Journal of Experimental Psychology:
Human Perception and Performance

0096-1523/16/$12.00 http://dx.doi.org/10.1037/xhp0000253

Out of the Corner of My Eye: Foveal Semantic Load Modulates Parafoveal Processing in Reading

Brennan R. Payne, Mallory C. Stites, and Kara D. Federmeier
University of Illinois at Urbana-Champaign

In 2 experiments, we examined the impact of foveal semantic expectancy and congruity on parafoveal word processing during reading. Experiment 1 utilized an eye-tracking gaze-contingent display change paradigm, and Experiment 2 measured event-related brain potentials (ERPs) in a modified flanker rapid serial visual presentation (RSVP) paradigm. Eye-tracking and ERP data converged to reveal graded effects of foveal load on parafoveal processing. In Experiment 1, when word *n* was highly expected, and thus foveal load was low, there was a large parafoveal preview benefit to word $n + 1$. When word *n* was unexpected but still plausible, preview benefits to $n + 1$ were reduced in magnitude, and when word *n* was semantically incongruent, the preview benefit to $n + 1$ was unreliable in early pass measures. In Experiment 2, ERPs indicated that when word *n* was expected, and thus foveal load was low, readers successfully discriminated between valid and orthographically invalid previews during parafoveal perception. However, when word *n* was unexpected, parafoveal processing of $n + 1$ was reduced, and it was eliminated when word *n* was semantically incongruent. Taken together, these findings suggest that sentential context modulates the allocation of attention in the parafovea, such that covert allocation of attention to parafoveal processing is disrupted when foveal words are inconsistent with expectations based on various contextual constraints.

Keywords: reading, foveal load, context, eye movements, event-related brain potentials

Skilled readers not only process fixated words in foveal vision but also preprocess words in parafoveal vision, corresponding to approximately 2°to 5° on either side of the vertical meridian. In English, the perceptual span (i.e., the useful field of vision obtained in a single fixation) is asymmetric, ranging from three to four characters to the left of fixation to about 14 to 15 characters to the right of fixation (McConkie & Rayner, 1975). A substantial literature exists examining the nature and scope of parafoveal visual representations during reading, largely focusing on the quality of the information that can be extracted from the parafovea (see Schotter, Angele, & Rayner, 2012). One of the current outstanding questions in the literature concerns how attention is allocated across this perceptual span within a given fixation.

Brennan R. Payne, Mallory C. Stites, and Kara D. Federmeier, Department of Psychology, Neuroscience Program, and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign.

Mallory Stites is now at Sandia National Laboratories, Albuquerque, New Mexico.

This work was supported by a James S. McDonnell Foundation Scholar Award and a National Institutes of Health Grant AG026308 to Kara D. Federmeier. Portions of this research were presented at the Annual Meeting of the Cognitive Neuroscience Society. We thank Denis Drieghe and two anonymous reviewers for their comments on a previous version of this article.

Correspondence concerning this article should be addressed to Brennan R. Payne, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, 405 North Mathews Avenue, Urbana, IL 61801. E-mail: payne12@illinois.edu

The foveal load hypothesis (Henderson & Ferreira, 1990) posits that parafoveal processing is modulated by concurrent foveal difficulty, such that the amount of information derived from the word to the right of fixation is reduced as the difficulty induced by the currently fixated word increases. Although virtually every model of reading posits some online trade-off between foveal and parafoveal processing, demonstrations that foveal difficulty modulates parafoveal processing are largely confined to effects of foveal word frequency (Drieghe, Rayner, & Pollatsek, 2005; Henderson & Ferreira, 1990; Kennison & Clifton, 1995; Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999; White, Rayner, & Liversedge, 2005; see Schotter et al., 2012, for a review). In the present experiments, our goal was to further elucidate the role of foveal load on parafoveal word processing. Toward this goal, our first aim was to examine the effects of semantic constraints on parafoveal word processing by probing the effects of foveal expectancy (i.e., target word predictability) and congruity (i.e., semantic anomaly) on the parafoveal preview benefit in naturalistic reading. Our second aim was to examine how foveal semantic load constrains parafoveal processing by probing the nature and time course of the covert deployment of visual attention into the parafovea via event-related brain potentials (ERPs).

Foveal Load and Parafoveal Word Processing

A number of eye-tracking paradigms have been developed to study the dynamics of the perceptual span and parafoveal processing in natural reading (see Schotter et al., 2012, for a review). In the gaze-contingent boundary change paradigm (Rayner, 1975), an invisible boundary is placed between words *n* and $n + 1$. A change in the display is triggered when the eyes cross the invisible boundary, which is used to manipulate the availability of useful parafoveal information. On selected trials, $n + 1$ is initially replaced by a masking stimulus of some sort (i.e., the invalid preview condition) and, as the reader saccades from n to $n + 1$, the mask is replaced with the target word. The logic of this paradigm is that if a reader obtains parafoveal information from $n + 1$ while fixating on n, any inconsistency between what is available parafoveally and what is available when $n + 1$ is fixated results in a change in processing of $n + 1$. Research using this paradigm has demonstrated a characteristic *parafoveal preview benefit*: When a reader receives a valid preview of word $n + 1$, fixations on that word are 30 ms to 50 ms shorter relative to trials in which word n was initially replaced with an invalid mask, suggesting that some level of preprocessing of $n + 1$ occurred when the eyes were fixating word n (Hyönä, Bertram, & Pollatsek, 2004; Schotter et al., 2012).

Across these studies, effects of the parafoveal manipulation are often not observed until the masked stimulus is fixated. That is, observing effects of parafoveal manipulations of $n + 1$ during foveal processing of word n—so-called parafoveal-on-foveal (POF) effects— are rare and controversial in eye tracking (see Rayner, 2009, for a review). More recently, it has been argued that such POF effects exist, but are obscured by the timing of eye-movement programming and are only observable at a delay, once $n + 1$ is fixated (Kliegl, Risse, & Laubrock, 2007; Risse & Kliegl, 2014). However, one example in which POF effects are often (but not always; e.g., White & Liversedge, 2006; White et al., 2005) observed is when the invalid preview is an orthographically illegal nonword (Drieghe, 2011; Drieghe, Brysbaert, & Desmet, 2005; Inhoff, Starr, & Shindler, 2000; Payne & Stine-Morrow, 2012; Pynte, Kennedy, & Ducrot, 2004), presumably because the system is able to parafoveally recognize that the preview violates the orthographic conventions of the language. Orthographically invalid previews also result in the largest preview benefits (Henderson & Ferreira, 1990; Hyönä et al., 2004), suggesting that at least some portion of the preview benefit from invalid previews may also result from a *cost* of early parafoveal recognition of the invalid preview (cf. Kliegl, Hohenstein, Yan, & McDonald, 2013).

Importantly, the amount of information available from parafoveal vision does not appear to be static, but instead varies both between individuals (Payne & Stine-Morrow, 2012; Rayner, Castelhano, & Yang, 2009; Veldre & Andrews, 2015) and dynamically within an individual from fixation to fixation, such that concurrent foveal processing dynamically modulates the amount of information derived from parafoveal vision. Henderson and Ferreira (1990) first demonstrated the foveal load effect: When word n was more difficult to process (e.g., low in word frequency), the parafoveal preview benefit on $n + 1$ was reduced relative to when word n was less difficult to process (e.g., high in word frequency). Although widely accepted in the eye-movement literature, there are surprisingly only a few narrow empirical demonstrations that foveal load impacts parafoveal processing during reading, with findings largely restricted to effects of foveal lexical frequency (Henderson & Ferreira, 1990; Reingold & Rayner, 2006; Schroyens et al., 1999; White et al., 2005). Thus, the overall scope of foveal load effects is not well understood. In E-Z Reader (Reichle et al., 2009; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 2003), a serial-attention-based computational model of eye-movement control, foveal load effects are modeled as a lengthening of lexical access time. This delay in lexical processing time (i.e., reaching the "L2" stage; see Reichle et al., 1998) results in a delay in the allocation of attention to $n + 1$, as lexical access triggers attentional shifts to $n + 1$ in the model. Thus, when the time to lexical access is delayed, there is less time between the shift of covert attention to $n + 1$ and the execution of a saccade to $n + 1$, resulting in reduced parafoveal processing of $n + 1$ when foveal load is high. Higher level aspects of semantic and syntactic analysis have been argued to impact processing only after lexical access has occurred and, thus, after attention has already been allocated to word $n + 1$. An important implication of this model is that foveal load effects must be isolated to processes that lengthen lexical access time, such as effects of word frequency.

In the current study, our focus was on examining whether semantic features beyond foveal lexical frequency induce foveal load effects and modulate parafoveal processing of the word to the right of fixation. Only three experiments have examined whether nonlexical foveal manipulations induce foveal-load effects. In their second experiment, Henderson and Ferreira (1990) found a reduced parafoveal preview benefit when word n marked a syntactically disambiguating point in a garden-path sentence. More recently, Payne and Stine-Morrow (2012) and White, Warren, and Reichle (2011) showed that word position effects modulate the magnitude of the preview benefit, such that words following clause and sentence boundaries can elicit reduced preview benefits relative to intrasentential words. Thus, there is some limited evidence that effects related to structural aspects of the language (e.g., syntactic integration and wrap-up) can modulate covert attention to parafoveal processing. Currently, however, models of eye-movement control make no clear distinction among sources of foveal load, despite differing assumptions about what kind of information becomes available before attention spreads to information outside of foveal vision. To expand this understanding, in the current study, we focused on the effects of foveal semantic expectancy (i.e., target word predictability) and semantic incongruity on parafoveal word processing.

Semantic Contextual Constraints in Reading

The influence of sentential semantic constraints on word recognition has been an important and oft-investigated topic in psycholinguistics, with a substantial literature demonstrating that word processing is facilitated in supportive contexts across measures of behavior (e.g., Fischler & Bloom, 1979; Schwanenflugel & LaCount, 1988), electrophysiology (see Kutas & Federmeier, 2011, for a review), and eye movements during reading (see Staub, 2015, for a review). The current study focuses on two forms of contextual constraints that readers appear particularly sensitive to: (a) effects of word expectancy or predictability (Kutas & Hillyard, 1984; Rayner & Well, 1996), that is, the probability with which a target word can be predicted based on its prior context, typically as determined by a word's offline cloze probability (Taylor, 1953); and (b) effects of semantic congruity (e.g., typicality, plausibility, anomaly; Kutas & Hillyard, 1980; Rayner, Warren, Juhasz, & Liversedge, 2004; Warren, 2011), the degree to which a target word can be integrated with its prior semantic context to form a coherent message-level semantic representation.

We introduce the distinction between these by way of an example. A highly constraining sentence frame such as "The rude waiter was not given a . . ." elicits *tip* as the best completion, with a high cloze probability rating of 94%. Consider two less predictable alternate completions for this sentence context: *tray* and *lung*. Although both of these completions have a cloze probability of near 0%, they differ substantially in their final message-level semantic representation, with implications for the way in which the comprehension system responds when presented with such completions. *Tray*, although not predictable from the prior sentence context, is still a plausible completion, and, as such, comprehension should be able to continue with some fluency (perhaps following the revision or reanalysis of the sentence context; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007; Wlotko & Federmeier, 2012). However, the completion *lung*, which is equally unpredictable (i.e., equally low cloze probability), is additionally not a plausible completion based on the prior context, yielding an incongruent or anomalous message-level semantic representation and likely disrupting normal comprehension.

Words that are predictable and congruent with their prior semantic context are fixated for less time and are skipped more frequently in natural reading (Rayner & Well, 1996). Similarly, the N400, an ERP component strongly linked with meaning processing and initial access to semantic memory (see Kutas & Federmeier, 2011), shows graded reductions in amplitude with increasing word predictability (Kutas & Hillyard, 1984) and accumulating contextual constraints (e.g., Payne, Lee, & Federmeier, 2015). The source of such facilitative effects across methods are theorized to be attributable to several possible mechanisms, including preactivation of related semantic features (Kutas & Federmeier, 2000) and increased lexical predictability with accumulating semantic constraints (see Staub, 2015, and DeLong, Troyer, & Kutas, 2014, for recent reviews). In E-Z Reader, effects of word predictability are modeled as facilitating early stages of lexical processing (Reichle et al., 2003). Along with a word's frequency, these two features are argued to be the largest contributors to the amount of time a word is fixated (see Rayner, 2009, for a review).

