drives the eyes forward, so that the absence of letter information in the attended location would cause this engine to stop and thereby halt the forward progression of saccades. To address this problem, we adopted the assumption that there is some maximal amount of time that the reader will wait without having successfully completed the familiarity check on a word before he or she will move his or her eyes to a new viewing location. This assumption is consistent with the simple notion that readers, upon encountering a blank space, will quickly move their eyes to the next word so as to continue reading. In the simulations reported below, we set this deadline for completing the familiarity check equal to 500 ms.¹⁰ It is important to note that neither this assumption nor the particular deadline value should be viewed as being essential parts of our model; they were instead adopted out of convenience, to simply make it possible to run the simulations. For this reason, the simulation results that are reported below should only be evaluated qualitatively—the relative differences among the conditions should be informative for evaluating the assumption (that was discussed in the preceding section) that attention is needed to encapsulate the features of words, but any absolute difference among the conditions depends upon the specific value of our processing deadline assumption.

Given the qualitative nature of our simulations, we also opted to focus on specific target words rather than all of the words in each sentence. That is, instead of having either one or two words disappear (i.e., the fixated word or both the fixated word and the next word) with each new fixation in each sentence, we opted to examine the consequences of having either one or two words disappear when a designated target word ($word_n$) was fixated. We did this primarily because it minimizes the potentially negative consequences of making a bad guess about the processing deadline (because the effects that result from incorrect values will not cascade across multiple words). In the simulations, the Schilling et al. (1998) target words (see Appendix A) were designated as being “ $word_n$ ” (i.e., the target words of interest in our simulations). The lengths of these words were then set equal to six letters, the mean length of the Rayner et al. (2005) target words. To minimize any contribution of predictability, the predictability of these words was also set equal to zero. Finally, the frequencies of these words were set equal to the mean values of the Rayner et al. (2005) target words: 105 per million (as tabulated in the CELEX database) in the three conditions involving high-frequency targets, and 1 per million in the three conditions involving low-frequency targets. Six simulations were then completed using 1000 statistical subjects per simulation and the same parameter values that were used in the other simulations reported in this article. Depending upon the condition being simulated, $word_n$ was either high- or low-frequency, and one of three things happened: (1) all words remained visible; (2) $word_n$ disappeared 60 ms after it was first fixated; or

¹⁰ Smaller values of the processing deadline lead to the same qualitative pattern of results.

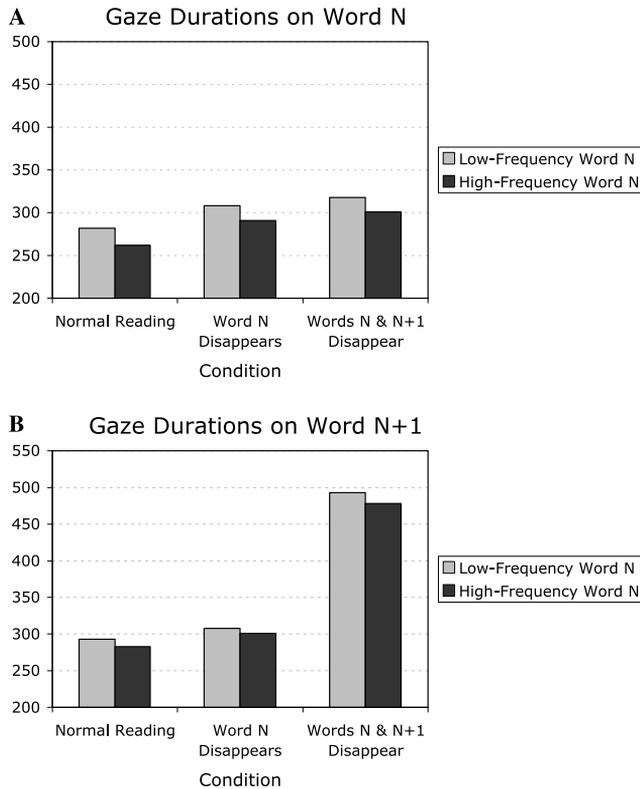


Fig. 5. Mean gaze durations predicted by E-Z Reader 9 in the disappearing-text paradigm. (A) The gaze durations on $word_n$, conditional upon that word's frequency of occurrence in printed text. (B) The gaze durations on $word_{n+1}$, also conditional upon the frequency of $word_n$. Both panels show gaze durations in three display-change conditions: (1) when no change occurs and both $word_n$ and $word_{n+1}$ remain visible (i.e., normal reading); (2) when $word_n$ disappears 60 ms after being first fixated; and (3) when both $word_n$ and $word_{n+1}$ disappear 60 ms after either word is first fixated.

(3) $word_n$ and $word_{n+1}$ disappeared 60 ms after either word was first fixated (usually $word_n$).

Fig. 5A shows the predicted gaze durations for $word_n$ as a function of that word's frequency of occurrence in printed text. (Because first-fixation durations show the same qualitative pattern as the gaze durations, we will limit our discussion to the latter.) The important points to note in the figure are that: (1) the gaze durations on high-frequency targets are shorter than those on the low-frequency targets; (2) this frequency effect is approximately the same size across the three display-change conditions; and (3) there is only a modest increase in the gaze durations going from the normal reading (no display change) condition to the condition in which $word_n$ disappears to the condition in which both $word_n$ and $word_{n+1}$ disappear. This pattern can be explained by the fact that in both of the display change conditions, approximately 14% of the trials resulted in a situation in which the model shifted attention to $word_n$

after that word had already disappeared. This is largely due to the fact that the saccades that are generated by the model are prone to error, so that eye movements that are directed towards $word_{n-1}$ sometimes overshoot their intended target and prematurely start the “clock” that will make $word_n$ disappear 60 ms later. (This predicted cost may be an overestimate because participants may move their eyes more carefully in this paradigm.) In such cases, the model must “wait” until the processing deadline (500 ms) has been met before it can initiate a program to move the eyes to the next word, which will in turn lengthen the gaze duration. (In this situation, both $word_{n-1}$ and $word_n$ are processed from $word_n$, which also causes the gaze duration on that word to be longer than it otherwise would.)

Despite this, the fact that the frequency of $word_n$ had a normal effect on the gaze durations indicates that $word_n$ was identified in the majority of the simulation trials and that the familiarity check on it was the trigger for moving the eyes to $word_{n+1}$. This frequency effect is consistent with the results reported by Rayner et al. (2005; see also Rayner et al., 2003; Liversedge et al., 2004), and it supports our contention that 50–60 ms is enough time to “fix” the visual codes of words being attended so that more abstract (and robust) lexical codes can be used for further processing. However, the model predicted that $word_{n+1}$ was not identified on the majority of the simulation trials. (On approximately 76% of the trials, the model shifted attention to $word_{n+1}$ after that word had already disappeared, thereby causing the model to wait for the processing deadline before proceeding to $word_{n+2}$.) This is indicated in Fig. 5B, which shows that the gaze durations on $word_{n+1}$ increased quite substantially in the conditions in which both $word_n$ and $word_{n+1}$ disappear. Thus, in contrast to the condition in which only $word_n$ disappears, the condition in which both $word_n$ and $word_{n+1}$ disappear would be expected to produce a general slowing in the overall reading rate—a prediction that is also consistent with the results reported by Rayner et al. (2005) and that again supports our contention regarding the role of attention is word identification.

7.2. Conclusions

The modeling above indicates that the sequential attention assumption of the E-Z Reader model is quite compatible with the Rayner et al. (2005) data. That is, it successfully predicts that there is little or no decrement in processing when only the fixated word disappears after 60 ms, but that there is an appreciable cost when the word to the right of fixation disappears as well. In these simulations, a key assumption, of course, was that attention to a word was necessary to “fix” that information in some sort of “store” so that lexical system could continue to work on it.

