Visual Cognition, 2014
http://dx.doi.org/10.1080/13506285.2013.873508

Routledge
Taylor & Francis Group

Rethinking parafoveal processing in reading: Serial-attention models can explain semantic preview benefit and N+2 preview effects

Elizabeth R. Schotter[1], **Erik D. Reichle**[2], **and**
Keith Rayner[1]

[1]University of California, San Diego, CA, USA
[2]University of Southampton, Southampton, UK

(*Received 23 September 2013; accepted 4 December 2013*)

During reading, some information about the word to the right of fixation in the parafovea is typically acquired prior to that word being fixated. Although some degree parafoveal processing is uncontroversial, its precise nature and extent are unclear. For example, can it advance up to the level of semantic processing? Additionally, can it extend across more than two spatially adjacent words? Affirmative answers to either of these questions would seemingly be problematic for serial-attention models of eye-movement control in reading, which maintain that attention is allocated to only one word at a time (see Reichle, 2011). However, in this paper we report simulation results using one such model, E-Z Reader (Reichle, Pollatsek, Fisher, & Rayner, 1998), to examine the two preceding questions. These results suggest the existence of both semantic preview and N+2 preview effects, indicating that they are not incompatible with serial-attention models. We discuss the implications of these findings for models of eye-movement control in reading and provide a new theoretical framework for conceptualizing parafoveal processing during reading and its influence on eye movement behaviour.

Keywords: Semantic preview benefit; $N+2$ preview benefit; Models of eye movement control; Reading.

There is considerable evidence that, during reading, information about the word to the right of the fixated word is processed to some degree. This is clear from the fact that, approximately 30% of the time, readers do not directly fixate words

---

Please address all correspondence to Elizabeth Schotter, Department of Psychology, University of California, San Diego, La Jolla, CA 92093-0109, USA. E-mail: eschotter@ucsd.edu

The work described in this paper was supported by the Atkinson fund at UCSD and grant HD065829.

(i.e., skip them; Rayner, 1998, 2009a) during first pass reading and from the fact that parafoveal processing of the upcoming word leads to preview benefit (for a review, see Schotter, Angele, & Rayner, 2012). What remains less clear, however, is the precise nature of such parafoveal processing. For example, although orthographic information about the next word (word *N*+1) is usually acquired while fixating the current word (word *N*; Rayner, 1975), there is some uncertainty about whether semantic information is also acquired and integrated with foveal processing after a saccade (see Schotter, 2013). Similarly, the temporal and spatial constraints on parafoveal processing are not well understood, raising questions about whether, for example, information about word *N*+2 can also be acquired from word *N* (see Radach, Inhoff, Glover, & Vorstius, 2013).

The answers to these questions have important ramifications for models of eye movement control during reading and the debate about whether attention is allocated to one or multiple words during reading (Reichle, Liversedge, Pollatsek, & Rayner, 2009). For example, according to the SWIFT model (Engbert & Kliegl, 2011; Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005; Risse, Hohenstein, Kliegl, & Engbert, this issue 2014; Schad, & Engbert, 2012), attention is distributed as a gradient to support the concurrent lexical processing of multiple words (typically 3–4). By this account, both semantic preview of word *N*+1 and some amount of *N*+2 preview should occur because a significant amount of lexical processing of both words *N*+1 and *N*+2 normally occurs while the eyes are still on word *N*. However, according to the E-Z Reader model (Pollatsek, Reichle, & Rayner, 2006; Rayner, Ashby, Pollatsek, & Reichle, 2004; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2012; Reichle, Rayner, & Pollatsek, 2003; Reichle, Warren, & McConnell, 2009), attention is allocated in a serial manner, to support lexical processing of only one word at a time. For that reason, it is less intuitive how the model would predict semantic preview benefit because lexical (and therefore semantic) processing of word *N*+1 only begins after lexical processing of word *N* has completed. And similarly, it is less intuitive how E-Z Reader would predict *N*+2 preview because lexical processing of word *N*+2 can only begin after lexical processing of word *N*+1 has completed. For this reason, any evidence that semantic information is acquired from word *N*+1 or that parafoveal lexical processing extends to word *N*+2 is generally taken to be at odds with models that posit strictly serial lexical processing (see Reichle, 2011).

That said, this paper reports simulations using E-Z Reader to estimate the probability and magnitude of these effects, demonstrating that they are not necessarily incompatible with the model. By doing this, we will attempt to answer the two previously raised questions: (1) Is semantic information acquired from the parafovea? (2) How many words typically receive some amount of parafoveal lexical processing? We will then introduce a theoretical framework for understanding the functional role of parafoveal processing during reading and the important constraints it places on any plausible model of readers' eye movements.

THE DEPTH AND SPATIAL EXTENT OF PARAFOVEAL PROCESSING

As indicated earlier, one controversial aspect of reading is whether semantic information is accessed from parafoveal words and integrated across saccades to facilitate lexical processing (i.e., whether semantic preview benefit exists). Although several studies report positive evidence for semantic preview benefit when reading German (Hohenstein & Kliegl, in press; Hohenstein, Laubrock, & Kliegl, 2010) and Chinese (Yan, Richter, Shu, & Kliegl, 2009; Yang, 2013; Yang, Wang, Tong, & Rayner, 2010), all but one of the studies involving the reading of English have failed to observe semantic preview benefit (Rayner, Balota, & Pollatsek, 1986; Rayner & Schotter, 2013; Rayner, Schotter & Drieghe, 2013).[1]

The one exception to the latter set of null results is a recent study by Schotter (2013). In that study, using the boundary paradigm (Rayner, 1975), the preview could either be the target word itself (*begin*), a synonym of that word (*start*), a semantically related word (*ready*), or an unrelated word (*check*). Whenever the subjects' eyes crossed an invisible boundary to the left of a target word, the preview was immediately replaced by the target word. There were two key findings from this study: (1) Fixation duration on the target was approximately the same in the synonym preview condition as in the identical preview condition; and (2) fixation duration on the target was significantly shorter in these two conditions than in the unrelated preview condition, which was no different from the semantically related preview condition.

As Schotter (2013) argued, the finding that the preview effect size varied with the amount of semantic overlap provides some insight into why the effect was not observed in previous studies (e.g., those studies did not control the type of semantic relatedness). Perhaps more importantly, her data—combined with studies from other languages showing semantic preview benefit—demonstrate that such effects may be not just be limited to specific languages or writing systems, provided certain conditions are satisfied. That said, it is important to know whether such effects are consistent with serial-attention models of eye-movement control that, like E-Z Reader, posit that only one word is attended and lexically processed at any given point in time (Reichle, 2011).

---

[1] Rayner and Schotter (2013) reported fully significant semantic preview benefit effects in later measures (e.g., go-past time) for target words with the first letter capitalized (as in German), but since such measures reflect postlexical processing they may not indicate semantic preview benefit, per se. They also found a hint of semantic preview benefit in earlier measures, such as gaze duration- but only if the preview/target was capitalized.

In a similar manner, there has recently been a debate about whether parafoveal processing can extend not just from word *N* to word *N*+1, but whether (under some conditions) it can also extend to word *N*+2. Several studies found no effects of word *N*+2 preview (Angele & Rayner, 2011; Angele, Slattery, Yang, Kliegl, & Rayner, 2008; Rayner, Juhasz, & Brown, 2007). Moreover, McDonald (2006) only found word *N*+1 preview if that word was the target of the current saccade, along with no cumulative preview benefit across successive saccades on the pretarget word, further contradicting the possibility of preview from words other than *N*+1. However, Radach et al. (2013) reported *N*+2 preview effects when words *N* and *N*+2 were short and high frequency and word *N*+1 was the short, high-frequency word *the* (but see Angele & Rayner, 2011). And although Kliegl, Risse, and Laubrock (2007) did not find an *N*+2 preview benefit effect per se, they observed an effect on word *N*+1, which they interpreted as evidence of parafoveal preprocessing across multiple words (see also Risse & Kliegl, 2011). In studies using Chinese (in which words *N*+1 and *N*+2 are usually 1–2 characters and thus more likely to be close to the fovea), an *N*+2 preview benefit effect has been reported, but only if word *N*+1 was not masked, was high-frequency (Yan, Kliegl, Shu, Pan, & Zhou, 2010), or was a function word (Yang et al., 2009); *N*+2 preview was not observed if the preceding character was low frequency (Yan et al., 2010; Yang, Rayner, Li, & Wang, 2012).

Although the evidence for *N*+2 preview effects is less equivocal than that for semantic preview benefit, there is enough uncertainty to warrant a more thorough determination of whether such effects are compatible with serial-attention models of eye-movement control. Therefore, in the next section of this paper we attempt to examine both of these effects using the E-Z Reader model (Reichle, 2011). We first provide a brief overview of the model and then report the results of two simulations—one examining semantic preview benefit and the other examining *N*+2 preview effects.