In contrast to effects of word expectancy in context, the effects of encountering semantic incongruities are less clearly understood (Abbott & Staub, 2015; Clifton, Staub, & Rayner, 2007; Van Petten & Luka, 2012; Warren, 2011). It is generally found that processing costs are observed for words that are semantically incongruent with their prior context, including increased fixation durations (Rayner et al., 2004; Warren & McConnell, 2007) and a biphasic ERP response, with larger N400s followed by late positive potentials (which, as discussed below, are different for unexpected but plausible words vs. semantically anomalous ones; DeLong, Quante, & Kutas, 2014; Federmeier et al., 2007; Van Petten & Luka, 2012). However, the scope and timing of these costs vary across studies, as does the nature of the semantic incongruity. For example, Rayner et al. (2004) manipulated the severity of a semantic plausibility violation by manipulating the contexts preceding a target word (*carrots*), such that, as in the following example, it was either plausible, as in the first sentence; implausible, as in the second sentence; or completely semantically anomalous, as in the third sentence:

1. John used a knife to chop the large *carrots* for dinner.

2. John used an axe to chop the large *carrots* for dinner.

3. John used a pump to inflate the large *carrots* for dinner.

Rayner et al. (2004) found that both semantic incongruities resulted in disruptions to reading, but they manifested in the eye-movement record in different ways, with effects for the third sentence occurring on early measures such as gaze durations (gd) and showing substantial spillover effects onto the following words, but effects for the second sentence seen only on measures believed to reflect later stages of processing, such as go-past time (the sum of all fixation durations beginning with the first fixation on the word until the eyes go past the word to the right, including any regressive fixations). Warren (Warren, 2011; Warren & McConnell, 2007) has argued that the time course of plausibility effects is largely determined by the severity of the semantic violation. More recently, Stites and Federmeier (2015) have argued that the time course of semantic plausibility effects in the eye-movement record are likely to be "smeared" across multiple words following the source of the violation (see also Dambacher & Kliegl, 2007; Matsuki et al., 2011), in part because of the multitude of ways in which readers may respond to semantic violations (e.g., an immediate slow down at the target word, a delayed spillover response on the following word, launching an immediate regression), effects that likely vary from reader to reader. Moreover, given that saccadic programming is initiated with incomplete information about the semantics of the currently fixated word (Reichle et al., 2003), it is likely that foveal and parafoveal semantic effects can be delayed, "spilling over" onto the following words (Matsuki et al., 2011; Rayner et al., 2004; Risse & Kliegl, 2014). Importantly, an additional possibility is that some of the variance attributed to delayed "spillover" effects (cf. Rayner et al., 2004) is actually due to foveal-load effects from word n that, in turn, impact the amount of information derived from word $n + 1$ (Henderson & Ferreira, 1990; Schroyens et al., 1999).

The findings introduced above suggest that effects of semantic congruity violations may be graded in nature, and are consistent with a large literature of electrophysiological work showing graded responses of the N400 to semantic fit. For example, in a variation of the related anomaly paradigm (Kutas, Lindamood, & Hillyard, 1984), Federmeier and Kutas (1999) found that in strongly constraining contexts that targeted a particular completion (e.g., *palms*), within-category semantic violations (e.g., *pines*) elicited less N400 activity than between-category violations (e.g., *tulips*). A more direct comparison of expectancy and plausibility manipulations within the same study comes from DeLong and colleagues (2014), who recorded ERPs to words completing highly constraining contexts (e.g., "It was difficult to understand the visiting professor. Like many foreigners, he spoke with a . . .") that were completed with (a) an expected word (*accent*), (b) an unexpected but plausible word (*lisp*), or (c) a semantically incongruent (*apron*) target word. They replicated the widely found graded response of the N400, with the largest N400s to the semantically incongruent target, followed by the unexpected but plausible target, then the expected target. In addition, the expectancy violation and the congruity violation showed qualitatively dissociable effects following the N400, with unexpected but plausible completions showing a slow anterior positivity, and semantically anomalous completions showing a P600-like posterior positivity (see also Van Petten & Luka, 2012). These findings suggest that sensitivity to variations in lexical expectancy are dissociable from effects of semantic plausibility, and that this dissociation may reflect distinct costs of revising disconfirmed predictions, as in the case of the unexpected but plausible completion (e.g., Federmeier et al., 2007), versus an outright failure of semantic integration, as in the case of the semantic incongruity (Delong et al., 2014; Federmeier et al., 2007; Federmeier, Kutas, & Schul, 2010; Thornhill & Van Petten, 2012).

Whether violations of semantic congruity engender early disruptions of semantic processing versus other processing costs in later stages of semantic integration is currently open to debate (Abbott & Staub, 2015; Lau, Phillips, & Poeppel, 2008; Reichle, Warren, & McConnell, 2009; Warren, 2011). The time course of the sensitivity to message-level expectancy and congruity violations has implications for understanding the nature of semantic processing and its interaction with the allocation of attention in reading. Recent attempts have been made to computationally model the impacts of higher level aspects of language processing on attention allocation and eye movements in reading, such as in cases in which a word is implausible or difficult to integrate based on its prior semantic or syntactic context (Reichle et al., 2009). In the latest version of E-Z Reader, such processes are modeled as occurring at a late "postlexical integration" stage that serially follows the completion of lexical access. Importantly, because the completion of lexical processing of word *n* triggers the shift of attention to a parafoveal word, and postlexical integration only begins after lexical processing has completed, foveal-load effects on parafoveal processing are largely structurally constrained to arise only from lexical sources of processing difficulty. In the current study, we explore whether contextual sources of processing difficulty can induce foveal-load effects and constrain parafoveal processing. An empirical investigation into how and when foveal load can impact parafoveal processing requires being able to continuously monitor covert attention to parafoveal words in real time, which is not easily accommodated with current eye-movement paradigms. ERP methods may, however, offer the capability to probe these intermediate covert stages of processing.

Event-Related Potentials (ERPs), Covert Monitoring, and Parafoveal Processing

One limitation of existing foveal-load studies is that they have almost entirely relied on paradigms that study the *aftermath* of parafoveal processing, drawing inferences about the nature of parafoveal processing based on measures derived from the foveal processing of words that were previously masked in parafoveal vision. Indeed, in almost all experiments utilizing the gaze-contingent boundary change paradigm, effects of foveal load on parafoveal processing are not observed until word $n + 1$ is foveated (i.e., on the parafoveal preview benefit of $n + 1$; Henderson & Ferreira, 1990; Payne & Stine-Morrow, 2012; White et al., 2011). Several important stages of parafoveal processing leading up to fixation durations on $n + 1$ are thus likely to be confounded, including perceptual and cognitive processing of the stimulus in parafoveal vision, modulation of covert visual attention, and integration of parafoveal and foveal visual representations across successive saccades. ERPs have a long and extensive history in successfully addressing topics that have been controversial and elusive to study in the behavioral literature related to covert attentional phenomena, including spatial selective attention (Mangun, 1995), serial and parallel visual search (Luck & Hillyard, 1990), attentional bottlenecks (Vogel, Luck, & Shapiro, 1998), and sustained and divided visual attention (Müller, Malinowski, Gruber, & Hillyard, 2003; see Luck, 2012, for a review).

Recently, a number of studies have examined parafoveal processing in reading by utilizing ERPs, which offer the opportunity to continuously and covertly monitor parafoveal processing in real time (e.g., Barber, Ben-Zvi, Bentin, & Kutas, 2011; Barber, van der Meij, & Kutas, 2013; Dimigen, Kliegl, & Sommer, 2012; Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011, Kretzschmar, Bornkessel-Schlesewsky, & Schlesewsky, 2009; Li, Niefind, Wang, Sommer, & Dimigen, 2015). One approach to examining ERPs in reading has been to coregister electroencephalogram (EEG) and eye-tracking data within the same experiment to examine *fixation-related potentials* (reviewed in Dimigen et al., 2011). An alternative approach, developed by Barber, Kutas, and colleagues (Barber et al., 2011, 2013; Barber, Doñamayor, Kutas, & Münte, 2010), has been to modify the traditional ERP–RSVP paradigm by adding hemifield flankers in the left and right visual field, corresponding to the prior and subsequent word in the sentence, respectively. In this paradigm, participants fixate on a centrally presented word that is flanked on the right and left visual field by the preceding and following word, separated by 2° of visual angle. On the immediately following trial, each word is shifted to the left, such that the sentence appears to shift successively across the screen from right to left. Although not naturalistic, this method has proven to be a useful tool in examining parafoveal processing during reading, providing a bridge between eye movement and ERP studies of language processing, while maintaining much of the experimental and EEG artifact control that is a strength of traditional ERP studies (i.e., homogeneity in visual input across trials, control and assessment of component overlap, reduction of eye-movement artifacts).

Current research activity within this paradigm has centered on the nature of parafoveal visual representations, largely focusing on the degree to which semantic information can be extracted from parafoveal vision. These studies have shown that semantic congruency effects can be observed for words in the parafovea, especially when sentence contexts are highly constraining, influencing both sensory/perceptual components (the N1 and P2; Barber et al., 2010; Kornrumpf & Sommer, 2015) as well as later components indexing semantic access, such as the N400 (Barber et al., 2011, 2013; Li et al., 2015). Other work (Dimigen et al., 2012; Li et al., 2015) has focused on how processing semantically related parafoveal previews impacts electrophysiological indices of subsequent foveal processing (e.g., preview benefit effects). These studies have identified a *preview positivity* effect following the fixation of word $n + 1$, such that a word that initially had a valid parafoveal preview shows a more positive potential over occipitotemporal channels between 200 ms and 300 ms following its direct fixation, relative to a case when the word was previously masked with a different word. This potential appears to index the activity associated with integrating parafoveal and subsequent foveal visual representations, and has been argued to reflect partial orthographic priming, but does not *directly* index parafoveal perception, as it is not observed until after word $n + 1$ is fixated.

It remains to be seen whether localized modulation of foveal load impacts parafoveal processing in the flanker ERP paradigm in the same way that foveal processing difficulty influences the parafoveal preview benefit in natural reading (cf. White et al., 2005). Moreover, no study has yet examined the electrophysiological response to orthographically illegal previews, which are arguably easier to detect in parafoveal vision (cf. Drieghe, 2011; Drieghe, Brysbaert, et al., 2005). As described above, in the classic foveal-load paradigm, identity previews are often contrasted with invalid nonword previews while foveal load is simultaneously manipulated. Although this paradigm has generated robust (frequency-based) foveal load effects in the eye-movement record, neither the neural response to orthographically illegal parafoveal previews nor the impact of foveal load on this processing has been characterized with ERPs.

Because ERPs yield continuous and direct measures of sensory and cognitive processing beginning at stimulus onset, we can monitor the direct effects of foveal load on parafoveal processing in real time, rather than inferring the effects of foveal load on parafoveal processing after it has already occurred (i.e., effects on the parafoveal preview benefit). At the same time, only one study employing the parafoveal flanker ERP paradigm has cross-validated their results with an eye-tracking paradigm (Li et al., 2015). Thus, a central goal of the current study was to examine whether ERP and eye-tracking indices of foveal load effects yield converging conclusions regarding the role of foveal semantic predictability and congruity on parafoveal processing.

Experiment 1

The aim of the first experiment was to examine the impact of foveal semantic expectancy and congruity on parafoveal word processing in natural reading. Toward this goal, we used a traditional gaze-contingent boundary change paradigm, wherein we simultaneously manipulated the foveal semantic expectancy and congruity of a target word n (expected, unexpected but plausible, and incongruent) embedded within a highly constraining sentence context while manipulating the validity of word $n + 1$ in parafoveal vision (valid identity preview vs. orthographically illegal invalid preview). If effects of prior semantic context are available to the system early enough to modulate covert attentional control, then we would expect to observe evidence of foveal-load effects on parafoveal processing, such that the magnitude of the parafoveal preview benefit to $n + 1$ is reduced when n is unexpected with the prior context relative to when n is expected, and possibly even further reduced when unexpected words are incongruent with the preceding sentence context.