8. Are two stages of lexical processing necessary?

As indicated in the introduction, the two stages of lexical processing in the E-Z Reader model were largely posited to explain (a) that preview benefit was modulated by foveal difficulty of processing (Henderson & Ferreira, 1990) and (b) spill-

over effects (i.e., that the frequency of a word influences processing after the word was fixated, most notably increasing fixation times on the following word; Rayner & Duffy, 1986). There has been some controversy over whether the first phenomenon is reliable (Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999), although a number of studies have now replicated the original results (see Drieghe et al., 2005; Kennison & Clifton, 1995; White et al., 2005). In addition, there could be other causes for spillover effects that are beyond the scope of the E-Z Reader model. That is, a low frequency word may not only be difficult to access lexically, but its meaning may only be known approximately, so that it causes difficulty in text comprehension—a part of the reading process we have not attempted to model. As a result, it seemed advisable to return to the question of whether positing two stages of lexical processing is necessary to explain the basic eye movement data (i.e., the time spent on a word—the gaze duration—and probability of skipping it as a function of its frequency).

A major constraint in all our modeling work has been to fit the data with parameters that are reasonable values for the processes involved. That is, it would be trivial to do much of the modeling work in reading if one can assume arbitrarily short times for lexical access (or other acts of language comprehension) or arbitrarily short times for the programming of eye movements. The necessity to have realistic parameter estimates for psychological processes places a greater constraint on a model that posits only a single stage of lexical processing (such as Morrison's, 1984) because it posits that full lexical access is the trigger for an eye movement program. In contrast, the E-Z Reader model has an extra degree of freedom, because the trigger for an eye movement program is the familiarity check stage, which, in our usual interpretation of the model, is a partial stage of word identification that is sufficiently advanced so that full identification of the word is likely to occur shortly, thereby providing a cue that it is "safe" to program an eye movement to the next word.¹¹ (That is, the familiarity check ensures that the saccade to the next word is unlikely to be premature and thus unlikely to result in processing difficulty that would result in a regression).

The above discussion raises the following question. If full lexical access is the trigger for both the saccade program and the shift of attention, is there enough time for lexical access to influence fixation durations during reading—given that most fixation durations are on the order of 200–250 ms? In our prior modeling work (Reichle et al., 1998), we attempted to answer this question by testing a one-stage model (which we will call 'Morrison's model') and it did not fare very well. However, this may not have been a fair test of the one-stage model, as some of the parameter restrictions on it may have been too restrictive. Most notably, the one-stage model

¹¹ Another logical possibility, of course, is that there is a single trigger for both the attention shift and the eye-movement program, but it is something like the familiarity check. We are not going to discuss this possibility in depth because we think that this would be a fairly unsatisfactory way to model reading unless there was some independent way to assess this familiarity check stage. Such a model would give parameter values for how long it takes to complete the familiarity check, but these values would be of limited utility unless there was some way to independently determine if they were reasonable or corresponded to processes or phenomena described in the word processing literature.

was required to have the same parameters as the E-Z Reader model in the equation describing the fall-off of efficiency of lexical processing as a function of eccentricity from the fixation point (i.e., the equivalent of Eq. (2) above). However, if one posits a less severe drop-off of efficiency for a one-stage model, then it might be able to fit the reading data with reasonable estimates for lexical access time by taking greater advantage of parafoveal processing than E-Z Reader does. Thus, it seemed worthwhile to be sure that a one lexical-stage model with reasonable parameters couldn't account for the reading data. The first goal was to see whether it could give a reasonable account of the Schilling et al. (1998) corpus of data, and then, if it did, to determine if the model, together with the parameters that were needed in this modeling exercise, do a reasonable job explaining other reading phenomena. In particular, we thought that a critical test would be to determine if such a one-stage model would give a good account of data in which parafoveal information was withheld (specifically boundary experiments) or whether it would predict too great a cost when this information is only first visible when a word is fixated because it relied too heavily on parafoveal information to be able to process words fast enough.

8.1. Simulations

To address these issues, it was first necessary to determine if a one (lexical) stage model can provide an adequate account of the Schilling et al. (1998) sentence corpus. In fitting the model to the data, it became clear that the model gave an overall fit to the data that was comparable to the one provided by E-Z Reader (see Table 4). The overall quality of the fit for the one-stage model was $RMSD = .148$ (as compared to E-Z Reader's $RMSD = .153$). It is important to note, however, that we were only able to obtain this fit with Morrison's model by relaxing the constraint that visual acuity slows the rate of lexical processing for long words and words that are far from the center of vision (cf., the values of the ε parameter for E-Z Reader vs. Morrison's model). However, apart from this one difference, both models achieve their fits using reasonable parameter values. If one includes the 50 ms that is necessary for visual processing time, then the E-Z Reader parameters specify the mean minimum (in cases where predictability < 1) and maximum times to identify words as being equal to 151 and 233 ms, respectively. (These estimates ignore the effects of visual acuity, which would in most instances increase the estimated values.) Likewise, Morrison's model (again assuming 50 ms for visual processing and ignoring the effect of visual acuity) estimates the minimum and maximum word identification latencies to be 173 and 250 ms respectively. Thus, on the basis of both how well the models predict the aggregate corpus data and the reasonableness of the parameters needed to do so, there is little basis for arguing that one model is vastly superior to the other.

The second simulation tested whether a one-stage model could explain the boundary paradigm (Rayner, 1975). That is, as indicated above, to achieve a reasonable fit with a one-stage model, the eccentricity parameters had to be relaxed (i.e., there was less of a decrement in processing as the distance from the fovea increased). The conditions we will contrast are those discussed earlier: (a) a *no preview* condition, in

which a designated target word is replaced by something like a string of random letters until the reader crosses an invisible boundary, at which point the preview is replaced by the target word (see Fig. 3, above); or (b) a *full preview* condition, in which there is no display change. We again used the Schilling et al. (1998) sentence corpus, and simply ran a pair of simulations using each model—one in which parafoveal processing was allowed to occur as it normally does (the full-preview condition), and one in which processing of designated target words did not begin until those words were actually fixated (the no-preview condition). The results of these simulations were quite diagnostic. The one stage model predicted a mean preview benefit of 69 ms for first-fixation durations and 87 ms for gaze durations. This contrasts both with the preview benefit values predicted by E-Z Reader 9 of 39 ms for first-fixation durations and 46 ms for gaze durations, and with the observed preview benefit values for gaze duration, which are typically around 40–50 ms (Hyönä et al., 2004). The results of our simulation thus indicate that a one-stage model needs to posit too much lexical processing in the parafovea to be able to identify words in a reasonable amount of time to provide a reasonable account of the standard reading data.¹²

In sum, the simulations reported in this section show that single-stage models such as Morrison's (1984) are ill suited to simultaneously handle the joint constraints of having: (1) reasonable word identification times; (2) reasonable saccadic programming times; and (3) limited visual acuity. Currents models of eye-movement control have been able to meet these constraints in two ways. The first approach (exemplified by the E-Z Reader model) is to posit that the signal to move the eyes from one word to the next is not full word identification, but is instead some preliminary stage of lexical processing (such as the familiarity check) that is sufficient to guarantee that most forward saccades will leave most words shortly after they have been identified. (This will insure that the majority of words are identified without having to make regressions; for an explanation for how skilled readers might learn how to do this, see Reichle & Laurent, 2005.) The second approach that has been adopted to meet these constraints is to simply assume that more than one word can be identified in parallel. By allowing the parallel processing of words, attention-gradient models like SWIFT (Engbert et al., 2002, 2005) and Glenmore (Reilly & Radach, 2003) sidestep the constraint that words have to be identified in a reasonable amount of time. However, as we will argue below, this “solution” is itself likely to be problematic.