THE E-Z READER MODEL AND SIMULATIONS OF SEMANTIC AND *N* + 2 PREVIEW BENEFIT

The E-Z Reader model

E-Z Reader is a computational model that simulates the eye movements observed during reading. Figure 1 is a schematic diagram showing the main components of the model, the flow of information, and how the control of processing is passed between components. Components can be grouped into five broad categories: (1) preattentive visual processing (V); (2) early (L_1) and late (L_2) stages of lexical processing; (3) attention (A); (4) postlexical integration of word meanings into the sentence representation (I); and (5) labile (M_1) and nonlabile (M_2) programming and execution of saccades (S).

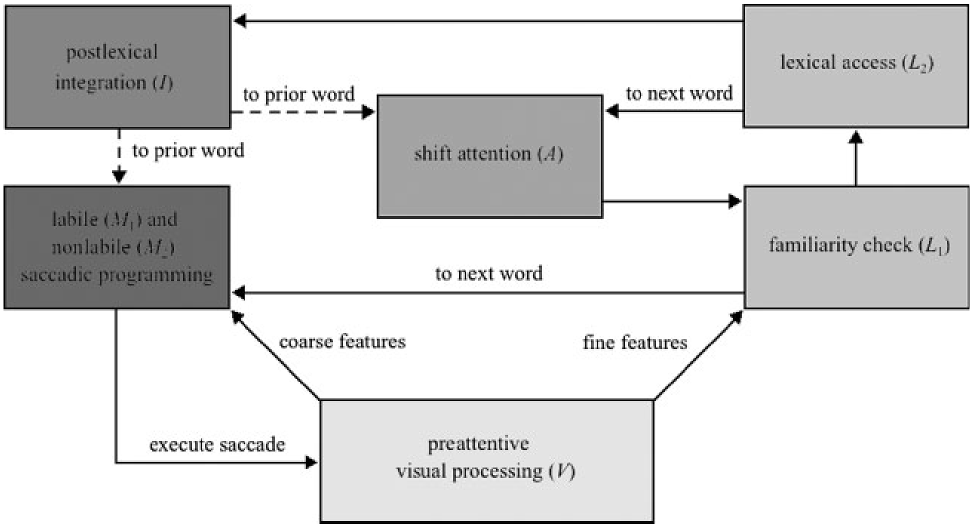

Figure 1. Schematic diagram of the E-Z Reader model of eye-movement control during reading. The boxes represent information processing components and the arrows represent the flow on information and the flow of control between these components. The shading shows how the components can be divided into five basic functional groups: (1) preattentive visual processing, (2) lexical processing, (3) attention, (4) postlexical processing, and (5) saccadic programming and execution.

The core assumptions of the model are: (1) completion of an early stage of lexical processing, called the familiarity check (i.e., L_1), is the signal to begin programming a saccade to move the eyes from one word to the next; (2) completion of lexical access (L_2) is the signal to shift attention; and (3) attention is allocated in a strictly serial manner. The model therefore stands in contrast to models that posit that attention is a gradient encompassing several words in parallel (e.g., SWIFT: Engbert & Kliegl, 2011; Engbert et al., 2002, 2005; Schad, & Engbert, 2012).

The time (in ms) required to complete L_1 for a given word is a function of its frequency (measured using corpora; e.g., Francis & Kucera, 1982) and within-sentence predictability (measured using cloze-task norms; Taylor, 1953), as specified by:

$$t(L_1) = \begin{cases} 0 & \text{with p} = \text{predictability} \\ \alpha_1 - \alpha_2\,\text{frequency} - \alpha_3\,\text{predictability} & \text{with } p = 1 - \text{predictability} \end{cases} \quad (1)$$

In Equation (1), α_1 (= 104), α_2 (= 3.5), and α_3 (= 39) are free parameters that were selected to maximize the model's capacity to simulate several dependent measures obtained from the Schilling, Rayner, and Chumbley (1998) sentence corpus. Important consequences of this are that the duration of L_1 can be set equal to 0 ms if a word is predictable (i.e., "guessed" from its preceding sentence context). However, in the majority of instances, the duration of L_1 equals some nonzero value that is inversely related to a word's frequency and predictability.

The time (in ms) required to complete L_2 for a given word is also a function of its frequency and predictability, equalling some fixed proportion of $t(L_1)$ as specified by:

$$t(L_2) = \Delta\, t(L_1) \tag{2}$$

It is important to note that Equations (1) and (2) give the mean values of $t(L_1)$ and $t(L_2)$ for words of a given frequency and predictability; however, during each Monte-Carlo run of the model, the actual values of $t(L_1)$ and $t(L_2)$ are sampled from gamma distributions having means equal to the values given by Equations (1) and (2) and a standard deviation equal to 0.22 of those means. Furthermore, the time required to complete L_1 is also a function of the mean absolute distance between each of the letters in the word being processed and the centre of vision (i.e., the fixation location), as specified by:

$$t(L_1) \leftarrow t(L_1)\varepsilon^{\sum_{i=1}^{n} |fixation - letter_i|/n} \tag{3}$$

In Equation (3), ε (= 1.15) is a free parameter that controls the extent to which visual acuity affects the rate of lexical processing, and the exponent is the mean absolute distance (in letter spaces) between the fixation location and the locations of each of the i letters in the word being processed, with n being the number of letters in the word. Together, the preceding assumptions allow the model to explain the findings that common (i.e., frequent), predictable, and/or short or proximal words are fixated less often and for shorter durations than less common, unpredictable, and/or long or distant words (for reviews, see Rayner, 1998, 2009a).

According to the model, saccades are programmed in two discrete stages—an initial labile stage (M_1) that is subject to cancelation if another saccadic program is subsequently initiated, followed by a nonlabile stage (M_2) that cannot be cancelled. The times needed to complete both stages are random deviates sampled from gamma distributions with means respectively equal to $t(M_1)$ = 125 ms and $t(M_2)$ = 25 ms and standard deviations equal to 0.22 of those means. Furthermore, M_1 can be divided into two substages that each subsume half of $t(M_1)$—an initial "readying" stage that prepares the oculomotor system to program a saccade, followed by a "conversion" stage that converts the spatial coordinates of a saccade target to the appropriate saccade length. Importantly, the cancellation of one labile saccadic programming by the initiation of another can sometimes result in a "savings" or reduction in the amount of time required to complete the second program. This occurs because whatever time was spent "readying" the oculomotor system to program a saccade to one location will be subtracted from the time that would otherwise be spent "readying" the oculomotor system to program the second saccade.

Finally, the executed saccades are subject to both random and systematic error, causing the fixation-location distributions to be approximately Gaussian in shape but with missing "tails" due to occasional saccades that under/overshoot their intended targets. The execution of a saccade from one word to the next can also result in the initiation of a "corrective" saccade that is intended to rapidly move the eyes from a poor viewing location (i.e., near either end of a word) to a better viewing location, near the centre of a word. These corrective saccades are initiated probabilistically, as an increasing function of the distance between the intended saccade target (i.e., the centre of the word) and the initial fixation location.

The aforementioned assumptions (and others that are not discussed here because they are less related to the topics at hand; e.g., assumptions about postlexical processing) are sufficient for E-Z Reader to explain the "benchmark" phenomena that have been used to evaluate models of eye-movement control in reading (see Reichle, 2011), including the finding that words typically receive some amount of parafoveal processing, with the amount modulated by the processing difficulty of the fixated word (i.e., the Foveal load × Preview interaction; Henderson & Ferreira, 1990; Kennison & Clifton, 1995; White, Rayner, & Liversedge, 2005). Because these findings are central to this paper, it is worthwhile considering their explanation. Figure 2 shows the sequence of events that—according to E-Z Reader—results in word *N*+1 preview from word *N*.

The sequence shown in Figure 2 begins with the first stage of lexical processing (i.e., L_1) on word *N*, which is the word being fixated. At some point, L_1 completes on word *N*, which initiates the labile stage of saccadic programming (i.e., M_1) to move the eyes to word *N*+1. The second stage of lexical processing (i.e., L_2) of word *N* also continues. If, as depicted in Figure 2, L_2 completes before the eyes actually move to word *N*+1, then attention (*A*) will shift to word *N*+1, allowing some amount of preview of word *N*+1 from word *N*. The precise amount will of course depend upon how rapidly L_2 finishes on word *N*, but also on how long it takes to complete the labile (M_1) and nonlabile (M_2) stages of saccadic programming that are necessary to execute a saccade (*S*) to word *N*+1. Although visual processing is suppressed during saccades, lexical processing continues using information acquired from the previous fixation. Furthermore, as Figure 2 also shows, parafoveal processing does not end with the onset of a fixation of word *N*+1, but instead continues up until new visual information from the fixation on word *N*+1 actually reaches the brain (*V*), which then allows lexical processing of word *N*+1 to continue more rapidly because visual acuity is enhanced by the more proximal viewing location. The amount of time available for parafoveal processing of word *N*+1 thus corresponds to the interval between when attention first shifts to word *N*+1 and when new visual input from the fixation location on word *N*+1 first reaches the brain.