Method

Participants. Twenty-four adults from the University of Illinois at Urbana-Champaign community participated in the experiment for course credit. All were right-handed native speakers of English, reported near 20/20 corrected or uncorrected vision, and had no prior history of neurological or psychiatric issues.

Eye-tracking recording and processing. Ten subjects' eye movements were monitored with an EyeLink II (500 Hz) head-mounted eye tracker and a 19-in. ViewSonic P225f monitor set to a resolution of 1,024 × 768 with a refresh rate of 120 Hz. Fourteen subjects eye movements were monitored with an EyeLink 1000 (1,000 Hz) desktop-mounted eye tracker and a 22-in. Cornerstone P1750 monitor set to a resolution of 1,024 × 768 with a refresh rate of 85 Hz. Viewing was binocular, but only movements of the right eye were recorded. Participants were seated 97 cm from the monitor. At the viewing distance, three letters subtended about 1° of visual angle. The instructions and passages were displayed in a white font (16 pt Courier New) on a black background. Fixations less than 80 ms and within a half degree of visual angle were merged. Remaining fixations <80 ms and >1,000 ms were discarded. All trials in which the display did not change by the start of the first fixation on the target were discarded. Overall, this trimming procedure resulted in an average of 22% of the experimental trials being excluded from analysis, which is on par with other studies using the boundary-change paradigm.

Despite the differing monitor refresh rates and eye-tracker sampling rates between the two systems, a comparison of display change latencies (i.e., the lag between the trigger of the boundary change and the completion of the monitor retrace) revealed no reliable difference between the two systems (EL-1000 = 11.69 ms; EL-II = 12.92 ms; t = 1.22). In addition, we compared target eye-movement measures between the two recording systems and found no reliable difference between the systems in first fixation durations (ffd; EL1000 = 235 ms; EL-II = 259 ms; $t < 1$), word skipping (EL1000 = .20; EL-II = .17; $t < 1$), or regressions out (EL1000 = .17; EL-II = .15; $t < 1$). Therefore, results are presented collapsed across the two recording systems.

Materials. The stimuli consisted of 60 sentence frames, modified from Federmeier et al. (2007). Sentences were identical up through word n across all conditions. Highly constraining sentence contexts were continued with either the most expected word n (mean cloze = 94%), an unexpected but plausible word (mean cloze <1%), or a semantically incongruent word of the same grammatical class (mean cloze = 0%). As can be seen in Table 1, expected, unexpected but plausible, and incongruent words were identically matched in word length on a single-trial basis, and were further matched on word frequency (derived from the Corpus of Contemporary American English [COCA]; Davies, 2008), as well as perceived concreteness and imageability, based on norms collected from the medical research council (MRC) Psycholinguistics Database (Wilson, 1988).

These conditions were factorially combined with the parafoveal preview of $n + 1$ (valid preview vs. invalid preview) to create six experimental conditions. Sentence stimuli were rotated through all conditions following a Latin square design to create six counterbalanced lists, such that all subjects saw each sentence frame in only one of the six conditions (i.e., no sentences were repeated within an experimental session). Sentences were presented in one fixed, but random, order. An example set is presented below:

1. *Predictable n*: Maria remembered to shut the front window and lock the back **door** before/qzrvvk leaving for vacation.

2. *Unpredictable but plausible n*: Maria remembered to shut the front window and lock the back **room** before/qzrvvk leaving for vacation.

Table 1
Word N Lexical Characteristics (Mean and Standard Deviation) in Experiments 1 and 2

Lexical Characteristics	Expected	Unexpected	Incongruent	*t*[E−U]	*t*[E−I]	*t*[U−I]
Experiment 1: Eye tracking						
Word length	4.75 (.92)	4.75 (.92)	4.75 (.92)	0	0	0
Word frequency	10.42 (1.25)	10.09 (1.68)	10.44 (.90)	1.20	1.43	.13
Concreteness	517 (105)	496 (94)	501 (68)	.25	.99	.33
Imageability	542 (86)	521 (74)	534 (63)	.15	.58	1.03
Experiment 2: ERP						
Word length	4.82 (1.20)	5.68 (1.99)	5.29 (1.37)	**4.96**	**3.46**	**2.16**
Word frequency	9.81 (1.46)	9.61 (1.42)	9.59 (1.22)	1.26	1.57	.22
Concreteness	522 (104)	508 (104)	511 (86)	1.27	1.09	.30
Imageability	547 (86)	530 (81)	534 (72)	1.93	1.56	.50

Note. Word length in characters. Word frequency is the natural log transformed rate of occurrence per million from the Corpus of Contemporary American English (Davies, 2008). Concreteness and imageability are measured on a scale from 100 (*low*) to 700 (*high*). Rightmost columns present *t* statistics comparing expected (E), unexpected (U), and incongruent (I) conditions. Statistically significant *t* statistics are bolded. ERP = event-related potential.

3. *Incongruent n*: Maria remembered to shut the front window and lock the back **note** before/qzrvvk leaving for vacation.

For $n + 1$, the invalid preview was always a random string of visually dissimilar consonants of the same length as the preview word. Words n and $n + 1$ varied in length from four to seven characters ($M = 5.35$ characters). Word $n + 1$ was never sentence-final. Because the stimuli were modified from sentence sets in which word n was originally sentence-final, word $n + 1$ always served as the beginning of a new phrase or clause and was always a semantically sparse closed-class word. Word $n + 1$ was as an adjective or adverb in 20% of the experimental trials, a conjunction in 40% of the trials, a determiner in 5% of trials, and a preposition in 21% of trials. All sentences were less than 80 characters and appeared on a single line.

Procedure. The gaze-contingent boundary change paradigm was used to manipulate parafoveal word information on word $n + 1$. An invisible boundary was placed between the penultimate and final letter of word n. The average delay between the boundary trigger and the display change was 12 ms (range = 4 ms to 21 ms). Following the eye-tracking session, participants were administered a delayed sentence recognition task (in which half of the items were old and half were new) in order to ensure that participants were attending. A debriefing interview was also administered that probed participant's awareness of the display change. Six participants reported noticing the display change more than 3 times. Removing these subjects did not change the pattern of results, so results are presented for the full sample.

Data analysis. Eye-movement analyses are presented first for word n, followed by $n + 1$. Measures include single fixation duration (sfd), first fixation duration (ffd), gaze duration (gd), regression path duration (rpd), probability of launching a first-pass regression (*p*Reg), and probability of skipping (*p*Skip). Linear mixed-effects models were fit to the data for fixation durations and generalized linear mixed models with a logit link function were fit to the dichotomous responses (*p*Reg, *p*Skip). Subjects and items were treated as completely crossed random effects, and all models were fit with variance parameters for the random slope of the within-subject Preview × Context contrast interactions (see below) across subjects and items (excluding correlations between random intercepts and slopes to reduce model complexity; Barr, Levy, Scheepers, & Tily, 2013). Although the fixation time distributions were positively skewed, models fit to the raw and natural log transformed fixation durations resulted in no critical difference in the pattern of findings. Thus, model results fit to the raw (untransformed) data are presented to facilitate interpretation of fixed-effect parameters. For both logistic models, the random-slope structure would not converge to a valid solution, so results are presented with random intercepts only. Outlier values of rpd (≤99th percentile) were removed prior to analysis.

Because our a priori focus was on interactions between the Context and Preview factors, when reliable interactions were found, these were reported and explored in more detail. When interactions were not found, lower order effects were reported in more detail. Because the Context factor has three levels, the expected condition was treated as the reference group to form two contrasts: the *expectancy effect* (C1)—expected versus unexpected; and the *congruity effect* (C2)—expected versus incongruent. To directly compare the unpredictable but plausible and incongruent conditions, models were refit, treating the unpredictable condition as the reference group, in order to estimate the contrast between unpredictable and incongruent conditions. Fixed-effect parameter estimates are presented to indicate effect size, along with 95% profile likelihood confidence intervals (CIs) for statistical inference.

Lastly, to directly test whether there was a graded effect of word n context on the magnitude of the preview benefit to $n + 1$, a linear trend analysis was conducted, testing whether there was a reliable linear trend of contextual fit and whether this effect reliably interacted with preview validity. To estimate the linear context trend, the context factor was treated as a continuous variable (expected < unexpected < incongruent) by assigning linear interval contrast weights to each level of the factor. This variable was then treated as a continuous regressor in the mixed-effects models to estimate the linear change in fixation durations (and log odds change in probability measures) per unit change in context. The slope estimate thus indicates a linearly graded change in the outcome per unit change in degree of semantic violation. Importantly, this variable can be treated as a moderator of the Preview factor to directly test whether the magnitude of the preview benefit decreases in a linearly graded fashion with an increasing degree of semantic violation (cf. Kutas & Federmeier, 2011; Warren, 2011).

Results and Discussion

Recognition memory. Participants correctly recognized an average of 66% of the experimental sentences ($SD = 18\%$) and false alarmed to an average of 9% of experimental sentences ($SD = 8\%$). Because a subset of participants showed a false alarm rate of zero, the parametric d' index was undefined. Thus, signal detection sensitivity was calculated using the A-index (see Zhang & Mueller, 2005). Mean A was .86, 95% CI [.84, .89], indicating that participants were successfully discriminating between old and new sentences. Thus, participants appeared to be attending to the experimental materials.

Word *n*. Table 2 presents fixation duration and eye-movement probability measures for word *n*. No interactions between context and preview validity were reliable for any measure, although for rpd, there was a numerical trend for a POF effect in the expected and unexpected conditions only.

Overall, fixations were longer on word *n* when it was unexpected compared with expected for sfd ($b = 44$ ms, 95% CI [19, 63]), ffd ($b = 31$ ms, 95% CI [12, 49]), gd ($b = 42$ ms, 95% CI [16, 70]), and rpd ($b = 72$ ms, 95% CI [33, 110]). Words were skipped slightly more frequently when *n* was expected relative to unexpected ($b = -.57$, 95% CI [−1.06, −.09]), and *p*Reg was marginally larger when *n* was unexpected compared with expected ($b = .65$, 95% CI [−.04, 1.37]). When *n* was incongruent, fixation durations were longer compared with expected words for sfd ($b = 62$ ms, 95% CI [31, 93]), ffd ($b = 46$ ms, 95% CI [27, 64]), gd ($b = 75$ ms, 95% CI [49, 101]), and rpd ($b = 114$ ms, 95% CI [76, 153]). Incongruent words also showed a larger proportion of regressions ($b = .86$, 95% CI [.19, 1.58]) and a smaller proportion of word skipping ($b = -.51$, 95% CI [−.99, −.03]) relative to expected words. The only reliable effect of $n + 1$ preview validity found on word *n* was that gd was reliably longer when $n + 1$ was an invalid preview compared with a valid preview ($b = 30$ ms, 95% CI [3, 57]). Directly comparing the unexpected and incongruent conditions revealed a marginal difference in ffd ($b = 15$ ms, 95% CI [−3, 33]) and reliable differences in gd ($b = 32$ ms, 95% CI [14, 50]), rpd ($b = 42$ ms, 95% CI [3, 81]), and *p*Reg ($b = 1.44$, 95% CI [.98, 1.92]). Lastly, a direct test of the graded effects of semantic violation severity, examined via a linear trend analysis (see Data Analysis above), revealed a reliable graded linear effect of context (expected < unexpected < incongruent) for sfd ($b = 27$ ms, 95% CI [15, 39]), ffd ($b = 18$ ms, 95% CI [11, 25]), gd ($b = 30$ ms, 95% CI [19, 41]), rpd ($b = 46$ ms, 95% CI [29, 63]), *p*Reg ($b = .34$, 95% CI [.09, 59]), and *p*Skip ($b = .22$, 95% CI [.02, .42]).

Overall, the pattern of findings suggest that the fit between a fixated word and its prior context has a robust effect on eye movements at that word, impacting the probability of word skipping and regressions, as well as both first-pass (sfd, ffd, gd) and later (rpd) fixation duration measures. Word skipping was greatest, fixation durations were smallest, and regressive eye movements were the least probable when *n* was expected relative to both the unexpected and incongruent conditions. Moreover, both a comparison of the unexpected and incongruent conditions and the results from the linear trend analysis were consistent with the claim that the severity of the semantic violation impacted fixation durations in a graded manner. Importantly, there was little evidence that the parafoveal status of $n + 1$ modulated eye movements while the eyes were fixated on word *n*, with effects only reaching traditional levels of statistical significance for gd. These findings are consistent with prior research utilizing the boundary change paradigm, which suggest that even striking parafoveal manipulations (i.e., of orthographic legality) do not routinely impact foveal fixation durations to that word, and are instead often observed downstream on fixation durations to word $n + 1$ (cf. Risse & Kliegl, 2014).