9. General discussion

The above modeling hopefully indicates that the E-Z Reader model, though far from being a complete model of the reading process, is both a satisfactory explanation of a wide variety of phenomena and a valuable heuristic in exploring the reading

¹² If the restrictions on lexical processing times are removed, a one-stage model can actually provide even a somewhat better fit to the overall data than E-Z Reader 9 (RMSD = .101). However, this is at the cost of (a) having quite improbable minimum and maximum lexical processing times (259 and 330 ms, respectively) and (b) predicting ridiculously large preview benefits (181 ms on gaze duration).

process. We think that part of the reason that the E-Z Reader is a valuable heuristic is that, if it is a good approximation to how word identification interacts with the attentional and eye movement systems, it allows eye movement data to be interpreted in a way so that they are of great use in understanding the cognitive and linguistic processes involved in reading. For example, even though it is an oversimplification within the framework of the E-Z Reader model to interpret gaze duration as equal to the time to identify a word, the model indicates that it is a good approximation to word identification time and that it is reasonable to interpret differences in gaze duration between experimental conditions as approximately equal to differences in word identification time. However, the model makes clear that it is unlikely that the duration of a gaze is simply equal to the time to identify the fixated word. First, word identification usually starts before a word is fixated, and second, it is not complete identification of the word that triggers the eye movement off of a word. (In addition, eye movement latencies are not zero, although there should be a reasonable trade-off between the latency of the eye movement that lands on a word—during which processing of the word is going on—and the latency of the eye movement off of the word—during which processing of the next word is going on.)

A major reason why the E-Z Reader model both makes it easy to think about the relationship between cognition and eye movement control and provides a conceptual basis for using eye movement data to understand psycholinguistic issues is the assumption that attention is directed serially, word by word. Although we adopted this assumption partially for reasons of parsimony, we think there are deeper reasons why it makes sense and is the preferred assumption (at least as a very good approximation to reality). The first, which we articulated near the beginning of this article, is that human language developmentally and evolutionarily is primarily a spoken language and thus the brain structures that have evolved to understand language evolved to understand spoken language, which is inherently serial. Languages do differ in the importance of word order and English is admittedly a language in which word order is particularly important. However, we are primarily modeling English, and *run home* means something quite different from *home run* in the spoken language. As a result, if the reader, in trying to process these two adjacent words in parallel encodes them in an order other than going left to right, the utterance will be misinterpreted. We do not wish to suggest that this concern is fatal to parallel models (i.e., models such as SWIFT, in which several words are simultaneously attended on all fixations), as either (a) they can include additional mechanisms that sort out the spatial order or (b) perhaps the attentional weights can be set so that word order confusions rarely occur. However, if the latter is the case, then attention is primarily directed to one word at a time and they may be quite indistinguishable from a serial attention model such as E-Z Reader.

The second difficulty that parallel models often gloss over is how several words can be processed in parallel. That is, current models of word recognition are having a hard enough time trying to account for how letters within an isolated word of 4–6 letters are put together and whether even this process is as parallel as would be inferred from phenomena such as the word superiority effect (Reicher, 1969; Wheeler, 1970). A major conceptual problem is how letter order is maintained in such models without positing some type of “magic” device that “tags” the ordinal position of

each letter. In fact, many current models explicitly incorporate uncertainty in the position of letters (e.g., Davis & Bowers, 2004; Perea & Lupker, 2004; Ratcliff, 1981). However, if the system is attempting to process three or four words in parallel, the problem is magnified and the question of how to minimize crosstalk from adjacent words becomes a serious issue. Although there is evidence that people experience such crosstalk (Mozer, 1983), much of this research is not in the context of normal reading. Moreover, to the extent that such between-word crosstalk phenomena exist, it argues for the usefulness of directing attention serially to minimize or eliminate such problems.

A major justification for parallel models appears to be that they are faster than an apparently time-consuming strictly serial model. However, this is by no means clear. That is, to many people, parallel means a process in which everything is processed at once, and as rapidly as each of the entities that are processed in a serial process. However, such a model often seems to be “magic” and there is little justification for how such a process might work beyond some vague idea that individual neurons are parallel processors. However, our argument above about crosstalk indicates that if more than one word at a time is to be processed, one needs to bundle the information in each word into a separate packet. Moreover, it is necessary to keep the computations in these packets distinct in language decoding systems that are unlikely to have any natural segregation systems such as may exist in “maps” in early visual cortical systems. As a result, if there is parallel processing of several words, it is likely to occur in a parallel system in which attentional resources are distributed rather than focused, and it is far from clear that such a system would be more efficient than a serial processing system.

We also think that some of our modeling work above points to situations that may be difficult to model with a parallel model. First, consider the boundary experiment data we have modeled. We showed that Morrison’s model, which relies more heavily on parafoveal processing, predicts much too big a difference between situations in which $word_{n+1}$ is normal and those in which it is replaced by random letters. Because parallel models are extracting information from $word_{n+1}$ from the beginning of a fixation, they are also in danger of predicting too big a cost from withholding the letter information from $word_{n+1}$. Moreover, to the extent that they avoid this problem by positing a steep attentional gradient so that not much information is extracted from $word_{n+1}$, they become less distinguishable from a serial model. A key test of the difference between the models on this assumption may be the time course of extraction of information from $word_{n+1}$ while fixating $word_n$. In addition, in some of the parallel models (Yang & McConkie, 2001, 2004), the programming of a saccade is a “push-pull” decision based on competing levels of excitation from words, largely from $word_n$ and $word_{n+1}$. If so, we wonder how these mechanisms would be disrupted when $word_{n+1}$ is random letters and would likely be exerting little attractive force for the eye to move to it. If such a push-pull mechanism has to be abandoned, this may remove one of the rationales for adopting a parallel word identification assumption.

Second, the disappearing text situations may pose a problem for an attentional gradient assumption. That is, how does a model that posits such a gradient explain why there is such a difference between the situation where $word_n$ disappears and the situa-

tion where $word_{n+1}$ disappears? If such a model posits that there is some “fixing” mechanism applied to $word_n$ that does not apply to $word_{n+1}$, then where does this fixing mechanism come from? If it relies on an all-or-none attentional mechanism, then this requires positing two attentional systems, which seems unparimonious.¹³ Another possibility is imposing a threshold on the attentional gradient such that if the gradient is above threshold, the visual information is fixed, whereas it is not fixed if the gradient is below threshold. However, this seems to push the model in the direction of a serial model, but with a largely undetectable gradient.

Perhaps we should close by briefly discussing a putative phenomenon that we think draws people to parallel models and we will argue why this phenomenon is not an “Achilles heel” for the E-Z Reader model. This is really a set of phenomena often termed “parafoveal on foveal” effects (Kennedy, 2000; Kennedy & Pynte, 2005; Murray & Rowan, 1998). These phenomena are demonstrations that aspects of stimuli not yet fixated affect fixation time on a word and are often stated to be problematic for the E-Z Reader model. Some of these criticisms seem to ignore that, in the E-Z Reader model, significant processing of $word_{n+1}$ occurs when $word_{n+1}$ is fixated. Nonetheless, the general thrust of the E-Z Reader model is indeed that processing of the word initially attended to on a fixation is the primary engine that drives the eyes forward and that the difficulty of processing words to the right of the attended word will have little or no effect on the time spent on that fixation.