Given this background, it is now possible to explain how E-Z Reader accounts for the interaction between foveal load and preview. Because the time required to complete lexical access, $t(L_2)$, is a fixed proportion of the time required to complete the familiarity check, $t(L_1)$, and because the time required to

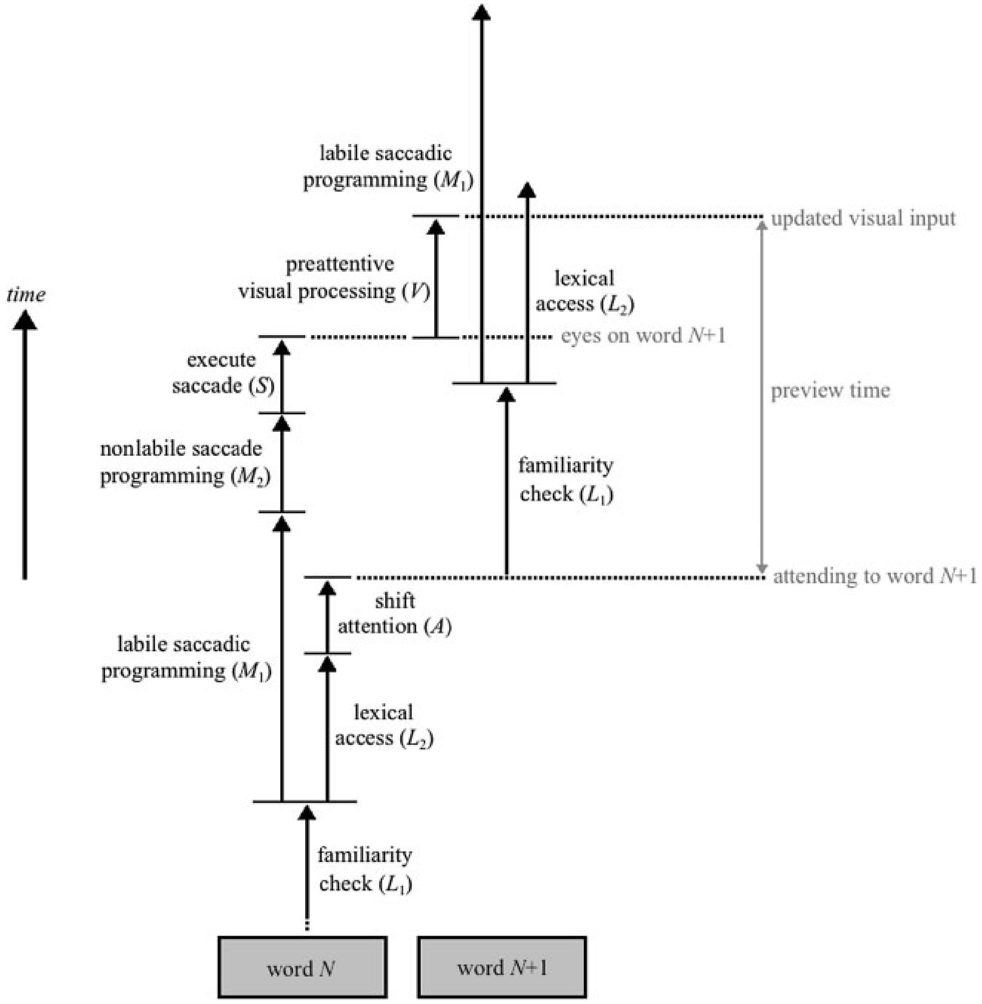

Figure 2. Schematic diagram showing the time course of processing in the E-Z Reader that can produce semantic preview benefit. In the diagram, the grey boxes represent two spatially adjacent words and time runs along the vertical axis, with processes (represented by labelled arrows) starting near the bottom of the diagram and completing near the top. The diagram shows a hypothetical sequence of events that can give rise to parafoveal processing of word N+1 from a fixation on word N. The preview time is denoted by the grey double-headed arrow, which shows the amount of time between when attention shifts to word N+1 (so that lexical processing of that word can begin) and when visual information from the new fixation location on word N+1 first reaches the brain. (See the text for a complete exposition of how this sequence may result in semantic preview benefit.)

program a saccade is on average a constant, the time available for previewing word N+1 will vary as a function of the processing difficulty of the fixated word, word N. This relationship is shown in Figure 3, where the mean time to complete L_1 and L_2 varies as a function of the processing difficulty of word N (i.e., its frequency, predictability, and length), thereby modulating the time available for preview of word N+1. Thus, the amount of time available for preview decreases with increasing foveal load. For that reason, one might reasonably predict that

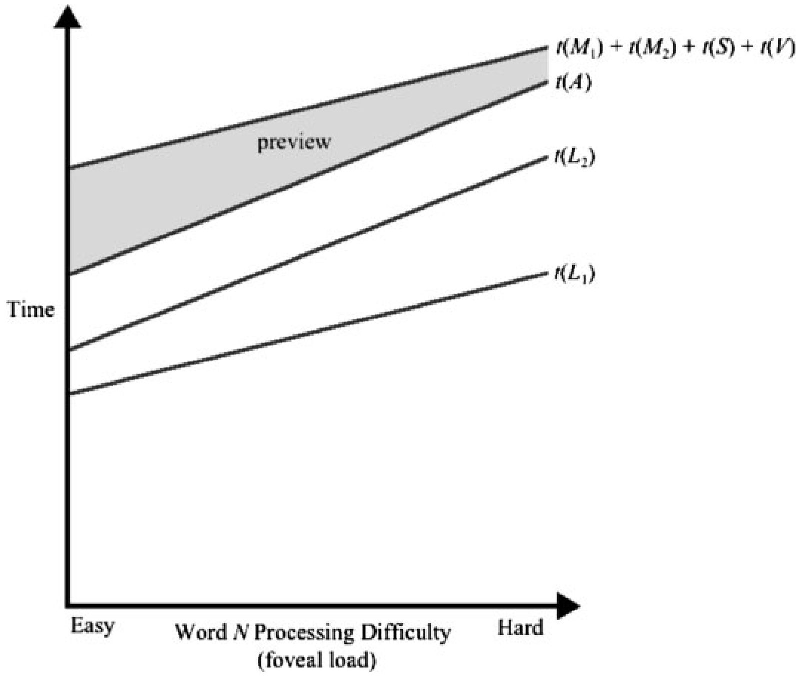

Figure 3. Schematic diagram showing how the time available for parafoveal processing of word *N*+1 varies as a function of the lexical processing difficulty of word *N* (i.e., foveal load). In the diagram, $t(L_1)$ and $t(L_2)$ corresponds to the time required to complete the two stages of lexical processing of word *N*, $t(A)$ corresponds to the time required to shift attention to word *N*+1, and $t(M_1) + t(M_2) + t(S) + t(V)$ corresponds to the time required to both move the eyes to word *N*+1 and for visual information from that new viewing location to be propagated from the eyes to the brain.

both semantic preview benefit and *N*+2 preview effects should be modulated by the lexical properties of word *N*, thus motivating the simulations reported later.

Simulations using E-Z Reader

The model was used to simulate a pair of "virtual experiments" to examine two phenomena of interest: (1) semantic preview benefit and (2) word *N*+2 preview effects. These simulations used the standard version of the model with all of its default parameter values (see Reichle et al., 2012), and were completed using the 48 sentences of the Schilling et al. (1998) corpus as "frames" within which the lexical properties of specific words of interest were manipulated to examine the consequences of those manipulations. Finally, both simulations were completed using 1000 statistical subjects per condition to provide reliable estimates of the simulated dependent measures that will be reported. The remaining details of our method are specific to each simulation and are therefore reported later. However, before doing this, it is important to discuss one final caveat about the simulations.

As described earlier, eye movement studies of parafoveal processing during reading have often used the boundary paradigm (Rayner, 1975) and considering how to approximate the boundary paradigm and other experimental manipulations in cognitive models is not trivial (see also Risse et al., this issue 2014). Although it is possible to simulate this paradigm by simply assuming that lexical processing does not begin until the eyes fixate the target word (e.g., see Pollatsek et al., 2006), this approach is not without its own problems. First and foremost, this assumption is at best an oversimplification because it is unlikely that lexical processing simply halts when—unbeknownst to the reader—attention shifts from the pretarget word to some preview other than the target word. Some available evidence (Murray, Rayner, & Wakeford, 2013) instead suggests that previews containing multiple letter changes produce some amount of inhibition of subsequent target processing, but that there may also be less inhibition or perhaps facilitation if the preview is similar to the target (e.g., contains a transposed letter pair or one different letter).