Word *n* + 1. Table 3 presents fixation duration and eye-movement probability measures for word $n + 1$. To preview the results, in contrast to the findings for word *n*, there were reliable interactions between the parafoveal validity of $n + 1$ and the contextual fit of word *n* on first-pass measures for word $n + 1$. The overall pattern that emerged was that the preview benefit (difference between valid and invalid parafoveal previews) was largest when *n* was expected, but was systematically reduced when *n* was unexpected or incongruent with the prior context. There were reliable interactions between parafoveal preview and both the C1 and C2 contrasts for sfd ($b_{C1 \times P} = 34$ ms, 95% CI [2, 66]; $b_{C2 \times P} = 54$ ms, 95% CI [12, 97]) and ffd ($b_{C1 \times P} = 26$ ms,

Table 2
Average Temporal and Probability Measures on Word N as a Function of Foveal Semantic Load and Preview Validity (Between-Subject Standard Errors)

Condition	SFD	FFD	GD	RPD	*p*(Reg)	*p*(Skip)
Expected *n*						
Valid *n* + 1	203 (12)	208 (12)	223 (18)	253 (32)	.08 (.06)	.23 (.09)
Invalid *n* + 1	211 (15)	219 (15)	250 (22)	278 (27)	.09 (.06)	.25 (.09)
Unexpected *n*						
Valid *n* + 1	229 (17)	238 (18)	265 (25)	292 (35)	.07 (.05)	.23 (.09)
Invalid *n* + 1	251 (26)	249 (23)	293 (36)	344 (42)	.14 (.07)	.17 (.08)
Incongruent *n*						
Valid *n* + 1	236 (15)	249 (19)	293 (28)	389 (53)	.20 (.08)	.15 (.07)
Invalid *n* + 1	269 (19)	267 (26)	327 (36)	397 (50)	.17 (.08)	.19 (.07)

Note. SFD = single fixation duration; FFD = first fixation duration; GD = gaze duration; RPD = regression path duration; *p*(Reg) = probability of regression; *p*(Skip) = probability of skipping.

Table 3
Average Temporal and Probability Measures on Word N + 1 as a Function of Foveal Semantic Load and Preview Status (Between-Subject Standard Errors)

Condition	SFD	FFD	GD	RPD	*p*(Reg)	*p*(Skip)
Expected *n*						
Valid *n* + 1	211 (12)	215 (22)	246 (23)	261 (23)	.05 (.04)	.23 (.09)
Invalid *n* + 1	272 (18)	269 (25)	296 (25)	312 (25)	.12 (07)	.17 (.07)
PB	61	54	50	51	.07	.06
Unexpected *n*						
Valid *n* + 1	220 (14)	221 (27)	258 (27)	300 (27)	.10 (.06)	.21 (.09)
Invalid *n* + 1	243 (15)	250 (29)	286 (30)	352 (30)	.19 (.08)	.17 (.07)
PB	23	29	28	52	.09	.04
Incongruent *n*						
Valid *n* + 1	273 (23)	251 (33)	306 (33)	486 (33)	.36 (.10)	.16 (.08)
Invalid *n* + 1	276 (21)	267 (36)	317 (36)	529 (36)	.48 (.08)	.17 (.08)
PB	3	16	11	43	.12	–.01

Note. PB for pSkip is inverted because word skipping tends to occur for words that are easier to process. SFD = single fixation duration; FFD = first fixation duration; GD = gaze duration; RPD = regression path duration; *p*(Reg) = probability of regression; *p*(Skip) = probability of skipping; PB = preview benefit (Invalid – Valid).

95% CI [3, 49]; $b_{C2 \times P}$ = 34 ms, 95% CI [8, 62]). For gd, the Preview × C1 interaction was marginal ($b_{C1 \times P}$ = 32 ms, 95% CI [−4, 71]), whereas the Preview × C2 interaction was reliable ($b_{C2 \times P}$ = 43, 95% CI [2, 84]). Comparisons of first-pass differences in the preview benefit between unexpected and incongruent conditions did not reach statistical significance (sfd: b = 20 ms, 95% CI [−22, 63]; ffd: b = 9 ms, 95% CI [−18, 35]; gd: b = 11 ms, 95% CI [−31, 52]).

No reliable interactions were found for rpd, *p*Reg, or *p*Skip. However, words with invalid parafoveal previews had longer rpds (b = 52 ms, 95% CI [3, 101]) and more regressions out (b = .93, 95% CI [.15, 1.79]) than words with valid previews. In addition, when *n* was incongruent, rpds on *n* + 1 were considerably longer (b = 218 ms, 95% CI [170, 266]), and there were more regressions out of *n* + 1 (b = 2.01, 95% CI [1.48, 2.56]) compared with when *n* was expected, irrespective of preview validity. There was no effect of semantic context on skipping rates for *n* + 1. However, words with valid previews were more likely to be skipped than those with invalid previews (b = .51, 95% CI [.04, .99]).

Examining the magnitude of the preview benefit for first-pass measures in each context condition separately revealed that there was a robust and reliable preview benefit on *n* + 1 when word *n* was expected (sfd = 61 ms, t = 5.21; ffd = 54 ms, t = 6.21; gd = 50 ms, t = 4.79). When word *n* was unexpected, preview benefits were numerically smaller and only statistically reliable in first fixation duration (sfd = 23 ms, t = 2.03; ffd = 29 ms, t = 3.78; gd = 28 ms, t = 1.87). When *n* was incongruent, the preview benefit was not reliable across any first-pass measures (sfd = 3 ms, t = .68; ffd = 16 ms, t = 1.80; gd = 11 ms, t = .99). In a model treating semantic context as a linear trend (expected < unexpected < incongruent), there were reliable interactions between parafoveal preview and the semantic context linear trend for sfd (b = 23 ms, 95% CI [7, 39]), first fixation duration (b = 14 ms, 95% CI [4, 25]), and gd (b = 18 ms, 95% CI [2, 34]), supporting the argument that effects of foveal semantic load on attentional allocation to parafoveal processing were graded in nature.

In summary, fixation durations and eye movements on word *n* were largely sensitive to the effects of semantic constraints, with only a little evidence that the parafoveal status of *n* + 1 modulated fixation durations. However, beginning on first-pass measures at *n* + 1, we saw clear evidence that parafoveal processing was modulated by foveal semantic load, with large parafoveal preview benefits for *n* + 1 when *n* was expected, reduced preview benefits when *n* was unexpected, and no first-pass preview benefit to *n* + 1 when *n* was incongruent. These findings clearly demonstrate that foveal semantic congruity acts to modulate some aspects of parafoveal processing. In Experiment 2, we used ERPs to probe the nature and time course of these semantic foveal load effects.

Experiment 2

In Experiment 2, we measured ERPs to track the time course of semantic load effects on parafoveal processing. More specifically, our aim was to investigate the nature of the effects of foveal semantic expectancy and congruity on both initial parafoveal word processing (at word *n*) and the subsequent integration of foveal and parafoveal visual representations (at word *n* + 1). Both eye tracking and ERPs provide excellent temporal resolution for the online investigation of foveal and parafoveal processing in reading. However, the added benefit of ERPs is that they provide a *continuous* measure of processing, one that begins prior to stimulus onset and is not based on end-state behaviors (fixation durations/saccades). Because, in the boundary change paradigm, definitive conclusions about the locus of parafoveal processing (i.e., processing derived while the stimulus appears in parafoveal vision) cannot be readily investigated, ERPs may be a particularly beneficial tool for examining the effects of foveal load on the allocation of attention to parafoveal processing.

By examining converging evidence across both methods, we additionally aimed to validate important assumptions in both methods. To the extent that results from a flanker ERP paradigm converge with a more naturalistic eye-tracking paradigm, this would provide evidence that the ERP paradigm is tapping into parafoveal attentional processes that take place during natural reading, as opposed to reflecting task-specific demands. At the same time, because the preview benefit measures in eye tracking likely index multiple stages of parafoveal processing, findings from the flanker ERP paradigm will offer an important, direct view of the locus of foveal load effects.

Method

Participants. Twenty-four adults from the University of Illinois at Urbana-Champaign community participated in the experiment for course credit. All were right-handed native speakers of English, reported near 20/20 corrected or uncorrected vision, and had no prior history of neurological or psychiatric issues. None of the participants had previously provided data for Experiment 1.

Materials. Experimental sentences included the same 60 items as in Experiment 1, plus an additional 120 items utilizing the same experimental design as Experiment 1. As in Experiment 1, sentences were identical up through word *n* across all conditions. Highly constraining sentence contexts were continued with either the most expected word *n* (mean cloze = 89%), an unexpected but plausible word of the same grammatical class (mean cloze <1%), or a semantically incongruent word (mean cloze = 0%). Semantically incongruent words were drawn from the same grammatical class and were controlled for word frequency (COCA; Davies, 2008), and perceived concreteness and imageability, based on norms collected from the MRC Psycholinguistics Database (Wilson, 1988). The expected words, however, were slightly shorter than the incongruent words, which, in turn, were slightly shorter than the unexpected words (see Table 1). Note, however, that because the parafoveal manipulation is defined based on degrees of visual angle in Experiment 2, there is no confound between word *n* length and degree of parafoveal eccentricity of $n + 1$ as would be present in eye tracking. These conditions were factorially combined with the parafoveal preview of $n + 1$ (valid preview vs. invalid preview) to create six experimental conditions. Sentence stimuli were rotated through all conditions following a Latin square design to create six counterbalanced lists, such that all subjects saw each sentence frame in only one of the six conditions (i.e., no sentences were repeated within an experimental session). Sentences were presented in one fixed, but random, order.

For $n + 1$, the invalid preview was always a random string of visually dissimilar consonants of the same length as the preview word. Words *n* and $n + 1$ varied in length from three to 10 characters ($M = 5.21$ characters). Word $n + 1$ was never sentence-final. Because the stimuli were modified from sentence sets in which word *n* was originally sentence-final, word $n + 1$ always served as the beginning of a new phrase or clause and was always a semantically sparse closed-class word. Word $n + 1$ was as an adjective or adverb in 17% of the experimental trials, a conjunction in 37% of the trials, a determiner in 3% of trials, and a preposition in 43% of trials. Sentences ranged in length from seven to 27 words ($M = 14.58$).

EEG recording and processing. EEG was recorded from 26 evenly spaced silver–silver chloride electrodes embedded in an Electro-Cap (following the same montage as in Federmeier et al., 2007). Electrodes were referenced online to the left mastoid and rereferenced off-line to the average of the right and left mastoids. In addition, one electrode (referenced to the left mastoid) was placed on the left infraorbital ridge to monitor for vertical eye movements and blinks, and another two electrodes (referenced to one another) were placed on the outer canthus of each eye to monitor for horizontal eye movements. Electrode impedances were kept below 5 kΩ. The continuous EEG was amplified through a bandpass filter of .02 to 100 Hz and recorded to hard disk at a sampling rate of 250 Hz. EEG epochs were examined and marked for artifacts (drift, muscle activity, eyeblinks, and eye movements). Epochs of EEG data were taken from 100 ms before stimulus onset to 1,400 ms poststimulus onset (i.e., containing words *n* and $n + 1$). On average, a total of 6% (range across subjects = 0 to 23%) of words were marked as artifacts and not included in subsequent analyses.