We want to make a few points about such parafoveal-on-foveal phenomena (see also Rayner & Juhasz, 2004; Rayner, Warren, et al., 2004; Rayner, White, Kambe, Miller, & Liversedge, 2003). The first is that quite a few of these findings are not in reading, but in other word identification tasks, where attentional strategies may be different. The second is that many of the strongest effects are not produced by linguistic aspects of $word_{n+1}$ such as its frequency, but instead occurs when it “looks funny” (e.g., due to a very strange sequence of letters, particularly at the beginning of the word; see Hyönä & Bertram, 2004). Third, some of the key results haven’t been replicated (Hyönä & Bertram, 2004; Rayner, Warren, et al., 2004). However, the most important point is that E-Z Reader explains that such phenomena can occur. The reason for this is that the E-Z Reader predicts that fixations can go to a word other than the word that was intended due to saccadic programming error. (This prediction is based both on detailed saccadic data, e.g., McConkie et al., 1988, and the general laws of motor control, and furthermore such occurrences are not rare, see also Rayner, Warren, et al., 2004.) Thus, if a saccade undershoots the targeted word (i.e., $word_{n+1}$), then $word_{n+1}$ is the attended word at the beginning of the

¹³ We realized when we were writing this section that there is a figure that we have used in several expositions of our model (e.g., Reichle et al., 2003, Fig. 3), that shows an “attentional spotlight” that is basically the area that is identified as being the field of useful vision from moving window experiments. This metaphor is potentially misleading because it suggests (a) a separate attentional mechanism than the one we are using in our simulations, and (b) that this whole area is attended simultaneously. Instead, this area is meant to represent information that *may be* attended to at some time during a fixation; that is, for words in this region, there is a probability greater than zero that (the 1-word) attentional spotlight will be directed on them on some fixations.

fixation but $word_n$ is the fixated word. As a result, in these cases, aspects of $word_{n+1}$ (such as its frequency) would be predicted by the E-Z Reader model to influence the fixation time on $word_n$. However, according to the E-Z Reader model, these phenomena should be largely limited to those cases when the fixation on $word_n$ comes fairly close to $word_{n+1}$, as most mistargeting errors are not large. This is in fact the typical pattern of results (e.g., Rayner, 1975; Rayner, Warren, et al., 2004). Moreover, we want to make clear that such effects are not rampant, and that, in general, information to the right of the fixated word has little or no effect on fixation time on that word—which would be the prediction of a model such as the E-Z Reader model which ascribes such effects to the mistargeting of saccades.

There is one arena in which the seriality of processing may need to be thought about more seriously. This is the question of how meaningful parts of words are processed. That is, there are data that indicate that long Finnish compound words (Hyönä & Pollatsek, 1998; Pollatsek, Hyönä, & Bertram, 2000) and long English words (Andrews, Miller, & Rayner, 2004; Juhasz, Starr, Inhoff, & Placke, 2003) are processed, at least in part through their constituents. That is, the time spent on these words is influenced both by the frequency of the word as a whole and by the frequency of the individual constituents. These data seem best accounted for by a model in which processing proceeds through two channels: a process involving recognizing the whole word and one that is recognizing the parts. If so, then there is some sense in which the components of these words are processed both in parallel and in series. However, there is evidence that the component processing may be chiefly in series. This comes from two within-word boundary experiments (Hyönä et al., 2004; Pollatsek & Hyönä, 2004) using such long Finnish compounds (which are obligatorily written without spaces), in which most of the second constituent of the word was replaced with random letters before the second constituent was fixated. Even though this display change produced a large increase in gaze time after the second constituent was fixated (relative to a condition where there was no display change), there was no effect of having an abnormal second constituent until it was fixated. This is very much a negation of parafoveal on foveal effects and seems inconsistent with parallel processing of even the constituents of a long compound word.

In summary, our goal in this article has been to refine and test the E-Z Reader model, and in doing so, to understand more fully the relationship between visual and language processing and eye movements in reading. We think that our modeling efforts point the way to several conclusions. First, the assumption that words in text are processed serially by skilled readers is not only a logical possibility but the most attractive hypothesis. Not only does our model employing this assumption fit the aggregate reading data, the serial processing assumption also appears to be the most parsimonious explanation of certain phenomena such as disappearing text (Liversedge et al., 2004; Rayner et al., 2003, 2005). It also accounts quite well for the decrement in reading observed in situations where the availability of text is restricted, such as moving window and boundary experiments. Second, using the framework of our model, we can draw certain conclusions about the nature of processing. Perhaps most importantly, we think our exposition has indicated that the parameters of our model are not arbitrary. The values that are best fitting are quite

constrained and are in accordance with what is known about visual and motor processing, such as the eye–brain lag and eye movement programming times. Of almost equal importance is the conclusion that a single process can not be the trigger for both an eye movement and a shift of covert spatial attention. In addition, our analyses show that lexical processing is essentially continuous throughout the reading process and does not experience downtime due to disappearance of the visual stimulus except in circumstances when attention does not get to the disappearing visual information in time. We acknowledge that parallel models are likely to be able to duplicate the predictions of the E-Z Reader model, as one can generally simulate a serial model at this level of description with a parallel model. However, we think that a parallel model that is able to simulate our model will be far less parsimonious and not nearly as good a heuristic device for using eye movements to understand language processing in reading.

We think that further tests of our model (or any model of reading) are going to require a more serious model of word recognition than we and other modelers have employed so far. We employed a “stand-in” (i.e., some equations positing relationships between the duration of L_1 and L_2 and word frequency and predictability) because our initial goal in modeling was to propose a simple and plausible architecture for how word recognition “talked to” the eye movement system. We believe that the work in the present article indicates that this architecture is on firm footing. However, the question of whether our two-stage model of lexical processing is the right elaboration of a one-stage model will likely revolve around fine detail of both the experimental data and of the assumptions about L_1 and L_2 . Consider the issue of spillover effects. The size of this effect in the E-Z Reader model is roughly equal to the duration of L_2 . (A one stage model can be viewed as a version of the E-Z Reader model with L_2 equal to zero and therefore with no predicted spillover effects.) Our current model’s assumptions about L_2 duration are crude—it is assumed to be a linear function of word frequency and predictability. However, there is evidence that many other factors affect lexical access. For example one variable that has been shown to have an influence on lexical processing is the number of *neighbors* a word has (i.e., other words that differ at only one letter position from a target word). In particular, in reading this has been shown to have an inhibitory effect and mainly on spillover (Perea & Pollatsek, 1998; Pollatsek, Perea, & Binder, 1999). Such a pattern is quite plausibly explained by a two-stage model of lexical access, such as an activation-verification model (Paap, Newsome, McDonald, & Schvaneveldt, 1982), in which neighborhood effects would only affect later stages of lexical access, and thus neighborhood effects would influence L_2 but not L_1 . Conversely, as we argued above, the two-stage model fits nicely with the Reingold and Rayner (2005) data, which indicated that large effects on gaze duration are accompanied by virtually no spillover effects, presumably because reducing the visibility of the word affects L_1 but not L_2 . Such an argument, at present, is just hand-waving, as no testable model of lexical processing is being offered. However, the argument was presented to indicate our view of how our modeling enterprise should go: using the E-Z Reader model format to examine models of word recognition in the context of reading.

Acknowledgments

This work was supported by Grants HD17246 and HD26765 from the National Institute of Health and by Grants R305H030235 and R305G020006 from the Institute of Education Sciences (Department of Education). We thank Alice Healy, John Henderson, Reinhold Kliegl, Gordon Logan, and Eyal Reingold for their helpful comments on an earlier draft of this article.

Appendix A.

Table 1 lists the parameters of the E-Z Reader 9 model together with their interpretations and their numeric values. (Unless otherwise noted in the text, these parameter values were used in all of the E-Z Reader 9 simulations reported in this article.) The values of seven of the parameters (V , M_1 , M_2 , S , Ψ , Ω_1 , Ω_2) were based on previous estimates (e.g., the saccade length that is not subject to systematic error, Ψ , was set equal to 7 character spaces, based on estimates from McConkie et al., 1988). The values of a second set of six parameters (ε , Δ , $\sigma\gamma$, ξ , η_1 , η_2) were selected so that the model would generate eye-movement behavior having certain properties; for example, the values of the parameters that control how random error increases with saccade length (η_1 and η_2) were selected so that the fixation landing-site distributions resembled those reported by McConkie et al. (1988). Finally, the values of the remaining five parameters (α_1 , α_2 , α_3 , λ , R) were selected by completing multiple parameter searches within a space that was defined by all possible orthogonal combinations of the parameters. The upper and lower limits (respectively) of the parameter space during the first search were: $\alpha_1 = \{90, 150\}$; $\alpha_2 = \{0, 10\}$; $\alpha_3 = \{0, 50\}$; $\lambda = \{0, 1\}$; $R = \{50, 150\}$. These limits were set by logical constraints (e.g., word identification has to complete within some reasonable amount of time).