The second potential problem has to due to with fixations that are mislocated due to saccadic error. Because lexical processing is the “engine” moving the eyes forward in E-Z Reader, a situation can arise when the eyes fall short of the target word because of saccadic error, which then prevents any subsequent movement of the eyes because lexical processing of the word that would otherwise move the eyes forward (i.e., the target word) is not possible. Although it is possible to either remove such trials or to impose a “deadline” after which the eyes automatically move forward, neither solution is completely satisfactory because, for example, the consequences of removing such trials or including a deadline are not well understood.

Because of these potential problems, we decided that, rather than using one of the aforementioned “solutions” to approximate a simulation of the boundary paradigm, we instead simply used the model to directly predict the time course over which parafoveal processing of the preview word occurs, estimated by the probability of entering the lexical processing stages (L_1 and L_2) on that word and the time spent in these stages. This method avoids the potential pitfalls associated with not knowing how the nature of the preview and its relation to the target influences subsequent processing of that word.

Simulation 1: Semantic preview benefit

The first simulation examined the feasibility of semantic preview benefit within the serial-attention architecture of E-Z Reader, and how this benefit—if present—might be modulated by the lexical properties of the preceding word. We therefore orthogonally manipulated the length, frequency, and cloze predictability of word *N* (arbitrarily designated as the fifth word of each sentence) using three values of each factor: (1) The length was set equal to two, five, and eight letters; (2) the frequency was set equal to one, 100, and 10,000 occurrences per million; and (3) the cloze predictability was set equal to 0.0, 0.4, and 0.8. These values span the range of normal values while avoiding extreme cases (e.g., one-letter words or words that are completely predictable). Finally, because the amount of time available for parafoveal processing of a target word is independent of that word's processing difficulty in a serial-attention model like E-Z Reader, the lexical properties of word *N*+1 were selected to maximize their identification times (i.e., frequency = 1, predictability = 0, length = 8 letters), providing a conservative estimate of any possible semantic preview and minimizing potential problems associated with the target words being rapidly identified and therefore skipped.

The results of Simulation 1 are displayed in Table 1, systematically collapsing across two of the three factors to examine the consequences of manipulating the values of each individual factor. For example, the first three rows show how the length of the pretarget word, word *N*, affected the preview of word *N*+1 as measured by three indices: (1) the probability of engaging in any amount of preview of word *N*+1; (2) the mean duration (in ms) of the total preview time; and (3) the probability of the preview advancing into the L_2 stage of lexical processing—a stage that, as indicated, might be regarded as corresponding to *semantic* processing in the E-Z Reader model.

The overall probability of engaging in parafoveal processing of word *N*+1 from word *N* is quite high (*M* = 0.89). On average, this preview—when it occurred—lasted approximately 173 ms. Although this amount of preview might seem longer than one might predict, it includes both the duration of the saccade

TABLE 1

How the lexical properties of word *N* affect the mean probability of previewing word *N* + 1 from word *N*, the mean time available to preview word *N* + 1, and the probability of preview advancing to semantic processing of word *N* + 1 (standard deviations are in parentheses)

		Dependent measure		
Lexical property	*Value*	p*(preview)*	*Preview time (ms)*	p*(semantic preview)*
Length	2	.73 (.06)	163 (5)	.01 (.00)
	5	.95 (.01)	173 (5)	.02 (.00)
	8	1.00 (.01)	181 (4)	.02 (.00)
Frequency	1	.89 (.13)	169 (8)	.02 (.01)
	100	.89 (.13)	172 (8)	.02 (.01)
	10,000	.89 (.13)	177 (9)	.02 (.01)
Predictability	0.0	.87 (.15)	174 (6)	.02 (.01)
	0.4	.89 (.13)	173 (9)	.02 (.01)
	0.8	.92 (.10)	171 (11)	.02 (.01)

The standard deviations of all three measures (across conditions, not statistical subjects) are shown only to give some sense of the variability of those values, and not for inferential purposes.

from word N to N+1 (25 ms) and the time required for visual information from the new fixation to reach the brain (50 ms). Thus, if one subtracts these two process durations, the remaining preview time (M = 98 ms) is remarkably consistent with previous estimates derived from both survival-curve analyses of when word frequency effects occur in the presence versus absence of preview (Reingold, Reichle, Glaholt, & Sheridan, 2012) and electophysiological evidence about the time course of lexical processing and its relation to saccadic programming (Reichle & Reingold, 2013). Additionally, there was a very low probability of the preview advancing to the second, L_2 stage of lexical processing on word N+1 (M = 0.02). However, it is important to remember that the lexical properties of word N+1 were intentionally set equal to values that would make the word maximally difficult to identify; as such, our estimate of semantic preview benefit is conservative (as will be demonstrated later). Finally, because both the probability and duration of the preview were differentially affected by the length, frequency, and predictability of word N, each of these relationships will be discussed in turn.

As Table 1 shows, the probability of word N+1 preview was markedly affected by the length of word N (r = .90), as was the preview time (r = .86). Both of these positive relationships reflect the fact that longer words are more likely to be refixated than shorter words, with the probability of refixating word N being positively related to both the probability of previewing word N+1 (r = .83) and the preview time (r = .70). The underlying reasons for these latter two relationships are depicted in Figure 4, which is similar to Figure 2 but with the two panels showing the sequence of events that result in word N being the recipient of one or more than one fixation. Panel A shows the situation is which an initial (automatic refixation) saccadic program completes, causing word N to be refixated; in this situation, the time required to complete the second labile saccadic program (i.e., M_1) is not reduced, thus allowing some amount of time for previewing word N+1. In contrast, Panel B shows the situation in which a program to refixate word N is initiated but cancelled by the completion of L_1 and the initiation of a new saccadic program to move the eyes to word N+1 occurs. Because the second labile program cancels the first, it takes less time to complete the second because the oculomotor system has already been made ready to program a saccade, thereby reducing the time available for previewing word N+1.

Although the probability of word N+1 preview was unaffected by the frequency of word N (r = .01), there was a modest relationship between word N's frequency and the time spent previewing word N+1 (r = .37). The reason for the latter relationship was already explained in our earlier exposition (e.g., see Figure 3) of why E-Z Reader accounts for the interaction between foveal load and preview. However, the lack of a relationship between word N's frequency and the probability of previewing word N+1 indicates that the latter is largely a function