Procedure. Participants were seated 85 cm from a 21-in. CRT computer monitor in a dim, quiet testing room. As in Barber et al. (2013), sentences were presented serially in triads, with the target word appearing at central fixation, flanked bilaterally by the upcoming word in the sentence to the right and the preceding word to the left. Participants were informed that multiple words and symbols would appear to the left and right of the central word, but to keep focus on the word presented at the center of the screen and read each sentence for comprehension. At the viewing distance, 3.5 letters subtended 1° of visual angle. The beginning letter of the right parafoveal word and the final letter of the left parafoveal word were anchored so as to appear at 2° of visual angle from the center of the screen. Each trial began with a series of fixation crosses ("+++++") that remained on the screen for a duration that was jittered from 500 ms to 1,500 ms. Each triad was visible on the screen for 100 ms, with an interstimulus interval of 350 ms (cf. Barber et al., 2013). This short stimulus duration also minimized the possibility that participants could make a saccade to the parafoveal word within the amount of time that it was visible on the screen. At the end of the sentence, a blank screen was presented for 1,950 ms, after which the fixation crosses appeared and the next sentence began. At the end of the experiment, participants completed a brief recognition memory test and were probed for their awareness of the display change. A total of 10 participants reported noticing the nonword in parafoveal vision. As in Experiment 1, these subjects did not show a markedly different pattern of results, so results are presented collapsed for the full sample. Analysis of the effects of processing expected, unexpected, and incongruent words in parafoveal vision (cf. Barber et al., 2013), derived from this same data set, are discussed in Stites, Payne, and Federmeier (2016).

Data analysis. In the current experiment, we were primarily interested in studying the neural activity indexing the discrimination between valid (identity) and invalid (orthographically illegal) parafoveal previews as a function of foveal semantic load. For items in foveal vision, processing differences between words and orthographically illegal strings of letters encompass multiple ERP components and effects, beginning (around 100 ms) on sensory potentials sensitive to factors such as bigram probability (e.g., N/P150, P200), and continuing (between 200 ms and 300 ms) through the N250, a component linked to orthographic processing, the N400 (300 ms to 500 ms), a component linked to semantic processing, and beyond, on late positivities sensitive to participants' evaluation of stimuli (for a review, see Holcomb & Grainger, 2006; Laszlo & Federmeier, 2014). However, none of these effects have been examined for items in parafoveal vision, nor have the effects of foveal semantic load previously been examined in ERP studies. Given the breadth of time windows and spatial locations across which effects might occur and the lack of prior data about how these effects manifest under the conditions employed here, the current study utilized a component-independent experimental approach (see Luck, 2014) to assess how activity associated with parafoveal preview discrimination was modulated by semantic foveal load. To explore these effects, we adopted two approaches.

The first approach was aimed at directly targeting activity associated with parafoveal discrimination between valid and invalid previews when foveal load was low, followed by examining how that specific activity changed under conditions of high foveal load. A time window was chosen based on visual inspection of the waveforms in the foveally expected condition, in which foveal load was minimal, and thus in which the valid versus invalid parafoveal effect should be largest. This window was then held constant across all semantic context conditions, to test whether activity specifically related to parafoveal target discrimination was moderated by foveal semantic load. Note that this approach does bias test statistics for the simple valid versus invalid test in the expected condition alone (cf. Kriegeskorte, Simmons, Bellgowan, & Baker, 2009), but, critically, it does *not* bias test statistics for the interaction between parafoveal preview and sentence contexts (i.e., foveal load effects), as the window choice is made independent of the data in the unexpected and incongruent conditions. Rather, this approach allows us to directly test our hypothesis that activity related to parafoveal preview discrimination in the low-load condition should be reduced when foveal load is increased (note that this approach is similar to the functional localizer approach applied widely in the fMRI literature; cf. Kanwisher, McDermott, & Chun, 1997).

Mean amplitudes in two separate windows, corresponding to effects following when word *n* and word $n + 1$ appeared at central fixation (see below), were submitted to linear mixed-effects models with subject and electrode site as random effects. These models are analogous to repeated-measures ANOVAs, with an additional random effect for channel site. Whole-head (all-scalp channel) analyses were conducted, given the widespread distribution of the parafoveal effect (see Figure 1b). This additionally reduces the need to pick particular channels for analysis, and generates estimates and test statistics that are robust across all channel sites, reducing the risk of false-positives through multiple-channel comparisons. Models were fit with random slope adjustments across subjects and channels, and for the critical preview by context interactions (see Barr et al., 2013).

In addition, we conducted an exploratory analysis of the valid versus invalid preview contrasts in each context type, utilizing a mass univariate analysis (Groppe et al., 2011a, 2011b). ERP mean

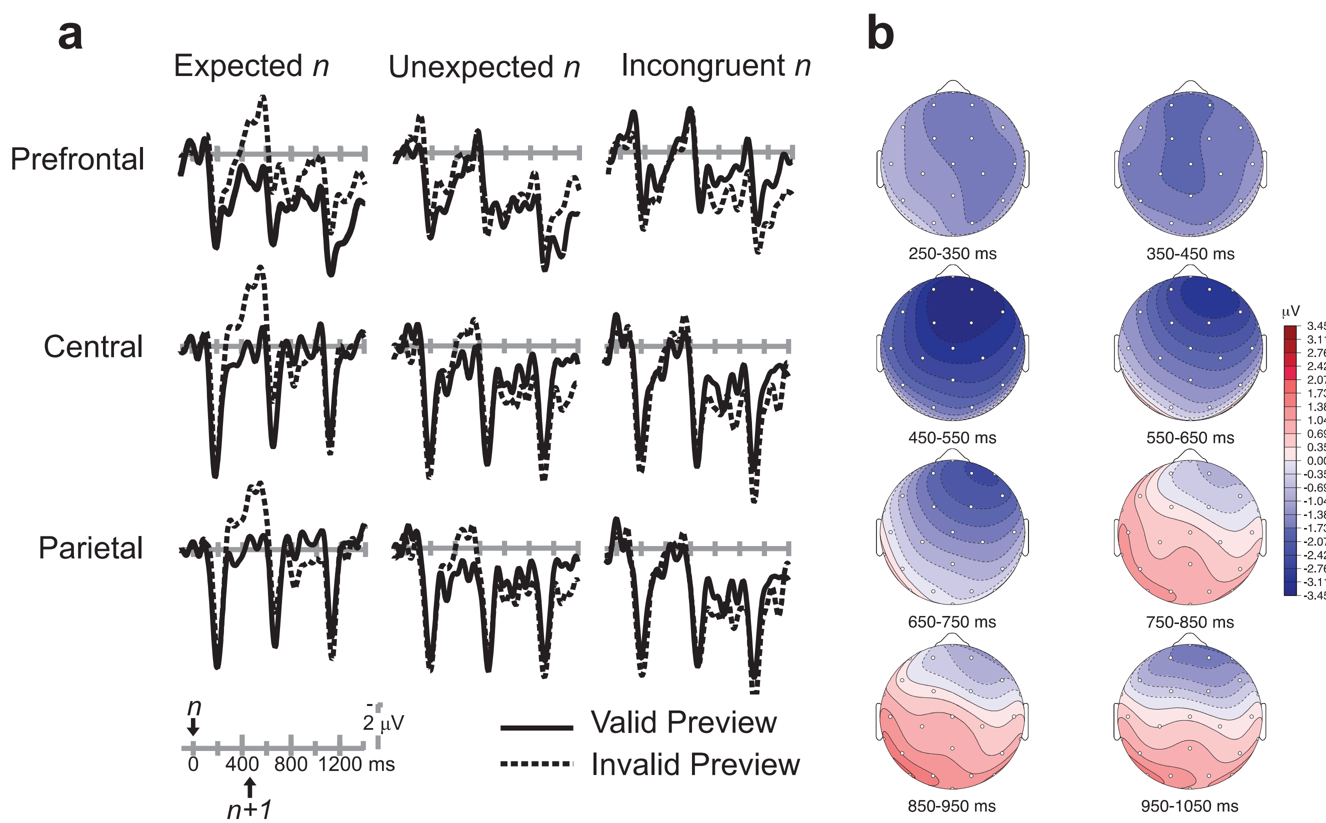

Figure 1. (a) Grand-average ERPs at representative midline electrode sites as a function of parafoveal word status (valid vs. invalid preview) and foveal semantic load (expected, unexpected, incongruent). (b) Topographic scalp map of the ERP nonword negativity (mean amplitude voltage difference between the invalid and valid preview conditions) when foveal load on word n was low (expected) over consecutive 100-ms time windows from 250 ms to 1,050 ms. A widely distributed, but anteriorly focused, negativity can be seen observed between 350 ms and 750 ms following the foveal presentation of word *n*. See the online article for the color version of this figure.

amplitudes were measured across a 50-ms moving window in the valid and invalid preview conditions separately in each context condition from 0 ms to 1,400 ms across all electrodes. The valid and invalid preview mean amplitudes were then submitted to repeated-measures *t* tests across all time windows, electrodes, and sentence types (2,184 comparisons in total). To protect against a large proportion of false-positives due to the massive number of multiple comparisons, we adopted a false discovery rate (FDR) control (see Benjamini & Hochberg, 1995). The local FDR was estimated as described in Strimmer (2008a), and *t* values that did not exceed the FDR-corrected critical threshold were considered nonsignificant. Analyses were conducted using the fdrtool package in R (Strimmer, 2008b). The mass univariate approach allows for a broader statistical exploration of ERP dynamics compared with traditional statistical approaches, but is limited by reduced statistical power compared with a priori selection of analysis parameters (see Groppe et al., 2011a, 2011b, for an in-depth discussion and tutorial in the context of ERP data).

Results and Discussion

Recognition memory. Participants correctly recognized an average of 53% of the experimental sentences ($SD = 16\%$) and false alarmed to an average of 13% of experimental sentences ($SD = 12\%$). Signal detection sensitivity was calculated using the A-index (see Zhang & Mueller, 2005). Mean A was .77, 95% CI [.69, .84], indicating that participants were successfully discriminating between old and new sentences. Thus, participants appeared to be attending to the experimental materials.

ERPs. Figure 1a shows grand-average ERPs at representative electrode sites (midline prefrontal, central, and parietal) as a function of foveal expectancy and parafoveal word status. Figure 1b plots the scalp topography of the valid versus invalid preview contrast over time in the expected *n* condition. A widespread negativity to invalid parafoveal targets was observed relative to valid parafoveal words when *n* was expected, beginning between 250 ms and 350 ms on average. This effect appears to signify parafoveal discrimination between orthographically illegal and valid parafoveal words. Note that although this effect is elicited within the traditional N400 latency band, its distribution is different from a typical visual word N400 effect, as it is broad but with an anterior focus (see Figure 1b); it is sustained for a longer period of time than the N400; and it also does not take the form that would be expected for an N400 effect to these conditions, as there is typically *less* N400 activity to foveally presented illegal strings than to orthographically legal words (e.g., Laszlo & Federmeier, 2011). Thus, the activity indexing the parafoveal recognition of nonwords is likely not an N400 effect.

Importantly, this parafoveal effect appeared to be modulated by foveal semantic load in a graded manner. When the foveal word was unexpected, the amplitude of the negativity to the invalid parafoveal target was reduced in magnitude. Furthermore, when *n* was incongruent, there did not appear to be any difference between valid and invalid parafoveal targets within the same latency band. Late differences between parafoveal words and nonwords were also observed, following the onset of word $n + 1$ (discussed in more detail below). Analyses were conducted separately on words *n* and $n + 1$ to quantify the magnitude of these observed differences; however, all analyses used the same prestimulus baseline to word *n*.

Word *n* in central position (parafoveal perception of $n + 1$). First, we conducted an analysis to examine whether ERP indices related to detecting the orthographically illegal string in parafoveal vision when foveal load was low (in the expected condition) was reliably reduced when foveal load was increased (unexpected and incongruent conditions). Mean amplitudes were measured from each subject and electrode site separately for each condition between 350 ms and 600 ms post onset of the triad with word *n* in central position. (Window selection was based on visual inspection of the valid vs. invalid contrast in the expected condition only, see Data Analysis above).