The first parameter search was completed using 10 statistical subjects per parameter set. During the search, new sets of parameter values were selected by systematically incrementing the value of each parameter orthogonally to the others (e.g., α_1 was incremented by 2 ms while holding the other parameter values constant). Each set of parameter values was then evaluated by computing the root-mean-squared-deviation (RMSD) of the normalized difference scores (for five frequency classes of words) between the mean observed and predicted first-fixation durations, gaze durations, probabilities of fixating once, and the probabilities of skipping (as described in Reichle et al., 1998). Additional parameter searches were then completed within a restricted part of the parameter space (i.e., within a part of the parameter space that was centered around the best-fitting parameter values obtained in the first search) using more statistical subjects (100 for the second grid search) and smaller parameter increments (e.g., α_1 was incremented by 1 ms in the second grid search). Three parameter searches were completed to find the parameter values that provided the best fit of the Schilling et al. (1998) data; these parameter values are listed in Table 1, and the simulation results are shown in Table 2.

It is important to demonstrate that E-Z Reader 9 accounts for the same phenomena that have been used to evaluate its predecessors. Several simulations of these benchmark phenomena (e.g., the word-based measures obtained from the Schilling et al., 1998, sentence corpus) that have been used to evaluate both the current and previous versions of our model are reported here (as in Table 2). The remaining benchmark simulations (Simulations 1–5), along with some new simulations that have not been previously reported (Simulation 6–9), are also reported here. Each simulation is based on 1000 statistical subjects and the same parameter values that were used in the other simulations reported in this paper.

A.1. Simulation 1: Fixation-duration distributions

Fig. 6 shows the observed and predicted distributions of first-fixation and gaze durations for five frequency classes of words. E-Z Reader 9 is clearly able to generate

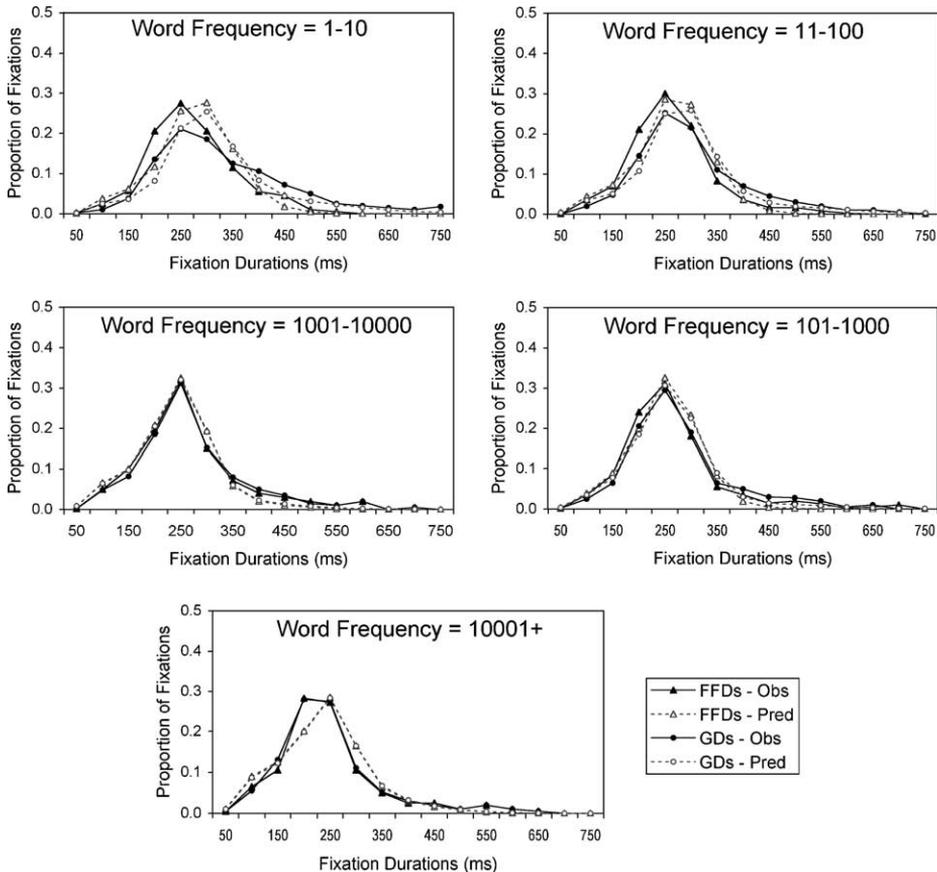


Fig. 6. Observed (Schilling et al., 1998) and predicted (E-Z Reader 9) first-fixation (FFD) and gaze duration (GD) distributions for five frequency classes of words.

distributions that closely resemble the observed distributions, thereby showing that the model can handle both aggregate (see Table 2) and distributional aspects of the Schilling et al., sentences.

A.2. Simulation 2: Word-frequency effects

The model was used to generate predictions for the specific target words that were used by Schilling et al., to examine the effects of word frequency. In their study, the first-fixation and gaze durations were 31 and 50 ms longer (respectively) on low-frequency words (mean = 2 per million) than on high-frequency words (mean = 141 per million). E-Z Reader 9 predicted 22 and 40 ms frequency effects on the first-fixation and gaze durations, respectively. These results demonstrate that E-Z Reader 9 can generate reasonable predictions for specific target words.

A.3. Simulation 3: Spillover effects

In this next simulation, we substituted the mean frequency of first the low-frequency targets (2 per million) and then the high-frequency targets (141 per million) into the word positions (designated as $word_n$ in the simulation) that were occupied by the Schilling et al. (1998) target words. The sentences were thus used as “frames” to determine whether the resulting frequency effect on $word_n$ would spill over onto the subsequent word ($word_{n+1}$). In this simulation, the length (=6 letters) and predictability (cloze probability = 0) of $word_n$ and $word_{n+1}$ were controlled, as was the frequency, predictability, and length of $word_{n+1}$. Consistent with reports that the spillover effect is typically a lot smaller than the actual frequency effect (Rayner & Duffy, 1986; Rayner et al., 1989), E-Z Reader 9 predicted a 26-ms frequency effect on the $word_n$ gaze durations, and a 7-ms spillover effect on the $word_{n+1}$ gaze durations.

A.4. Simulation 4: Skipping costs

Fixation durations on $word_n$ are typically longer if $word_{n-1}$ is skipped than if $word_{n-1}$ is fixated (Pollatsek, Rayner, & Balota, 1986; cf. Kliegl & Engbert, 2005). E-Z Reader predicts that this should happen because, in cases where $word_{n-1}$ is skipped, the parafoveal processing that is completed on $word_n$ is done from a more distant viewing position (where the quality of the visual information is less good because of poorer retinal acuity). E-Z Reader 9 predicted a 29-ms fixation duration “cost”, which is roughly consistent with the 50-ms effect observed in the Schilling et al. corpus.

Earlier versions of E-Z Reader also made a unique prediction (that was subsequently confirmed; Pollatsek et al., 1986; Pynte, Kennedy, & Ducrot, 2004; Rayner et al., 2004; Reichle et al., 1998) that fixation durations on $word_n$ should be longer if $word_{n+1}$ is skipped than if $word_{n+1}$ is fixated. This prediction stems from the fact that, in the model, the saccadic program that causes $word_{n+1}$ to be skipped cancels an earlier program and thus requires some additional time to be re-started. E-Z Reader 9 predicted a 18-ms fixation duration cost, which is again reasonably consistent with the 38-ms effect observed in the Schilling et al., corpus.