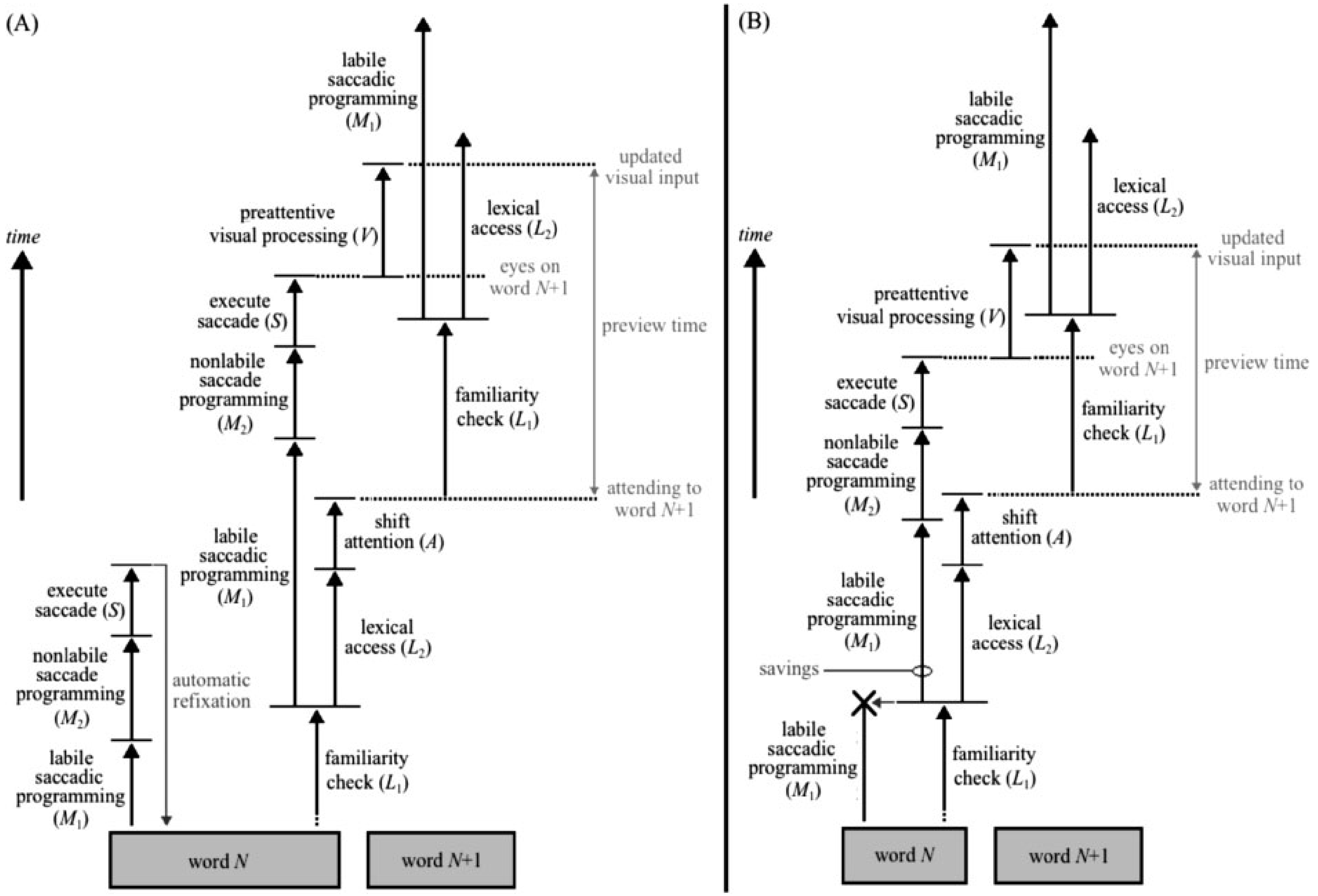

Figure 4. Schematic diagram showing the time course of processing in the E-Z Reader that results in parafoveal processing of word *N*+1 and how the probability and duration of this preview are modulated by the length of word *N*. In the diagram, the grey boxes represent two spatially adjacent words and time runs along the vertical axis, with processes (represented by labelled arrows) starting near the bottom of the diagram and completing near the top. Panel A depicts a situation in which word *N* is automatically refixated, thereby leaving both the time required to program a saccade to word *N*+1 and the time available for previewing word *N*+1 unaffected. Panel B depicts a situation in which a saccadic program to refixate word *N* is cancelled by the initiation of a saccadic program to move the eyes to word *N*+1, resulting in a reduction in the time required to program the second saccade (indicated by "savings") and thus a reduction in the time available for previewing word *N*+1.

of the probability of the pretarget word being fixated, as was discussed in relation to word *N*'s length and as will become evident by what will be discussed next.

As Table 1 also shows, the predictability of word *N* was actually positively related to the probability of previewing word *N*+1 ($r = .18$) but negatively related to the preview time ($r = -.13$). The reason for this has to do with the fact that, as the predictability of word *N* increases, so does its probability of being skipped ($r = .61$), causing the eyes to—in an increasing proportion of simulation trials—move directly from word *N*−1 to word *N*+1. This increasingly "rich" mixture of skipping trials reduces (but does not completely eliminate) the relationship between refixation probability and preview probability that was discussed earlier, in relation to word *N*'s length. Furthermore, the sequence of events that result in word *N* being skipped also causes a reduction in the amount of time available to preview word *N*+1. The reason for this reduction is depicted in Figure 5. In Panel A, the labile saccadic program to move the eyes from word *N*−1 to word *N* completes, resulting in both a fixation on word *N* and a certain amount of preview of word *N*+1. In Panel B, the first stage of lexical processing (i.e., L_1) is markedly reduced because word *N* is highly predictable (i.e., see the upper branch of Equation 1), resulting in the initiation of a saccadic program to move the eyes directly from word *N*−1 to word *N*+1. This often results in word *N* being skipped, but it also reduces the time available for previewing word *N*+1 because the cancelation of the first labile saccadic program by the second reduces the time necessary to complete the second program because of the savings gained from the oculomotor system already being ready to program a saccade.

The results of Simulation 1 are informative because they indicate that—within the framework of the E-Z Reader model—parafoveal processing of word *N*+1 from word *N* occurs more often than not, and that the resulting preview time was fairly substantial in duration. Although the finding that the preview time was usually insufficient for lexical processing of word *N*+1 to advance to the L_2 stage might be taken to preclude semantic preview benefit, this conclusion is premature because, as we indicated earlier, the lexical properties of word *N*+1 were deliberately selected to maximize its processing difficulty. That being the case, one might predict a significant amount of semantic preview benefit in cases where the preview word is less difficult to process, necessitating less time to reach an L_2 (i.e., semantic) stage of processing.

The latter point was demonstrated by running a variant of Simulation 1 using the mean lengths, frequencies, and predictabilities of the pretarget and synonym preview words that were used by Schotter (2013) to examine semantic preview benefit. Thus, the length, frequency, and predictability of the pretarget word *N* were set equal to five letters, 2019 occurrences per million, and .0, respectively. Similarly, the length, frequency, and predictability of the synonym preview word *N*+1 were respectively set equal to five letters, 527 occurrences per million, and .05. Using these values, the mean probability of previewing word *N*+1 was .94

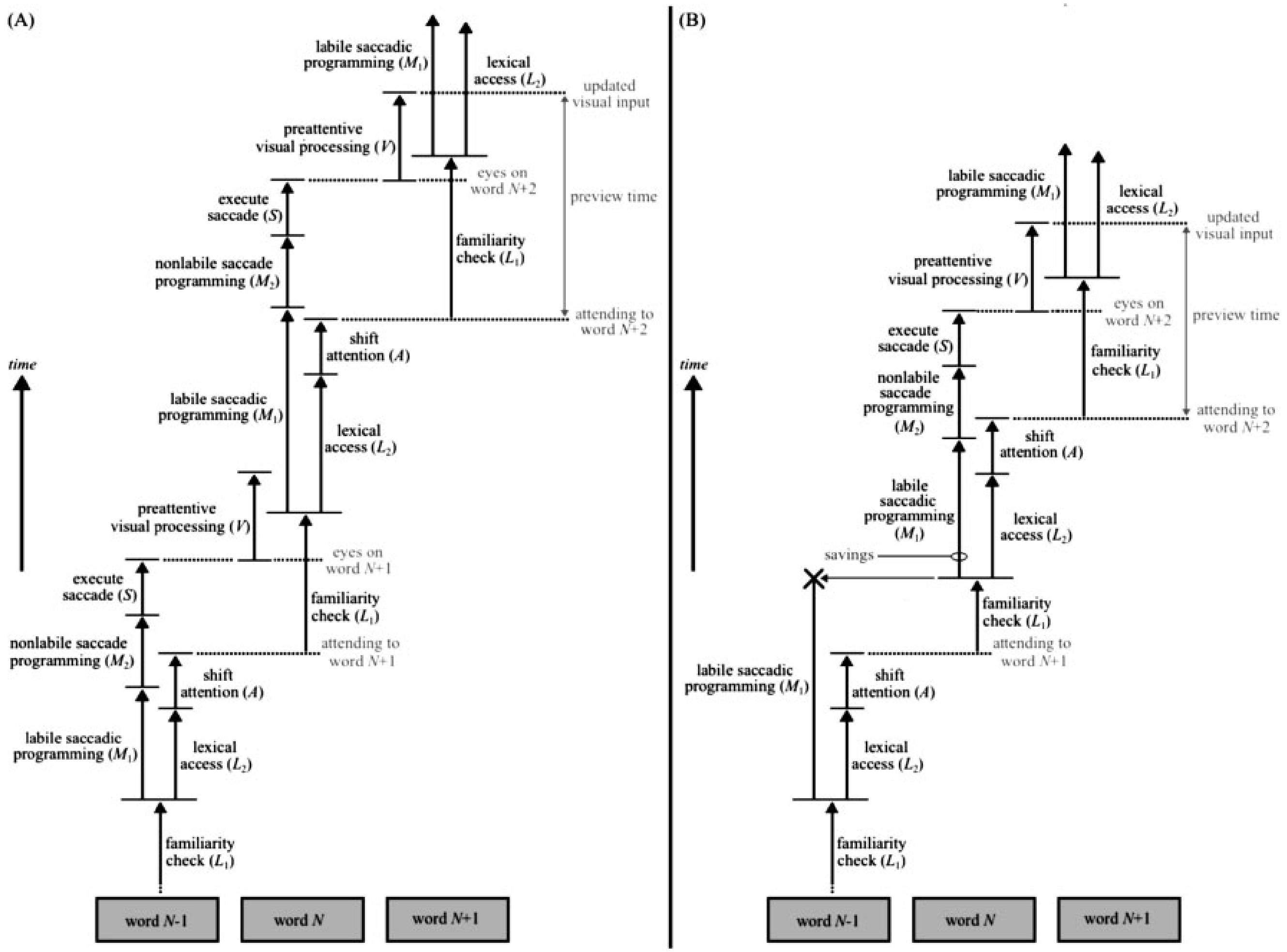

Figure 5. Schematic diagram showing the time course of processing in the E-Z Reader that results in parafoveal processing of word *N*+1 and how the probability and duration of this preview are modulated by the predictability of word *N*. In the diagram, the grey boxes represent three spatially adjacent words and time runs along the vertical axis, with processes (represented by labelled arrows) starting near the bottom of the diagram and completing near the top. Panel A depicts a situation in which word *N* is fixated, resulting in a certain amount of preview of word *N*+1. Panel B depicts a situation in which word *N* is predictable and thus more likely to be skipped, causing the eyes to move from word *N*–1 to word *N*+1 and resulting in a reduction of word *N*+1 preview because of the reduction in the time required to program the eye movement (indicated by "savings").

and the mean duration of that preview was 177 ms, and most importantly, the mean probability of the word N+1 preview reaching the L_2 stage of lexical processing was .08—often enough to possibly explain the semantic preview effects that have recently been reported by Schotter and others (e.g., Hohenstein & Kliegl, in press; Hohenstein et al., 2010; Yan et al., 2009; Yang, 2013; Yang et al., 2010). We will provide a theoretical "sketch" of these effects in the final section of this paper.