Mean amplitudes for each condition are presented in the upper portion of Table 4. There were reliable interactions between parafoveal word validity and the C1 and C2 contrasts ($b_{C1 \times P} = .93$ μV, 95% CI [.38, 1.48]; $b_{C2 \times P} = 1.28$ μV, 95% CI [.63, 1.95]), indicating that semantic expectancy and congruity modulated the difference in ERPs to valid and invalid parafoveal targets. The difference between the valid and invalid previews did not significantly differ between the unexpected and incongruent conditions ($b = .35$ μV, 95% CI [−31, 52]). When the centrally presented word was expected, there was a reliable difference between valid and invalid parafoveal targets of −1.44 μV ($t = -3.71$). However, when the foveated word was unexpected, this effect was reduced to a marginally significant effect of −.51 μV ($t = -1.79$). Finally, there was no reliable difference between parafoveal words and nonwords when word *n* was incongruent (−.15 μV, $t = -.41$). In a model treating foveal semantic load as a continuous linear trend (i.e., expected > unexpected > incongruent), there was a reliable interaction between parafoveal validity and the linear trend ($b = .53$ μV, 95% CI [−.45, 1.59]), supporting the argument that effects of foveal load on parafoveal processing were graded in nature. As can be seen in Figure 2, which plots both the preview benefit effects in Experiment 1 (Panel a) and the parafoveal negativity effects in Experiment 2 (Panel b), both eye tracking and ERPs show a very similar pattern of a graded reduction in sensitivity to parafoveal information as a function of increasing foveal load, consistent with the foveal load hypothesis. Importantly, these results suggest that increasing foveal load directly constrains the perception of the parafoveal word, as indexed by the reduced

Table 4
Mean Amplitude (Across All Scalp Sites) in Early Time Window (Word N) and Late Time Window (Word N + 1)

Context	Word *n* central (early window)	
	Valid $n + 1$	Invalid $n + 1$
Expected	.42 (.42)	−1.02 (.47)
Unexpected	1.03 (1.92)	.53 (.41)
Incongruent	.57 (.51)	.42 (.49)
	Word $n + 1$ central (late window)	
Expected	.82 (.47)	1.00 (.46)
Unexpected	1.99 (.53)	2.76 (.48)
Incongruent	1.96 (.58)	2.91 (.56)

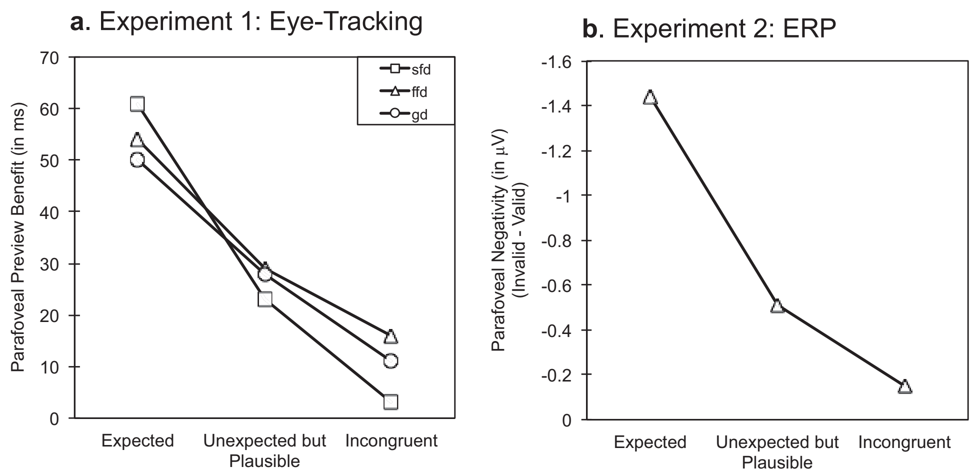

Figure 2. (a) Preview benefit effect (invalid preview – valid preview) on n 1 when word n was expected, unexpected, or incongruent for single fixation duration, first fixation duration, and gaze duration. (b) Mean amplitude of the nonword negativity (mean amplitude voltage difference between the invalid and valid preview conditions) between 350 ms and 600 ms poststimulus onset of word n in Experiment 2 when word n was expected, unexpected, or incongruent. Both experiments show the same pattern of decreases in parafoveal processing of word n 1 with increasing foveal semantic load.

discrimination between valid and orthographically illegal parafoveal previews.

Word $n + 1$ in central position (foveal processing of parafoveal target). A late difference was observed within the time window corresponding to the period following the sensory components to word $n + 1$, such that words that previously appeared with invalid targets elicited a larger relative positivity compared with words that had a valid parafoveal preview. This effect was measured from a broad time window corresponding to 700 ms to 1,100 ms following the onset of word n. Note that this corresponds to a time window of 250 ms to 650 ms following the onset of word $n + 1$. (Window selection was based on visual inspection of the valid vs. invalid contrast only in the expected condition; see Data Analysis above). However, we maintained the time locking and baseline to word n for this analysis, given that the large effects seen at word n in the expected and unexpected conditions would make the baselines inequivalent across conditions in an $n + 1$ locked average. Mean amplitudes for each condition are presented in the lower portion of Table 4.

The overall effect of parafoveal validity was reliable ($b = .27$ μV, 95% CI [.06, .48]), indicating that previously invalid parafoveal words showed a greater positivity once foveated relative to previously valid parafoveal targets. This is particularly striking given that, at this stage of processing, the visual information is identical within context conditions, so that this effect must be driven by prior parafoveal preview status, possibly indexing the integration of foveal and parafoveal visual representations of the same word across trials. The expected condition was also overall more negative than the unexpected ($b = 1.46$ μV, 95% CI [1.30, 1.62]) and incongruent conditions ($b = 1.52$ μV, 95% CI [1.36, 1.69]) during this time period, which appeared to reflect the continued sustained negative potential seen to invalid parafoveal targets over anterior scalp sites in the expected condition. There was no interaction between word n context and word $n + 1$ preview within this time window. Thus, this effect appears to be elicited to the same degree in all three semantic context conditions.

Mass univariate analysis. Figure 3 graphically represents the results from the mass univariate analysis (containing both central word n and central word $n + 1$ presentation; see Figure 1a). The figure is a raster-style heat map, plotting statistically thresholded t statistics contrasting the valid and invalid previews for each electrode, time bin, and sentence type. Those values exceeding the 95% FDR region are considered statistically significant. Values falling underneath the critical threshold are set equal to $t = 0$ and plotted in white. Darker coloration reflects larger differences between the valid and invalid conditions (see color figure online: Greater negative effect sizes to the invalid preview relative to the valid preview are blue, and more positive relative amplitudes to the invalid preview are plotted in red).

When word n is expected, a broadly distributed bilateral sustained negative potential is observed for the invalid previews, onsetting around approximately 300 ms and sustained until about 750 ms, mapping onto the sustained negativity seen in Figures 1a and 1b. This effect is seen most strongly over anterior and central scalp sites, but is less reliable over lateral parietal and occipital channels. Following the sustained parafoveal negativity, there is a reliable positive potential seen between 900 ms and 1,000 ms, primarily over left parietal and occipital channels, corresponding to the smaller positivity seen when the previously invalid $n + 1$ preview appeared in foveal vision.

When word n was unexpected, the parafoveal negativity seen in the expected condition is substantially reduced in magnitude and delayed in onset, indicating the reduced sensitivity to the invalid preview in parafoveal vision with increasing foveal load. However, once word $n + 1$ appeared in foveal vision, the system did appear to differentiate between the valid and invalid previews in the form of a sustained positive potential, albeit occurring at a delay relative to the expected condition, that is, only once word

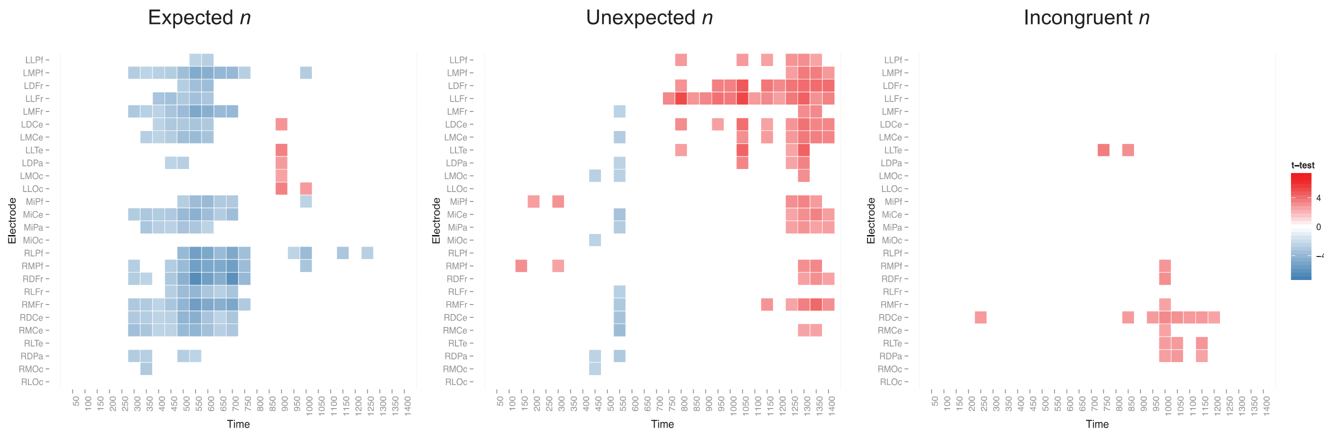

Figure 3. Raster-style heat-map plotting the FDR controlled *t* tests of the invalid–valid preview comparison when word *n* was expected, unexpected, or incongruent. Results are presented in 50-ms bins. Left-hemisphere electrodes are depicted in the upper portion of the figure, midline electrodes are presented in the center, and right-hemisphere electrodes are presented in the lower portions of each panel. Significant *t* tests for negative ERP differences are represented in blue and those for positive differences in red. Tests that did not exceed the FDR critical values are plotted in white. See the online article for the color version of this figure.

n + 1 appeared in foveal vision and was no longer invalid. Lastly, when word *n* was incongruent, there was no evidence that the system differentiated between the valid and invalid previews in parafoveal vision. Only after word *n* + 1 appeared in foveal vision was there evidence for a sustained positivity to invalid previews relative to valid previews (largely over right hemisphere electrodes). Thus, when foveal load was maximal (word *n* was semantically incongruent), the system did not appear to readily distinguish between valid and invalid parafoveal targets. However, once word *n* + 1 was foveated, there was a delayed response to the previously parafoveally invalid target, suggesting that some information about the orthographic identity of word *n* + 1 was processed in parafoveal vision, but only assessed at a delay. This exploratory analysis largely replicates the analyses reported above.

In summary, the ERP findings in Experiment 2 replicated the finding of disrupted parafoveal processing as foveal semantic load increased, which was found in Experiment 1 utilizing eye tracking. Moreover, Experiment 2 extended findings in Experiment 1 by shedding light on the underlying attentional constraints induced by foveal load. We found a large negativity to invalid parafoveal words relative to valid parafoveal words, an effect that was modulated by foveal semantic load. This parafoveal negativity was largest when *n* was expected, and thus foveal load was low, and was systematically reduced as semantic foveal load increased. When *n* was unexpected but plausible, the parafoveal negativity was reduced, and when *n* was incongruent, no parafoveal negativity was seen within the epoch corresponding to parafoveal processing of word *n* + 1. However, a later effect was observed once word *n* + 1 appeared in foveal vision, which was dependent upon prior parafoveal status (valid vs. invalid). This effect occurred regardless of foveal load, even appearing when word *n* was incongruent, and thus foveal load was greatest. This contrast corresponds to a time period in which the foveal and parafoveal conditions were physically identical, so that effects in this time period were driven by the prior lexical status of word *n* + 1, before it appeared at central fixation. Importantly, this effect was observed in the incongruent condition as well, despite the lack of evidence that valid and invalid parafoveal previews were processed differently in this condition.

General Discussion

We investigated the impact of semantic foveal load on subsequent parafoveal word processing in two experiments, one using eye tracking and another using ERP methods. Both experiments revealed clear influences of foveal semantic load on the allocation of attention to the word to the right of fixation, broadly consistent with foveal-load theories of attentional control during reading (cf. Henderson & Ferreira, 1990; Payne & Stine-Morrow, 2012; White et al., 2005). When a foveated word was highly expected/predictable based on the prior context, foveal load was low, and thus parafoveal processing of the word to the right of fixation was robust in both experiments. Experiment 1 showed a large parafoveal preview benefit (derived from fixation durations on word *n* + 1) when foveal load was low, and Experiment 2 revealed that orthographically invalid targets presented in parafoveal vision elicited a widespread and sustained negative potential (with an anterior maximum) relative to valid parafoveal previews.