A.5. Simulation 5: Landing-site distributions

Fig. 7A shows the landing-site distribution on 6-letter words that were generated by E-Z Reader 9. (The model generates very similar distributions for 4-, 5-, and 7-letter words.) Note that the distributions are (a) centered on the words, (b) approximately normal in shape (the truncated ends reflect saccades that were predicted to either undershoot or overshoot the target word), and (c) shift towards the beginnings of the words and become more variable (flatter) as saccade length increases. Fig. 7B

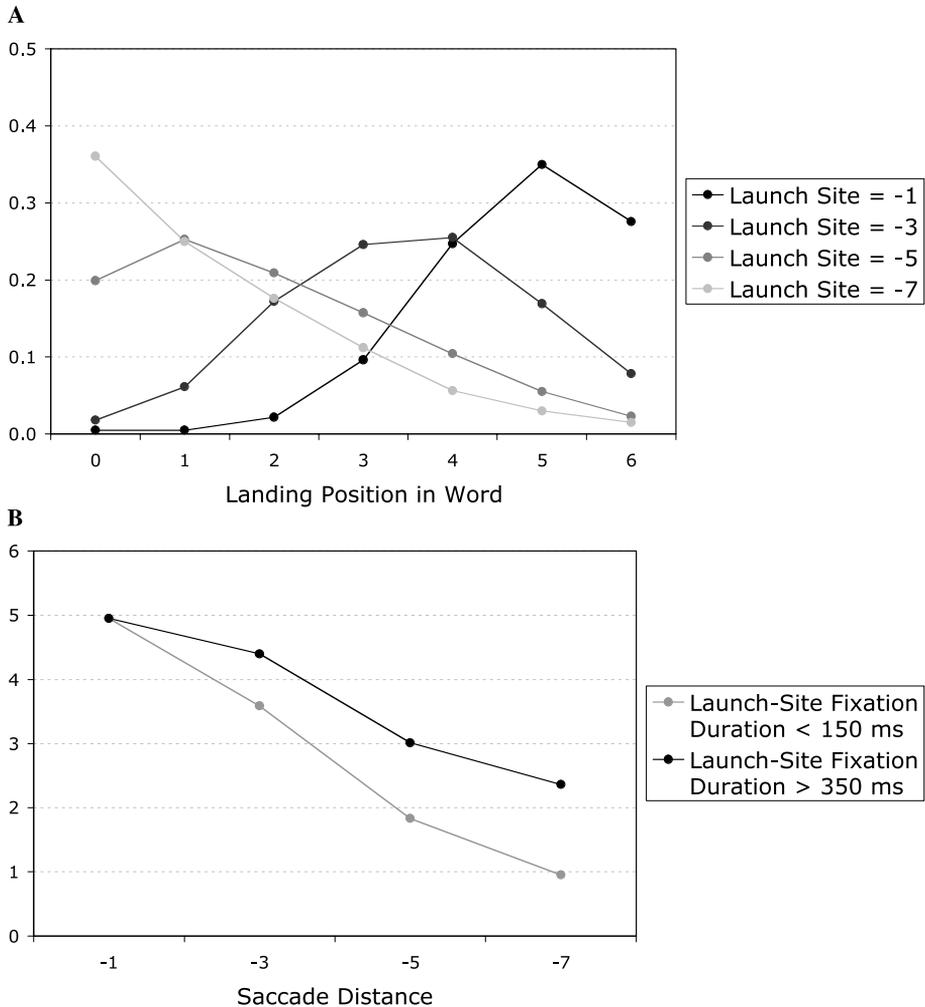


Fig. 7. (A) The landing-site distributions on 6-letter words that are predicted by E-Z Reader 9. (B) The means of this predicted landing-site distributions become more variable (i.e., the spread among the means increases) following shorter launch-site fixation durations (<150 ms) than following longer fixations (>350 ms).

also shows that the means of the landing-site distributions become less variable as the fixation duration on the launch-site word increases. All of these predictions are consistent with results in the literature (McConkie et al., 1988, 1991; Rayner et al., 1996).

A.6. Simulation 6: Word-predictability effects

E-Z Reader 9 predicts that words that are more predictable from their preceding sentence context should be the recipients of both shorter fixations and fewer fixations than words that are less predictable. Such predictability effects have been reported in the literature (Altarriba et al., 1996; Ehrlich & Rayner, 1981; Inhoff, 1984; Rayner & Well, 1996). Fig. 8A shows the effects of word predictability generated by E-Z Reader 9 when both word frequency (=100 per million) and length (=6 letters) are controlled.

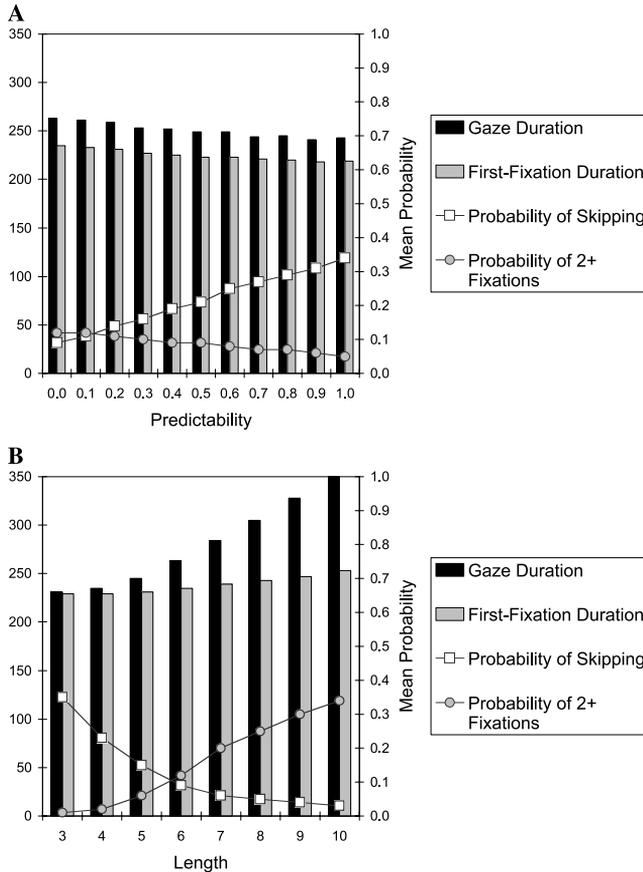


Fig. 8. (A) The effect of word predictability generated by E-Z Reader 9, controlling for both word frequency and length. (B) Predicted the word-length effects, controlling for frequency and predictability.

A.7. Simulation 7: Word-length effects

E-Z Reader 9 also predicts that longer words should be fixated more often and for more time than shorter words. Again, this is consistent with the literature (for a review, see Brysbaert & Vitu, 1998). Fig. 8B shows the predicted effects of word length when both word frequency (= 100 per million) and predictability (cloze probability = 0) are controlled. This simulation demonstrates that—counter to the claims of Brysbaert and Drieghe (2003)—our model predicts robust effects of word length under conditions where this variable is not negatively correlated with frequency (as it is in natural text).

A.8. Simulation 8: Foveal-load \times preview interaction

The processing difficulty of the fixated word has been shown to reduce the amount of processing that is completed on the parafoveal word (Drieghe et al., 2005; Henderson & Ferreira, 1990; Inhoff et al., 1989; Kennison & Clifton, 1995; White et al., 2005). E-Z Reader 9 predicts this interaction because of the disparity between the time that is needed to complete L_1 and L_2 and how this affects the amount of parafoveal processing that can be completed [see Eq. (3) and Fig. 1]. In the simulation, we inserted words that either had a frequency of 1 per million or 500 per million into the target word location in the Schilling et al. sentences and examined how this manipulation affected the parafoveal processing of $word_{n+1}$. In this simulation, the length (=6 letters) and predictability (cloze probability = 0) of $word_n$ and $word_{n+1}$ were controlled, as was the frequency (=1 per million) of $word_{n+1}$. E-Z Reader 9 predicted a 7-ms decrease in the size of the preview benefit (i.e., the difference between the predicted first-fixation durations on $word_{n+1}$ in the preview vs. no-preview conditions) when $word_n$ was difficult to process (i.e., low-frequency). This simulation demonstrates that the model predicts a foveal-load \times preview interaction, but that the absolute size of the decrement in preview benefit that is due to foveal load is modest.