Simulation 2: *N* + 2 preview effects

The second simulation was similar to the first in both purpose and method, but examined the amount of parafoveal lexical processing of word N+2 from word N. To better understand how E-Z Reader might account for such an effect, consider the sequence of events that—according to the model—might result in word N+2 being attended from word N. This sequence is shown in Figure 6. As indicated, the eyes are on word N, which is rapidly identified, causing attention to shift to word N+1. If this word is processed rapidly enough, it too will be identified, causing attention to shift to word N+2 before the saccadic program to move the eyes to word N+1 can be completed, resulting in some amount of preview of word N+2 from word N. The presence of some amount of N+2 preview is thus compatible with the assumptions of the model; the main goals of Simulation 2, therefore, were to determine how often such N+2 preview effects actually occur, and the amount of time—if any—available for such preview.

As with Simulation 1, we were interested in knowing how preview of a target word (i.e., word N+2) would—if present—be modulated by the lexical properties of the pretarget words. We therefore manipulated the lengths, frequencies, and predictabilities of words N and N+1, using the same values of these factors as in Simulation 1: (1) lengths = two, five, and eight letters; (2) frequencies = 1, 100, and 10,000 occurrences per million; and (3) cloze predictabilities = 0.0, 0.4, and 0.8. As with Simulation 1, it was important to minimize the probability that the target words (i.e., word N+2) would be rapidly identified to avoid potential problems associated with it, for example, being skipped; for that reason, the length, frequency, and predictability of word N+2 were respectively set equal to eight letters, one occurrence per million, and 0.0. Finally, as with Simulation 1, three measures of parafoveal processing were calculated: (1) the probability of engaging in any amount of preview of word N+2; (2) the mean duration (in ms) of the total preview time; and (3) the overall probability of the preview advancing into the L_2 or semantic stage of lexical processing.

The results of Simulation 2 are shown in Table 2. Because of the large number of conditions (N = 729), the simulation results are shown in summary form, collapsing across five of the six factors to show how each of the lexical properties

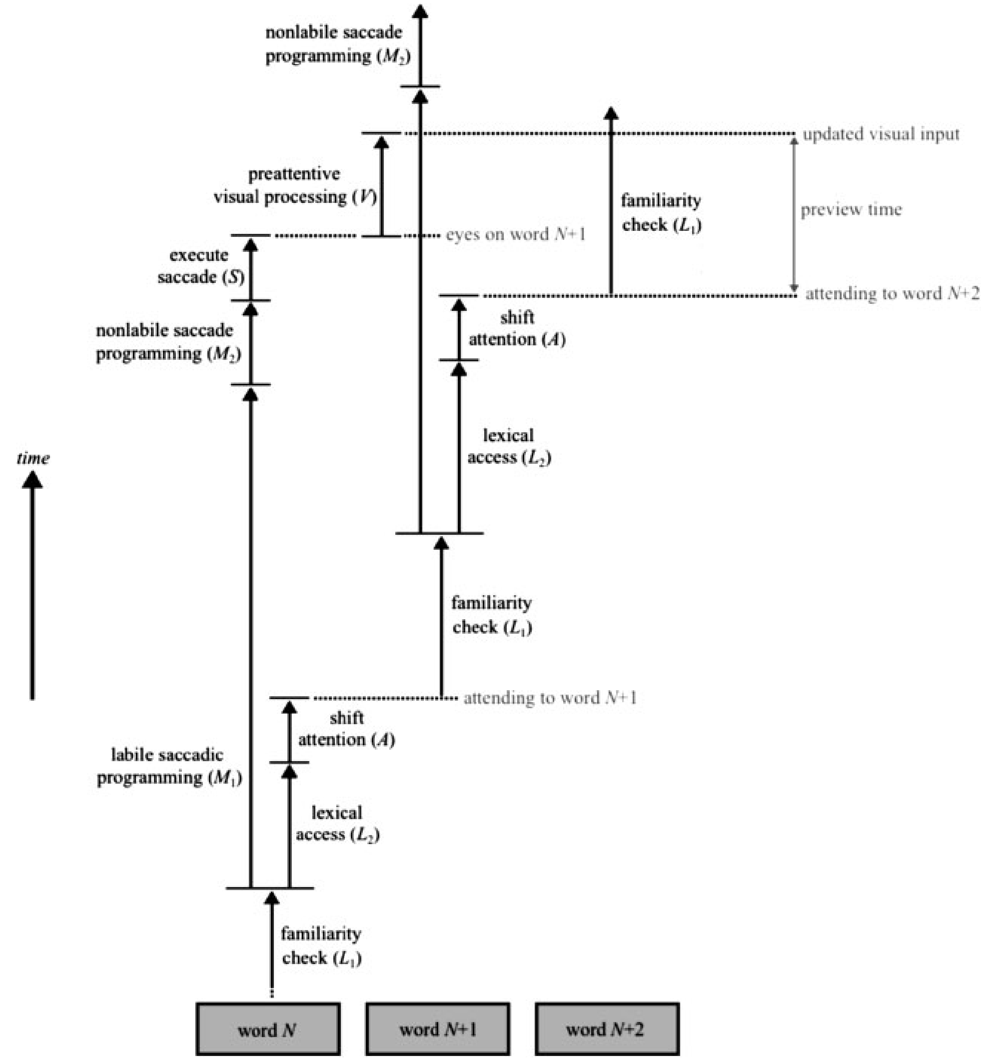

Figure 6. Schematic diagram showing the time course of processing in the E-Z Reader that can produce *N* +2 preview benefit. In the diagram, the grey boxes represent three spatially adjacent words and time runs along the vertical axis, with processes (represented by labelled arrows) starting near the bottom of the diagram and completing near the top. The diagram shows a hypothetical sequence of events that can give rise to parafoveal processing of word *N*+2 from a fixation on word *N*. The preview time is denoted by the grey double-headed arrow, which shows the amount of time between when attention shifts to word *N*+2 (so that lexical processing of that word can begin) and when visual information from the new fixation location on word *N*+1 first reaches the brain. (See the text for a complete exposition of how this sequence may result in *N*+2 preview effects.)

of words *N* and *N*+1 modulated preview of word *N*+2. For example, the top three rows show how the length of word *N* modulated preview of word *N*+2 when collapsing across the other lexical properties of words *N* and *N*+1.

The three key findings from Simulation 2 are straightforward. The first is that parafoveal preview of word *N*+2 did occur from word *N* during a modest but

TABLE 2

How the lexical properties of words *N* and *N* + 1 affect the mean probability of previewing word *N* + 2 from word *N*, the mean time available to preview word *N* + 2, and the probability of preview advancing to semantic processing of word *N* + 2 (standard deviations are in parentheses)

		Dependent measure		
Lexical property	*Value*	p*(preview)*	*Preview time (ms)*	p*(semantic preview)*
Word				
Length	2	*N*	137 (15)	.00 (.00)
	5	.21 (.14)	145 (20)	.00 (.00)
	8	.23 (.14)	149 (21)	.00 (.00)
Frequency	1	.19 (.13)	142 (21)	.00 (.00)
	100	.20 (.13)	143 (20)	.00 (.00)
	10,000	.20 (.14)	146 (18)	.00 (.00)
Predictability	0.0	.19 (.14)	142 (21)	.00 (.00)
	0.4	.20 (.13)	143 (19)	.00 (.00)
	0.8	.20 (.13)	145 (18)	.00 (.00)
N+1				
Length	2	.23 (.13)	143 (13)	.00 (.00)
	5	.19 (.13)	141 (18)	.00 (.00)
	8	.17 (.13)	146 (25)	.00 (.00)
Frequency	1	.19 (.13)	141 (20)	.00 (.00)
	100	.20 (.13)	145 (20)	.00 (.00)
	10,000	.21 (.14)	145 (18)	.00 (.00)
Predictability	0	.05 (.03)	168 (13)	.00 (.00)
	0.4	.19 (.04)	131 (7)	.00 (.00)
	0.8	.36 (.06)	132 (6)	.00 (.00)