Importantly, both experiments showed that the degree of parafoveal processing was modulated by foveal semantic load in a graded manner, revealing evidence of message-level semantic foveal load effects. In Experiment 1, the parafoveal preview benefit observed on *n* + 1 was reduced when word *n* was unexpected but plausible, and eliminated in first-pass measures when word *n* was incongruent with the prior semantic context. Experiment 2, which used ERPs to monitor parafoveal processing in real time, showed a corresponding pattern (see Figure 2). The negativity observed for invalid parafoveal previews was reduced in magnitude when the foveal word was unexpected, and was absent when the foveal word was incongruent (Figures 1–3). Collectively, the findings in the current set of experiments have important implications for understanding the influence of higher level language processes (i.e., beyond the scope of lexical, visual, or occulomotor constraints) on the allocation of attention in normal reading, and for understanding the link between eye-movement control and electrophysiological indices of online sentence processing. These points are discussed in turn below.

Semantic Foveal Load: Implications for Models of Attentional Control in Reading

Across both eye-tracking and ERP studies of language processing, the time course of semantic context effects (i.e., the earliest point at which message-level context can exert influences on word processing), and their implications for the functional nature of semantic processing, have been widely debated (Hagoort, Baggio, & Willems, 2009; Kutas & Federmeier, 2011; Reichle et al., 2009; Smith & Levy, 2013). Prior eye-tracking work suggests an early role of context, with word predictability modulating early lexical stages of processing during reading (Reichle et al., 2003). However, the costs associated with processing words that are inconsistent with a prior context have been argued to occur at a "postlexical" integration stage, following lexical access (Abbott & Staub, 2015; Reichle et al., 2010). Others have suggested that the time course of detecting semantic incongruities is itself modulated to by the severity of the semantic violation (Warren, 2011), such that strong violations (e.g., selectional restriction violations, animacy violations, impossibility) can impact processing at a very early stage, whereas weaker plausibility violations show delayed influences, "spilling over" onto following words. The current findings suggest that one source of these observed spillover effects may be due to a reduction in parafoveal processing of $n + 1$ due to foveal load effects at word n (cf. Schroyens et al., 1999). This muddies the investigation of the time course of semantic context effects, because such effects of context can show spillover-like effects onto the following word simply because the amount of attention allocated to parafoveal words was differentially reduced in high foveal-load cases of semantic violations. Experiments monitoring ERPs in conjunction with eye tracking (either in separate experiments or through coregistration; Dimigen et al., 2011) may help to further elucidate the time course of semantic violations (see also Stites & Federmeier, 2015).

The E-Z Reader model of eye-movement control has adopted a serially staged architecture whereby lexical processing of a word must complete before the word can begin to be assessed relative to its syntactic and semantic context. Importantly, in the E-Z Reader model, shifts of attention to parafoveal words only begin after lexical processing of word n has completed (Reichle et al., 2003). In the most recent update to E-Z Reader, so-called postlexical integration processes are modeled as beginning following full lexical access of word n. However, shifts of attention to word $n + 1$ (initiating parafoveal processing) are argued to be triggered by the lexical access of word n (Reichle et al., 2009). Under such a model, with strict serial linguistic and attentional processing, traditional foveal-load effects can only be induced at a lexical stage of processing. Collectively, the findings from Experiments 1 and 2 show clear effects of semantic fit on the allocation of attention to word $n + 1$—effects that are graded by the severity of semantic violations, with the greatest costs to incongruent words. The additional data provided by ERPs show that the locus of foveal load effects indeed occurs during parafoveal processing of word $n + 1$, suggesting early influences of foveal semantic load. Thus, the current findings suggest that message-level semantic processing, including the detection and processing of semantically unexpected or incongruent words, is available to the comprehension system early enough to modulate attentional allocation to parafoveal word processing. For a serially staged lexical processing model like E-Z Reader to explain the current findings, the effects of semantic context would have to occur rapidly, influencing early stages of lexical processing of word n (prior to attentional shifts to $n + 1$), or perhaps immediately following the completion of lexical processing, such that attention is immediately withdrawn from the processing of word $n + 1$ when the system encounters foveal load from higher level language processes. However, this suggests a qualitative difference between lexical and postlexical sources of foveal-load effects—effects that are currently not considered in serial or parallel attention models of eye-movement control in reading.

Another framework in which to cast the current findings is through understanding the degree to which message-level semantic processing and visual attention allocation draw on similar resources during reading (cf. Caplan & Waters, 1999; Just & Varma, 2007). In the most recent version of E-Z Reader, the authors note that their model is agnostic with respect to "the issue of how to characterize the attention and/or memory resources used during lexical and postlexical processing; for example, is there one pool of resources divided between the two, or are there separate resources for each" (Reichle et al., 2009, p. 18). Although violations of semantic expectancy and semantic congruity have been shown to elicit different electrophysiological indices of costs following the N400 (DeLong et al., 2014; Federmeier et al., 2007; Van Petten & Luka, 2012), the functional role of these "post-N400-positivites" (Thornhill & Van Petten, 2012) in language processing have not been clearly explored. Some have argued that such post-N400 positivities may reflect the recruitment of additional attentional resources when encountering difficulty in semantic integration (Federmeier et al., 2007; Van Petten & Luka, 2012). For instance, the observation that the posterior P600 effect is strongly modulated by attention and task demands has led some to argue that it may be part of a larger family of P3b responses (Batterink & Neville, 2013; Coulson, King, & Kutas, 1998), a family that is well characterized by its sensitivity to attentional demands (see Nieuwenhuis, Aston-Jones, & Cohen, 2005, for a review). Moreover, some have speculated that expectancy and congruity violations recruit different attentional resources (DeLong et al., 2014; Federmeier et al., 2007). Our findings suggest that aspects of message-level semantic processing that are indexed by expectancy and congruity violations are demanding of central visual attention, inducing foveal load effects. Thus, the revision of message-level semantic representations appears to share at least some overlapping attentional resources with early covert attentional allocation in parafoveal vision, insofar as that the semantic fit of word n modulates the time course of attention allocation to early lexical processing of word $n + 1$.

The Time Course of Foveal Load Effects in Eye-Tracking and Event-Related Potentials

In Experiment 1, we observed a reduced parafoveal preview benefit to words immediately following semantically unexpected and incongruent words. However, from these findings alone, we could not determine what aspects of parafoveal processing were impacted by increases in semantic foveal load. Because the gaze-contingent boundary change paradigm relies on downstream effects at word $n + 1$ in order to infer parafoveal processing at word n, it was not previously clear if effects of foveal load directly influenced parafoveal processing of $n + 1$, or instead disrupted integration of parafoveal and foveal visual representations across successive saccades, which would also influence foveal processing of word $n + 1$. The temporal resolution of eye-tracking methods is not the limitation for inference on parafoveal processing, but instead the method is limited because inference must be made after parafoveal processing has already happened. In contrast, because ERPs provide a continuous moment-to-moment "in-line" measure of processing (Kutas & King, 1999), we were able to dissociate between effects of foveal load that operate at parafoveal perception, occurring during the foveation of word n, from later processes occurring during the foveation of $n + 1$, effects that additionally index integration of parafoveal and foveal visual representations.

By combining orthographically invalid targets (for which there has been prior evidence of early parafoveal detection; e.g., Drieghe et al., 2005; Kliegl et al., 2013), with ERPs, we could monitor the earliest point in time in which valid and orthographically illegal words were distinguished in parafoveal vision. Although we did not observe strong POF effects in the eye-tracking data from Experiment 1 (effects of the invalid preview only reached significance on word n for gd), we did observe a POF-like effect of the invalid preview in the ERPs when foveal load was low. The timing of this effect, onsetting between 250 ms and 300 ms, was such that it would likely spill over onto fixation durations of word $n + 1$ during natural reading (cf. Risse & Kliegl, 2014). Indeed, the timing of the foveal load effects are quite compatible with a delayed parafoveal perception account—the preview benefit effect on fixation durations to word $n + 1$ patterned directly with the nonword negativity effects seen following the presentation of word $n + 1$ in parafoveal vision. The ERP data thus suggest that the effects observed on fixation durations at $n + 1$ were driven by a reduction in parafoveal perception of the invalid target beginning at word n and spilling over onto the following word (see also Dambacher & Kliegl, 2007, for evidence of a lag-1 relationship between ERPs and eye-movement measures, and Ditman, Holcomb, & Kuperberg, 2007, for similar effects in self-paced reading). In sum, the effects of foveal load on parafoveal processing in reading may have multiple sources—partially derived from modulation of attention during parafoveal perception of word $n + 1$ and partially derived from the increased ease of integration of $n + 1$ in foveal vision when it was parafoveally valid versus invalid. However, using ERPs, we were able to dissociate between effects that began during the parafoveal perception of word $n + 1$ from effects that occurred once $n + 1$ was foveated (and parafoveal and foveal visual representations were subsequently integrated).

Once word $n + 1$ appeared in foveal vision, we observed a late relative positivity contingent on the prior parafoveal status of $n + 1$ in all three context conditions in Experiment 2. This was particularly striking because this effect occurred even though the physical stimuli were identical (within sentence context conditions). Thus, this effect appeared to be driven by the prior parafoveal status of $n + 1$. Importantly, this late relative positivity persisted even in the incongruent condition, presumably when foveal load was the greatest, even though the sustained negativity to invalid previews was absent in this condition. Therefore, despite the initial reduction in the parafoveal perceptual processing of word $n + 1$ in the high load (incongruent n) condition, there did not appear to be a complete loss of the parafoveal visual representation. Critically, this suggests that even when foveal load was maximal, the system gained some access to parafoveal information, but that increases in foveal load acted to delay the assessment of this information, rather than solely gating sensory processing of word $n + 1$ (see Figure 3).

Foveal load thus did not appear to completely gate the sensory processing of word $n + 1$, but instead may have induced a temporal bottleneck in which processing of word n completed before the system began to distinguish between the valid and invalid parafoveal visual representations. The timeline of this effect, occurring long after the 100-ms window in which the stimulus was actually presented in the parafovea, suggests that parafoveal word information may be buffered—perhaps in visual iconic/short-term memory (STM; cf. Luck & Vogel, 1997)—with an increasing delay in the assessment of this information as foveal load increases. Currently, models of eye-movement control in reading do not readily include mechanisms to account for such temporal bottleneck or visual STM-dependent effects of foveal load. For example, in E-Z Reader, word processing occurs in serial from word to word by shifting visual attention (a "spotlight" model of attention; cf. Posner, 1980) to parafoveal targets prior to their fixation and increased foveal load acts to delay the allocation of visual-spatial attention to word $n + 1$, such that less perceptual information is transferred to the visual system before its fixation (Reichle et al., 2003). Although we did observe that high foveal load eliminated the parafoveal negativity, consistent with reduced parafoveal perception with increasing load, some information about parafoveal orthographic validity was transmitted to the system during high foveal load. Under a strict serial visual attention foveal-load account, we should not have observed such an effect, as foveal load should have delayed visual attention from shifting to $n + 1$ (i.e., acting as a sensory gate).

Dimigen and colleagues (Dimigen et al., 2012; Li et al., 2015) have reported an effect they termed the "preview positivity," which is an occipitotemporal positivity peaking between 200 ms and 300 ms following the direct fixation of word $n + 1$ (i.e., following parafoveal processing of word $n + 1$) when that word had an identical parafoveal preview relative to an invalid but orthographically legal preview. It is not clear whether the late positive potentials observed in the current study to words with previously orthographically illegal previews are functionally related to Dimigen and colleagues' preview positivity effect, given the differing time course and the fact that the preview positivity has only been found via contrasts between identity previews and invalid, but orthographically legal, previews. One critical difference between our findings and those of Dimigen and colleagues is that the orthographically illegal previews in the current study induced an earlier negativity during parafoveal perception, preceding the fixation of word $n + 1$ (i.e., prior to the onset of the preview positivity). These findings thus indicate that, in addition to being able to discriminate between parafoveal and foveal representations of the same word (as indicated by the preview positivity; Dimigen et al., 2012), the system can also begin to detect orthographically implausible strings in parafoveal vision prior to their fixation (cf. Drieghe et al., 2005).