A.9. Simulation 9: Initial landing-position effects

Our final simulation examined the effect that the initial landing position has on fixation durations and the probability of making a refixation. As in prior simulations, we used the Schilling et al. (1998) sentences as “frames” to examine the predicted effects of the initial landing position on $word_n$ when its frequency (= 1 per million) and predictability (cloze probability = 0) were held constant. Fig. 9A shows the probability of making a refixation on $word_n$ as a function of the initial fixation location on the word and its length (3-, 5-, and 7-letter words). (Other word lengths exhibited similar patterns but are not shown for brevity.) As has been reported in the literature, shorter words tended to be the recipients of fewer refixations than longer words, and refixations were more likely to occur following initial fixations near the beginnings and—to a lesser degree—ends of words (McConkie, Kerr, Reddix, Zola, & Jacobs, 1989; Rayner et al., 1996; Vitu et al., 2001). These results stem from two assumptions in E-Z Reader 9. The first is that the probability of initiating a refix-

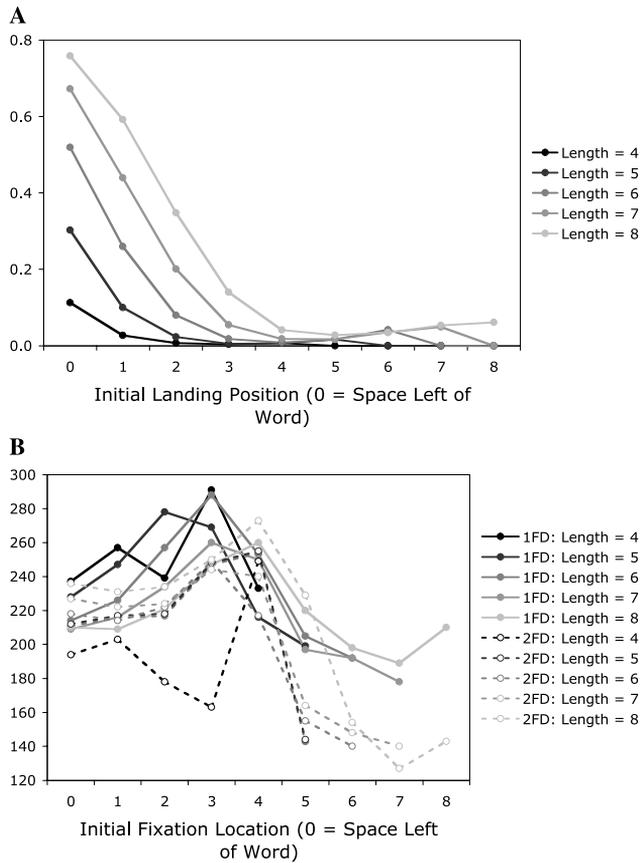


Fig. 9. (A) The probability of making a refixation on a word, as a function of the initial fixation location on the word and its length. (B) The mean first (1FD) and second (2FD) fixation durations on a word, as a function of the initial fixation location on the word and its length.

ation saccade is directly proportional to the distance between the initial fixation location on a word and its center (i.e., Eq. (6)). The second is that the completion of L_1 (which causes the oculomotor system to program a saccade to the next word, thereby canceling the refixation saccade) takes longer to complete from fixations located near the ends of words because of visual acuity limitations [i.e., Eq. (2)].

Because L_1 is most likely to cancel the refixation saccade following a first fixation located near the center of a word, refixation saccades starting from near the centers of words only occur in cases where L_1 takes a long time to complete (due to the variability associated with lexical processing). This bias tends to inflate the first of two fixations for initial fixations located near the centers of words, causing the inverted U-shaped fixation-duration function shown in Fig. 9B. E-Z Reader 9 thus predicts an inverted-optimal viewing position effect (Vitu et al., 2001). Unfortunately, the model does not provide a complete account of the initial landing-position effects; Fig. 9B also shows the durations of the second fixations, again conditional upon

the location of the first fixation. Although the duration of the second of two fixations is shorter than the duration of the first fixation (consistent with the findings of Vitu et al., 2001), the simulated second-fixation durations do not increase following initial fixations near the ends of words. This limitation of the model may reflect the fact that—in contrast to readers—E-Z Reader 9 always directs its refixation saccades towards the optimal viewing position. Future work is needed to determine if different metrics for refixation saccades will improve this aspect of the model's performance.

References

- Alpern, M. (1953). Metacontrast. *Journal of the Optical Society of America*, 43, 648–657.
- 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 & Cognition*, 24, 477–492.
- Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, 16, 285–311.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database. [CD-ROM]*. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.
- Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, 17, 364–390.
- Balota, D. A., & Rayner, K. (1983). Parafoveal visual information and semantic contextual constraints. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 726–738.
- Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research*, 19, 967–983.
- Brysaert, M., & Drieghe, D. (2003). Please stop using word frequency data that are likely to be word length effects in disguise. *Behavioral and Brain Sciences*, 26, 479.
- Brysaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 125–148). Oxford: Elsevier.
- Campbell, F. W., & Wurtz, R. H. (1978). Saccadic omission: Why we do not see a grey-out during a saccadic eye movement. *Vision Research*, 18, 1297–1303.
- Chace, K. H., Rayner, K., & Well, A. D. (2005). Eye movements and phonological parafoveal preview: Effects of reading skill. *Canadian Journal of Experimental Psychology*, in press.
- Clark, V. P., Fan, S., & Hillard, S. A. (1995). Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Human Brain Mapping*, 2, 170–187.
- Davis, C. J., & Bowers, J. S. (2004). What do letter migration errors reveal about letter position coding in visual word recognition? *Journal of Experimental Psychology: Human Perception & Performance*, 30, 923–941.
- Drieghe, D., Rayner, K., Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance* (in press).
- Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, 20, 641–655.
- Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, 42, 621–636.
- Engbert, R., Nuthmann, A., Richter, E., Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, in press.
- Eriksen, C. W., & Lappin, J. S. (1964). Luminance summation—contrast reduction as a basis for certain forward and backward masking effects. *Psychonomic Science*, 1, 313–314.
- Foxe, J. J., & Simpson, G. V. (2002). Flow of activation from V1 to frontal cortex in humans: A framework for defining “early” visual processing. *Experimental Brain Research*, 142, 139–150.

- Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.
- 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.
- Hanes, D. P., & Carpenter, R. H. S. (1999). Countermanding saccades in humans. *Vision Research*, *39*, 2777–2791.
- Henderson, J. M., Dixon, P., Petersen, A., Twilley, L. C., & Ferreira, F. (1995). Evidence for the use of phonological representations during transsaccadic word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 82–97.
- 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*, 417–429.
- Hyönä, J., & Bertram, R. (2004). Do frequency characteristics of nonfixated words influence the processing of fixated words during reading? *European Journal of Cognitive Psychology*, *16*, 104–127.
- Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement contingent display change study. *Memory & Cognition*, *32*, 523–532.
- 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*, 1430–1440.
- 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*, 1612–1627.
- 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*, 612–624.
- Inhoff, A. W., Eiter, B., & Radach, R. (2005). Temporal overlap in the extraction of linguistic information from consecutive words during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, in press.
- 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.
- Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, *40*, 431–439.
- Irwin, D. E. (1998). Lexical processing during saccadic eye movements. *Cognitive Psychology*, *36*, 1–27.
- 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 Science*, *6*, 1624–1632.
- Juhász, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology*, *94*, 223–244.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329–354.
- Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology*, *53A*, 429–456.
- Kennedy, A., Brooks, R., Flynn, L., & Prophet, C. (2003). The reader's spatial code. In J. Hyona, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research*. Amsterdam: North Holland.
- Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, *45*, 153–168.
- Kennison, S. M., & Clifton, C. (1995). Determinants of parafoveal preview 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.
- Kliegl, R., & Engbert, R. (2005). Fixation durations before word skipping in reading. *Psychonomic Bulletin & Review*, *12*, 132–138.