The standard deviations of all three measures (across conditions, not statistical subjects) are shown only to give some sense of the variability of those values, and not for inferential purposes.

nontrivial proportion of trials ($M = 0.20$).[2] The second key finding is that, when there was preview, its duration was a nontrivial 144 ms. If one subtracts the saccade duration (25 ms) and eye-mind lag (50 ms) from the mean preview time, the remaining preview time is a fairly modest 69 ms. Together, these first two findings suggest that *N*+2 preview effects are possible, but that they are likely to be small in size. Finally, the modest preview time was never sufficient for lexical processing of word *N*+2 to advance to the L_2 stage of processing, which by implication suggests that—at least with words having lexical properties similar to

---

[2] Although our finding that the probability of observing *N*+2 preview in Simulation 2 ($M = 0.20$) was greater than the probability of observing semantic preview of word *N*+1 in Simulation 1 ($M = 0.02$) might appear to be in error, it is due to the fact that, in Simulation 1, the lexical properties of word *N*+1 were selected to make the word as difficult as possible to process, thereby minimizing the amount of observed semantic preview. In contrast, in Simulation 2, the lexical properties of word *N*+1 (which modulated the likelihood of previewing word *N*+2) were manipulated, often affording very rapid processing of word *N*+1 and thus substantial preview of word *N*+2.

those used in our simulation—parafoveal processing of word *N*+2 is not predicted to result in semantic preview benefit. Thus, our simulation shows that, although the E-Z Reader model predicts *N*+2 preview effects, these effects are likely to be modest in size and unlikely to be semantic in nature. This prediction is consistent with the *N*+2 preview effects reported by Radach et al. (2013, p. 628), who concluded that "readers sought graphemic information from *N*+2 during *N* viewing".

Finally, Table 2 indicates that both the probability of word *N*+2 preview and its duration are modulated by the lexical properties of both word *N* and *N*+1. As indicated, the preview probability was modestly related to word *N*'s length (r = .21) but was largely unaffected by either its frequency (r = .03) or predictability (r = .03). Similarly, the preview time was also modestly related to word *N*'s length (r = .24) but was largely unaffected by its frequency (r = .08) or predictability (r = .07). The influence of word *N*+1's lexical properties were more complex in that the preview probability was negatively related to word *N*+1's length (r = −.17), unaffected by word *N*+1's frequency (r = .04), and was positively related to word *N*+1's predictability (r = .94). Furthermore, preview time was largely unaffected by word *N*+1's length (r = .06) and frequency (r = .06), but was negatively related to its predictability (r = −.75). As with Simulation 1, these relationships between preview probability and time and the lexical properties of the preceding word(s) are extremely complex because they reflect both the duration of and interactions among those processes that—in Simulation 2—are completed in between the times when attention first shifts to word *N*+2 and when the eyes first move to the right of word *N*. The durations of these processes influence the time available for preview both directly (e.g., as shown in Figure 3) and indirectly, by influencing the proportion of trials containing word skipping versus refixations. These complexities underscore the importance of actually running simulations to evaluate one's intuitions about a model's "predictions" (Rayner, 2009b)—even a relatively simple model like E-Z Reader is highly interactive in nature and can generate highly complex behaviour. That being said, the final section of this paper will discuss our simulation results in relation to what has been learned about parafoveal processing and how it constrains models of eye movement control in reading. Our goal will be to develop a conceptual framework for thinking about parafoveal processing during reading.

GENERAL DISCUSSION

The simulations reported in this paper clearly demonstrate that the E-Z Reader model is not inconsistent with modest-sized semantic preview benefit and *N*+2 preview effects, thereby showing that these effects are not necessarily incompatible with the more general hypothesis that attention is allocated serially during reading. However, the effects may be incompatible with serial-attention models that posit a tight coupling between the signals to shift attention and move the eyes (e.g., EMMA; Salvucci, 2001) because, in such models, lexical processing may not occur rapidly enough to permit some amount of semantic parafoveal processing of the upcoming word and/or some lesser amount of parafoveal processing of word N+2. Conversely, the fact that such effects appear to be limited in scope (e.g., there does not appear to be significant semantic preprocessing of word N+2) may provide an important point of contrast with attention-gradient models like SWIFT (Engbert & Kliegl, 2011; Engbert et al., 2005; Schad & Engbert, 2012). For example, because these models posit a distributed gradient of attention that supports the concurrent lexical processing of 3–4 words, they may predict semantic preview benefit for word N+2. Thus, any firm conclusions about the compatibility of these models and semantic preview benefit or N+2 preview effects require validation via simulations (see Risse et al., this issue 2014). Indeed, the finding that E-Z Reader was consistent with both effects underscores the importance of running simulations rather than relying upon one's intuitions about the outcomes of hypothetical experiments (Rayner, 2009b).

Although both of the aforementioned parafoveal-processing effects are compatible with the E-Z Reader model, it is important to acknowledge that our simulations are limited in scope because the model is completely silent about many issues related to these effects. For example, Simulation 1 indicated that preview reached the L_2 (i.e., semantic) stage of lexical processing during only 8% of the trials in our simulation of Schotter's (2013) experiment. And furthermore, the model says nothing about how time spent engaged in parafoveal processing of upcoming words is converted into the patterns of fixation duration that are observed with different types of previews, such as the graded effects of semantic similarity reported by Schotter. Although this is on some level understandable because E-Z Reader does not provide a "deep" account of lexical processing, the fact that the model's assumptions are (relatively) transparent makes it an ideal framework for exploring possible explanations of parafoveal-processing effects. That will be our goal in the remainder of this paper.

Because of the model's assumption that decisions about when to move the eyes forward are based on early (i.e., partial) lexical information (i.e., whatever information is predicted from the prior sentence context and/or is available after completing L_1), fixation durations can be conceptualized as reflecting a reader's "confidence" or degree of certainty that the identity of an attended word *will be*—typically within the time required to program a saccade—available for further linguistic processing (e.g., see Reichle & Laurent, 2006). From that perspective, parafoveal processing can be understood in relation to three sources of information that are available about an attended word: (1) information generated from the preceding text that makes a word more or less predictable, (2) information obtained from the parafoveal preview itself, and (3) information obtained from the word after it has actually been fixated. Because these sources of information enter the word identification system at different times (typically in the order listed), they probably differentially influence decisions about when to move the eyes during first pass reading.

For example, predictions about an upcoming target word constrain the word that will be perceived (Bicknell & Levy, 2012) and consequently decisions about whether to fixate it and—if so—for how long. For example, many studies have demonstrated that predictable words are less likely to be fixated and/or the recipients of shorter fixations than less predictable words (Balota, Pollatsek & Rayner, 1985; Drieghe, Rayner & Pollatsek, 2005; Ehrlich & Rayner, 1981; Fitzsimmons & Drieghe, 2013; Kliegl, Grabner, Rolfs, & Engbert, 2004; Rayner et al., 2004; Rayner, Slattery, Drieghe, & Liversedge, 2011; Rayner & Well, 1996; Zola, 1984). Because the various linguistic constraints that make a word predictable are available before the word is actually fixated, these constraints may be sufficient to assume the word will sometimes be identifiable in the absence of visual information about that word (see Equation 1). This would explain why readers sometimes skip strings of *X*s (Rayner, Well, Pollatsek, & Bertera, 1982) and inappropriate words (Ehrlich & Rayner, 1981) if the sentences in which they are embedded are sufficiently predictive of what the "words" should be. It also partially explains why semantic preview benefit is found with synonyms of target words (Schotter, 2013): Because parafoveal processing of a synonym preview is sufficient to activate a contextually appropriate meaning, and because this meaning is completely congruent with the overall meaning of the sentence and any prediction about what the word should be, the introduction of the target word and its new orthographic and phonological forms may not even "register" with the reader (within the word identification system, at the level of conscious awareness, or perhaps both).