- 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*, 918–923.
- Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., & Kentridge, R. W. (2004). Eye movements when reading disappearing text: is there a gap effect in reading? *Vision Research*, *44*, 1013–1024.
- McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations in words. *Vision Research*, *28*, 1107–1118.
- 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 & Psychophysics*, *46*, 245–253.
- McConkie, G. W., Zola, D., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). Children's eye movements during reading. In J. F. Stein (Ed.), *Vision and visual dyslexia 13*. London: MacMillan.
- McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, *17*, 578–586.
- McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics*, *25*, 221–224.
- McDonald, S. A., Carpenter, R. H. S., & Shillcock, R. C. (2005). An anatomically-constrained, stochastic model of eye movement control in reading. *Psychological Review*, in press.
- McPeck, R. M., Skavenski, A. A., & Nakayama, K. (2000). Concurrent processing of saccades in visual search. *Vision Research*, *40*, 2499–2516.
- Meseguer, E., Carreiras, M., & Clifton, C. J. (2002). Overt reanalysis strategies and eye movements during reading on mild garden sentences. *Memory & Cognition*, *30*, 551–561.
- Mielliet, S., & Sparrow, L. (2004). Phonological codes are assembled before word fixation: Evidence from boundary paradigm in sentence reading. *Brain and Language*, *90*, 299–310.
- Molker, A., & Fisher, B. (1999). The recognition and correction of involuntary prosaccades in an antisaccade task. *Experimental Brain Research*, *125*, 511–516.
- 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*, 667–682.
- Mouchetant-Rostaing, Y., Giard, M.-H., Bentin, S., Aguera, P.-E., & Pernier, J. (2000). Neurophysiological correlates of face gender processing in humans. *European Journal of Neuroscience*, *12*, 303–310.
- Mozer, M. C. (1983). Letter migration in word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 531–546.
- Murray, W., & Rowan, M. (1998). Early, mandatory, pragmatic processing. *Journal of Psycholinguistic Research*, *27*, 1–22.
- O'Regan, J. K. (1979). Eye guidance in reading: Evidence for linguistic control hypothesis. *Perception & Psychophysics*, *25*, 501–509.
- O'Regan, J. K. (1990). Eye movements and reading. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 395–453). Amsterdam: Elsevier.
- O'Regan, J. K. (1992). Optimal viewing position in words and the strategy-tactics theory of eye movements in reading. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 333–354). New York: Springer-Verlag.
- O'Regan, J. K., & Lévy-Schoen, A. (1987). Eye movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 363–383). London: Erlbaum.
- Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: The word superiority effect. *Psychological Review*, *89*, 573–594.
- Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language*, *51*, 231–246.
- Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 767–779.
- Pollatsek, A., & Digman, L. H. (1977). Dependent spatial channels in visual processing. *Cognitive Psychology*, *9*, 326–352.
- Pollatsek, A., & Hyönä, J. (2004). The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes*.

- 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, 820–833.
- Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 148–162.
- Pollatsek, A., Perea, M., & Binder, K. S. (1999). The effects of “neighborhood size” in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1142–1158.
- Pollatsek, A., Rayner, K., & Balota, D. A. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception & Psychophysics*, 40, 123–130.
- Pollatsek, A., Rayner, K., & Collins, W. E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General*, 113, 426–442.
- Pollatsek, A., Reichle, E., & Rayner, K. (2003). Modeling eye movements in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eyes: Cognitive and applied aspects of eye movement research* (pp. 361–390). Amsterdam: Elsevier.
- Pynte, J., Kennedy, A., & Ducrot, S. (2004). The influence of parafoveal typographical errors on eye movements in reading. *European Journal of Cognitive Psychology*, 16, 178–202.
- Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, 88, 552–572.
- Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 7, 65–81.
- Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition*, 4, 443–448.
- Rayner, K. (1979). Eye guidance in reading: Fixation locations in words. *Perception*, 8, 21–30.
- Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, 41, 211–236.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.
- Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 720–732.
- Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science*, 206, 468–469.
- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14, 191–201.
- Rayner, K., Inhoff, A. W., Morrison, R., Slowiaczek, 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, 167–179.
- Rayner, K., & Juhasz, B. J. (2004). Eye movements in reading: Old questions and new directions. *European Journal of Cognitive Psychology*, 16, 340–352.
- Rayner, K., Liversedge, S. P., & White, S. J. (2005). Eye movements when reading disappearing text: The importance of the word to the right of fixation. *Vision Research*.
- Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading disappearing text: Cognitive control of eye movements. *Psychological Science*, 14, 385–389.
- Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements. *Vision Research*, 16, 829–837.
- Rayner, K., McConkie, G. W., & Ehrlich, S. F. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 529–544.
- Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology*, 12, 206–226.
- Rayner, K., & Morrison, R. M. (1981). Eye movements and identifying words in parafoveal vision. *Bulletin of the Psychonomic Society*, 17, 135–138.
- Rayner, K., Pollatsek, A., & Reichle, E. D. (2003). Eye movements in reading: Models and data. *Brain and Behavioral Sciences*, 26, 507–526.
- Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review*, 3, 238–244.

- Rayner, K., Reichle, E. D., & Pollatsek, A. (2005). Eye movement control in reading and the E-Z Reader model. In G. Underwood (Ed.), *Cognitive processes in eye guidance*. Oxford: Oxford University Press.
- Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, 4, 21–49 (Special issue).
- 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, 1188–1200.
- Rayner, K., Slowiaczek, M. L., Clifton, C., & Bertera, J. H. (1983). Latency of sequential eye movements: Implications for reading. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 912–922.
- Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). Eye movements and plausibility effects in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 1290–1301.
- Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, 3, 504–509.
- Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception & Psychophysics*, 31, 537–550.
- Rayner, K., White, S. J., Kambe, G., Miller, B., & Liversedge, S. P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J. Hyona, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213–234). Amsterdam: North Holland.
- Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 81, 275–280.
- Reichle, E. D. & Laurent, P. (2005). Using reinforcement learning to understand the emergence of “intelligent” eye-movement behavior during reading (manuscript submitted for review).
- Reichle, E., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, 105, 125–157.
- 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, 4403–4411.
- Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparison to other models. *Brain and Behavioral Sciences*, 26, 445–476.
- Reichle, E. D., Tokowicz, N., Liu, Y., & Perfetti, C. A. (2005). Using ERP to Examine Eye-Movement Control During Reading. Manuscript submitted for review.
- Reilly, R., & O'Regan, J. K. (1998). Eye movement control in reading: A simulation of some word-targeting strategies. *Vision Research*, 38, 303–317.
- Reilly, R., & Radach, R. (2003). Foundations of an interactive activation model of eye movement control in reading. In J. Hyona, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 429–455). Oxford: Elsevier.
- Reingold, E., & Rayner, K. (2005). Examining the word identification stages hypothesized by the E-Z Reader model. Submitted.
- Salvucci, D. D. (2001). An integrated model of eye movements and visual encoding. *Cognitive Systems Research*, 1, 201–220.
- Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, 26, 1270–1281.
- 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, 1021–1046.
- Sereno, S. C. (1992). Early lexical effects when fixating a word in reading. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 304–316). New York: Springer-Verlag.
- Sperling, G. (1960). The information available in brief presentations. *Psychological Monographs*, 74 (No. 498).
- Suppes, P. (1990). Eye-movement models for arithmetic and reading performance. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 455–477). Amsterdam: Elsevier.

- Suppes, P. (1994). Stochastic models of reading. In J. Ygge & G. Lennerstrand (Eds.), *Eye movements in reading* (pp. 349–364). Oxford: Pergamon.
- Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*, 107–141.
- Van Rullen, R., & Thorpe, S. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience*, *13*, 454–461.
- Vergilino, D., & Beauvillain, C. (2000). The planning of refixation saccades in reading. *Vision Research*, *40*, 3527–3538.
- 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*, 3513–3533.
- Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, *1*, 59–85.
- White, S. J., Rayner, K., & Livensedge, S. P. (2005). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A re-examination. *Psychonomic Bulletin & Review*, in press.
- Yang, S.-N., & McConkie, G. W. (2001). Eye movements during reading: A theory of saccade initiation time. *Vision Research*, *41*, 3567–3568.
- Yang, S., & McConkie, G. W. (2004). Saccade generation during reading: Are words necessary? *European Journal of Cognitive Psychology*, *16*, 226–261.