The second source of information about a word comes from the preview itself. Note that this information must be less precise than the information obtained upon fixation due to decreased visual acuity in the parafovea. In most instances, this information is probably used to initiate lexical processing, thus making the precise nature of the preview—and not just its similarity to the target word—a crucially important variable. For example, in the absence of strong constraint form the prior sentence, a preview that obviously does not correspond to an identifiable word (e.g., a string of *X*s or a random, unpronounceable sequence of letters) will probably cause the word identification system to stall out, delaying the initiation of saccadic programming and lengthening any subsequent fixation.[3] This will also be likely to make the decision about when to move the eyes more dependent upon (i.e., sensitive to) the lexical properties of the target word itself, after it has been fixated. If this conjecture is correct, then it would explain why nonword previews result in longer subsequent fixations on the target word than, for example, a completely unrelated but real word; whereas

---

[3] A nonword that is an orthographic neighbour of a real word (e.g., *sorp*) should be excluded from this consideration because it may be coerced by the word identification system into its word equivalent (e.g., *song*) due to the lower fidelity visual information entering the system during preview.

the latter type of preview can be processed to various levels of lexical representation, nonwords are unlikely to be processed in this manner. This account suggests that the 40–50 ms estimates of "preview benefit" obtained by comparing nonword previews with target-identical previews are inflated (e.g., see Hyönä, Bertram, & Pollatsek, 2004) because they probably include a significant "cost" due to the word identification system being prevented from accumulating any useful lexical information (Murray et al., 2013).

The last source of information about a word comes from the word itself, after it has been fixated. However, as already mentioned, there is a significant temporal lag (50–60 ms; see Reichle & Reingold, 2013) between when a word is initially fixated and when information obtained from that fixation impacts cognitive processing (see Figure 2). One consequence of this lag is that, on some occasions, the information that is used to decide when to move the eyes will be based on information obtained during the preview—even after the eyes have already moved to the target word. On such occasions, the initial fixation on the target word would be entirely influenced by the preview rather than the target itself because there would not have been sufficient time for information about the latter to reach the brain and those systems that mediate the decision about when to initiate saccadic programming. Such instances might also help explain the finding of semantic preview benefit for synonyms (Schotter, 2013): Because the preview word is on some proportion of trials processed to the level of its meaning, the evidence available to the word identification system is sufficient for the eyes to continue their normal progression through the text. This account might also explain the finding (Risse & Kliegl, in press) that unrelated previews that are higher frequency than the target (and that fit into the sentence) lead to shorter fixations on the target than when the preview is lower frequency than the target; this apparent "reverse frequency effect" indicates that decisions about when to move the eyes are sometimes based on properties of the preview more than the actual target word.

Although much remains to be learned about the nature of parafoveal processing during reading, we believe that the simulations and theoretical "sketch" reported in this paper represent a significant step towards providing a more concrete account of what may happen during such processing. However, we also acknowledge that efforts to understand parafoveal processing would also benefit from thinking about how existing word identification models (e.g., the Dual-Route Cascaded model—Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; or the Parallel Distributed Processing model—Seidenberg & McClelland, 1989) might accommodate semantic preview benefit and *N*+2 preview effects if they were embedded within the frameworks of models of eye movement control during reading. Such embedded models would, for example, be useful for examining precisely how the processing of synonyms across successive fixations might give rise to the semantic preview effects observed by Schotter (2013). Barring such models, however, we maintain that, although our simulations do not provide a comprehensive account of either semantic preview benefit and/or

N+2 preview effects, they do show that these effects are not necessarily inconsistent with the serial allocation of attention during reading.

References

Angele, B., & Rayner, K. (2011). Parafoveal processing of word n+ 2 during reading: Do the preceding words matter? *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1210–1220.

Angele, B., Slattery, T. J., Yang, J., Kliegl, R., & Rayner, K. (2008). Parafoveal processing in reading: Manipulating n + 1 and n + 2 previews simultaneously. *Visual Cognition*, *16*, 697–707.

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

Bicknell, K., & Levy, R. (2012). The utility of modelling word identification from visual input within models of eye movements in reading. *Visual Cognition*, *20*, 422–456.

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

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 954–969.

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., & Kliegl, R. (2011). Parallel graded attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *Oxford handbook on eye movements* (pp. 787–800). Oxford: Oxford University Press.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, *42*, 612–636.

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777–813.

Fitzsimmons, G., & Drieghe, D. (2013). How fast can predictability effects influence word skipping during reading? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1054–1063.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston, MA: Houghton Mifflin.

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.

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 166–190.

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1150–1170.

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Do frequency characteristics of nonfixated words influence the processing of fixated words during reading? *European Journal of Cognitive Psychology*, *16*, 104–127.

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., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*, 262–284.

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word N+2. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1250–1255.

McDonald, S. A. (2006). Parafoveal preview benefit in reading is only obtained from the saccade goal. *Vision Research*, *46*, 4416–4424.

Murray, W. S., Rayner, K., & Wakeford, L. (2013). Preview benefit or preview cost? Paper presented at the 17th European conference on Eye Movements, Lund, Sweden.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*, 1–56.

Radach, R., Inhoff, A. W., Glover, L., & Vorstius, C. (2013). Contextual constraint and N+2 preview effects in reading. *Quarterly Journal of Experimental Psychology*, *66*, 619–633.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*, 65–81.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

Rayner, K. (2009a). The Thirty Fifth Sir Frederick Bartlett Lecture: Eye movements and attention during reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.

Rayner, K. (2009b). Eye movements in reading: Models and data. *Journal of Eye Movement Research*, *2*, 1–10.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (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., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic processing during eye fixations in reading. *Canadian Journal of Psychology*, *40*, 473–483.

Rayner, K., Juhasz, B. J., & Brown, S. J. (2007). Do readers obtain preview benefit from word n+2? A test of serial attention shift versus distributed lexical processing models of eye movement control in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 230–245.

Rayner, K., & Schotter, E. R. (2013). Semantic preview benefit in reading English: The effect of initial letter capitalization. *Manuscript under review.*

Rayner, K., Schotter, E. R., & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit revisited. *Psychonomic Bulletin & Review*, in press.

Rayner K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 514–528.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin and Review*, *3*, 504–509.

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics*, *31*, 537–550.

Reichle, E. D. (2011). Serial attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *Oxford handbook on eye movements* (pp. 767–786). Oxford: Oxford University Press.

Reichle, E. D., & Laurent, P. A. (2006). Using reinforcement learning to understand the emergence of "intelligent" eye-movement behavior during reading. *Psychological Review*, *113*, 390–408.

Reichle, E. D., Liversedge, S. P., Pollatsek, A., & Rayner, K. (2009). Encoding multiple words simultaneously in reading is implausible. *Trends in Cognitive Sciences*, *13*, 115–119.

Reichle, E. D., 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., Pollatsek, A., & Rayner, K. (2012). Using E-Z Reader to simulate eye movements in non-reading tasks: A unified framework for understanding the eye-mind link. *Psychological Review*, *119*, 155–185.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, *26*, 445–476.

Reichle, E. D., & Reingold, E. M. (2013). Neurophysiological constraints on the eye-mind link. *Frontiers in Human Neuroscience*, *7*(361), doi: 10.3389/fnhum.2013.00361

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher-level language processing on eye movements during reading. *Psychonomic Bulletin and Review*, *16*, 1–21.

Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from survival analysis of fixation durations. *Cognitive Psychology*, *65*, 177–206.

Risse, S., Hohenstein, S., Kliegl, R., & Engbert, R. (2014). A theoretical analysis of the perceptual span based on SWIFT simulations of the n+2 boundary paradigm. *Visual Cognition*, doi: 10.1080/13506285.2014.881444

Risse, S., & Kliegl, R. (2011). Adult age differences in the perceptual span during reading. *Psychology and Aging*, *26*, 451–460.

Risse, S., & Kliegl, R. (in press). Dissociating preview validity and preview difficulty in parafoveal processing of word n+1 during reading. *Journal of Experimental Psychology: Human Perception and Performance*.

Salvucci, D. D. (2001). An integrated model of eye movements and visual encoding. *Cognitive Systems Research*, *1*, 201–220.

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition*, *20*, 391–421.

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

Schotter, E. R. (2013). Synonyms provide preview benefit in English but other semantic relationships do not. *Journal of Memory and Language*, *69*, 619–633.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, and Psychophysics*, *74*, 5–35.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 525–568.

Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, *30*, 415–433.

White, S. J., Rayner, K., & Liversedge, S. P. (2005). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A reexamination. *Psychonomic Bulletin and Review*, *12*, 891–896.

Yan, M., Kliegl, R., Shu, H., Pan, J., & Zhou, X. (2010). Parafoveal load of word N+1 modulates preprocessing effectiveness of word N+2 in Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 1669–1676.

Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Readers of Chinese extract semantic information from parafoveal words. *Psychonomic Bulletin and Review*, *16*, 561–566.

Yang, J. (2013). Preview effects of plausibility and character order in reading Chinese transposed words: Evidence from eye movements. *Journal of Research in Reading*, *36*, S18–S34.

Yang, J., Rayner, K., Li, N., & Wang, S. (2012). Is preview benefit from word n+ 2 a common effect in reading Chinese? Evidence from eye movements. *Reading and Writing*, *25*(5), 1079–1091.

Yang, J., Wang, S., Tong, X., & Rayner, K. (2010). Semantic and plausibility effects on preview benefit during eye fixations in Chinese reading. *Reading and Writing*, *25*, 1031–1052.

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