Although all three conditions showed some sign of the delayed positivity to (previously) invalid previews at $n + 1$, the results from the mass univariate analysis indicated that the timing and spatial distribution of these effects appeared to vary across conditions. Although this may suggest functionally different neural processes, three complicating factors obscure this interpretation. First, the preceding sustained negativity during the parafoveal processing of word n results in substantial component overlap with the positivity at word $n + 1$, an effect that differs across conditions with increasing foveal load. Second, overall differences in component amplitude can distort topological interpretations in terms of underlying neural generators (Urbach & Kutas, 2006). Lastly, the mass univariate approach is only powered to detect large effect sizes (like the parafoveal nonword negativity), and thus the true underlying time course and scalp topography of this delayed positivity may not be well characterized. Given that Experiment 2 was the first electrophysiological investigation into parafoveal processing of orthographically illegal words during reading, more work is needed to understand the functional significance of the reported parafoveal and foveal effects and their relationship to prior ERP work on parafoveal processing, including semantic processing of parafoveal targets (e.g., Barber et al., 2013) and ERP indices of the preview benefit during the fixation of word $n + 1$ (e.g., Dimigen et al., 2012).

Importantly, the current study demonstrated converging findings across ERP and eye-tracking studies of parafoveal word processing, providing cross-validating data for assumptions in both methodologies and bridging the gap between two historically divisive methodologies (cf. Sereno & Rayner, 2003). By conducting parallel experiments utilizing traditional eye-tracking methods and modified ERP methods, our findings provide the first direct evidence that foveal load effects are driven by reductions in early stages of parafoveal perception, with additional distinct later influences when word $n + 1$ was fixated. We take this pattern of findings as validating evidence for the assumption in boundary change eye-tracking studies on the parafoveal preview benefit that effects of foveal load may partially reduce the parafoveal perception of word $n + 1$ (as observed by the early effects of foveal load on ERPs when word $n + 1$ appeared in parafoveal vision), but may also additionally act to delay the assessment of the parafoveal visual representation of $n + 1$ until the integration of word n into its semantic context has completed.

One important criticism of the flanker ERP paradigm is that it is not naturalistic and, as such, that task-related demands may result in effects that are not associated with normal reading. Importantly, our findings provide confirmatory evidence that the flanker ERP paradigm yields evidence of parafoveal preview and foveal load effects similar to that seen during normal reading, substantiating the use of such RSVP paradigms for investigating parafoveal perception and uncovering the neural mechanisms underlying normal reading (cf. Li et al., 2015). Development of a valid method for examining ERP indices of parafoveal word perception in reading is necessary, as the majority of ERP research on visual language processing has been conducted using single-word RSVP paradigms, precluding the ability to examine many interesting aspects of the reading process. Indeed, efforts to bridge the theoretical and conceptual gaps between eye movement and neurophysiological investigations of reading have pointed to a critical role for parafoveal processing (see Reichle & Reingold, 2013, for a discussion).

Conclusion

As it is unlikely that all factors that drive foveal increases in processing time also modulate parafoveal processing (cf. Reingold & Rayner, 2006), understanding the full range of factors that modulate attentional control in reading has direct implications for theories of covert attention in models of reading. The findings in the current study indicate that ERP studies of parafoveal processing will be useful in providing converging evidence with eye-tracking methods, as well as providing novel information regarding the role of visuospatial and nonspatial attentional mechanisms in reading. Indeed, although the foveal-load theory has been widely accepted in the eye-movement literature, more work is needed to understand both the nature of attention constraints induced by foveal load as well the foveal factors that influence parafoveal processing, in order to delineate the complex foveal–parafoveal dynamics of reading.

References

Abbott, M. J., & Staub, A. (2015). The effect of plausibility on eye movements in reading: Testing E-Z Reader's null predictions. *Journal of Memory and Language, 85,* 76–87. http://dx.doi.org/10.1016/j.jml.2015.07.002

Barber, H. A., Ben-Zvi, S., Bentin, S., & Kutas, M. (2011). Parafoveal perception during sentence reading? An ERP paradigm using rapid serial visual presentation (RSVP) with flankers. *Psychophysiology, 48,* 523–531. http://dx.doi.org/10.1111/j.1469-8986.2010.01082.x

Barber, H. A., Doñamayor, N., Kutas, M., & Münte, T. (2010). Parafoveal N400 effect during sentence reading. *Neuroscience Letters, 479,* 152–156. http://dx.doi.org/10.1016/j.neulet.2010.05.053

Barber, H. A., van der Meij, M., & Kutas, M. (2013). An electrophysiological analysis of contextual and temporal constraints on parafoveal word processing. *Psychophysiology, 50,* 48–59. http://dx.doi.org/10.1111/j.1469-8986.2012.01489.x

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255–278.

Batterink, L., & Neville, H. J. (2013). The human brain processes syntax in the absence of conscious awareness. *The Journal of Neuroscience, 33,* 8528–8533.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B: Methodological, 57,* 289–300.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22,* 77–94. http://dx.doi.org/10.1017/S0140525X99001788

Clifton, C., Jr., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. Van Gompel, M. Fisher, W. Murray, & R. L. Hill (Eds.), *Eye movement research: A window on mind and brain* (pp. 341–372). Oxford, England: Elsevier Ltd.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13,* 21–58.

Dambacher, M., & Kliegl, R. (2007). Synchronizing timelines: Relations between fixation durations and N400 amplitudes during sentence reading. *Brain Research, 1155,* 147–162. http://dx.doi.org/10.1016/j.brainres.2007.04.027

Davies, M. (2008). *The Corpus of Contemporary American English: 450 Million Words, 1990–present*. Retrieved from http://corpus.byu.edu/coca/

DeLong, K. A., Quante, L., & Kutas, M. (2014). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia, 61,* 150–162. http://dx.doi.org/10.1016/j.neuropsychologia.2014.06.016

DeLong, K. A., Troyer, M., & Kutas, M. (2014). Pre-processing in sentence comprehension: Sensitivity to likely upcoming meaning and structure. *Language and Linguistics Compass, 8,* 631–645. http://dx.doi.org/10.1111/lnc3.12093

Dimigen, O., Kliegl, R., & Sommer, W. (2012). Trans-saccadic parafoveal preview benefits in fluent reading: A study with fixation-related brain potentials. *NeuroImage, 62,* 381–393. http://dx.doi.org/10.1016/j.neuroimage.2012.04.006

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: Analyses and review. *Journal of Experimental Psychology: General, 140,* 552–572. http://dx.doi.org/10.1037/a0023885

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology, 44,* 927–935. http://dx.doi.org/10.1111/j.1469-8986.2007.00593.x

Drieghe, D. (2011). Parafoveal-on-foveal effects in eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *Oxford handbook on eye movements* (pp. 839–855). Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199539789.013.0046

Drieghe, D., Brysbaert, M., & Desmet, T. (2005). Parafoveal-on-foveal effects on eye movements in text reading: Does an extra space make a difference? *Vision Research, 45,* 1693–1706. http://dx.doi.org/10.1016/j.visres.2005.01.010

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 954–969. http://dx.doi.org/10.1037/0096-1523.31.5.954

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41,* 469–495. http://dx.doi.org/10.1006/jmla.1999.2660

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language, 115,* 149–161.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research, 1146,* 75–84. http://dx.doi.org/10.1016/j.brainres.2006.06.101

Fischler, I., & Bloom, P. A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning & Verbal Behavior, 18,* 1–20. http://dx.doi.org/10.1016/S0022-5371(79)90534-6

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011a). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48,* 1711–1725. http://dx.doi.org/10.1111/j.1469-8986.2011.01273.x

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011b). Mass univariate analysis of event-related brain potentials/fields II: Simulation studies. *Psychophysiology, 48,* 1726–1737. http://dx.doi.org/10.1111/j.1469-8986.2011.01272.x

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (4th ed., pp. 819–836). Boston, MA: MIT Press.

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. http://dx.doi.org/10.1037/0278-7393.16.3.417

Holcomb, P. J., & Grainger, J. (2006). On the time course of visual word recognition: An event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience, 18,* 1631–1643. http://dx.doi.org/10.1162/jocn.2006.18.10.1631

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. http://dx.doi.org/10.3758/BF03195844

Inhoff, A. W., Starr, M., & Shindler, K. L. (2000). Is the processing of words during eye fixations in reading strictly serial? *Perception & Psychophysics, 62,* 1474–1484. http://dx.doi.org/10.3758/BF03212147

Just, M. A., & Varma, S. (2007). The organization of thinking: What functional brain imaging reveals about the neuroarchitecture of complex cognition. *Cognitive, Affective & Behavioral Neuroscience, 7,* 153–191. http://dx.doi.org/10.3758/CABN.7.3.153

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience, 17,* 4302–4311.

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. http://dx.doi.org/10.1037/0278-7393.21.1.68

Kliegl, R., Hohenstein, S., Yan, M., & McDonald, S. A. (2013). How preview space/time translates into preview cost/benefit for fixation durations during reading. *Quarterly Journal of Experimental Psychology (2006), 66,* 581–600. http://dx.doi.org/10.1080/17470218.2012.658073

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. http://dx.doi.org/10.1037/0096-1523.33.5.1250

Kornrumpf, B., & Sommer, W. (2015). Modulation of the attentional span by foveal and parafoveal task load: An ERP study using attentional probes. *Psychophysiology, 52,* 1218–1227. http://dx.doi.org/10.1111/psyp.12448

Kretzschmar, F., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). Parafoveal versus foveal N400s dissociate spreading activation from contextual fit. *NeuroReport: For Rapid Communication of Neuroscience Research, 20,* 1613–1618. http://dx.doi.org/10.1097/WNR.0b013e328332c4f4

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience, 12,* 535–540. http://dx.doi.org/10.1038/nn.2303

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4,* 463–470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology, 62,* 621–647. http://dx.doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203–205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307,* 161–163. http://dx.doi.org/10.1038/307161a0

Kutas, M., & King, J. W. (1999). In-line measures of syntactic processing using event-related brain potentials. *Behavioral and Brain Sciences, 22,* 104–105. http://dx.doi.org/10.1017/S0140525X99331782

Kutas, M., Lindamood, T. E., & Hillyard, S. A. (1984). Word expectancy and event-related brain potentials during sentence processing. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes* (pp. 217–237). Hillsdale, NJ: Erlbaum.

Laszlo, S., & Federmeier, K. D. (2011). The N400 as a snapshot of interactive processing: Evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology, 48,* 176–186. http://dx.doi.org/10.1111/j.1469-8986.2010.01058.x

Laszlo, S., & Federmeier, K. D. (2014). Never seem to find the time: Evaluating the physiological time course of visual word recognition with regression analysis of single item ERPs. *Language, Cognition and Neuroscience, 29,* 642–661. http://dx.doi.org/10.1080/01690965.2013.866259

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience, 9,* 920–933. http://dx.doi.org/10.1038/nrn2532

Li, N., Niefind, F., Wang, S., Sommer, W., & Dimigen, O. (2015). Parafoveal processing in reading Chinese sentences: Evidence from event-related brain potentials. *Psychophysiology, 52,* 1361–1374. http://dx.doi.org/10.1111/psyp.12502

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique* (2nd ed.). Cambridge, MA: MIT Press.

Luck, S. J., & Hillyard, S. A. (1990). Electrophysiological evidence for parallel and serial processing during visual search. *Perception & Psychophysics, 48,* 603–617. http://dx.doi.org/10.3758/BF03211606

Luck, S. J. (2012). Electrophysiological correlates of the focusing of attention within complex visual scenes: N2pc and related ERP components. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford handbook of event-related potential components* (pp. 329–360). New York, NY: Oxford University Press.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–281. http://dx.doi.org/10.1038/36846

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology, 32,* 4–18. http://dx.doi.org/10.1111/j.1469-8986.1995.tb03400.x

Matsuki, K., Chow, T., Hare, M., Elman, J. L., Scheepers, C., & McRae, K. (2011). Event-based plausibility immediately influences on-line language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 913–934. http://dx.doi.org/10.1037/a0022964

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17,* 578–586. http://dx.doi.org/10.3758/BF03203972

Müller, M. M., Malinowski, P., Gruber, T., & Hillyard, S. A. (2003). Sustained division of the attentional spotlight. *Nature, 424,* 309–312. http://dx.doi.org/10.1038/nature01812

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus–norepinephrine system. *Psychological Bulletin, 131,* 510–532. http://dx.doi.org/10.1037/0033-2909.131.4.510

Payne, B. R., Lee, C.-L., & Federmeier, K. D. (2015). Revisiting the incremental effects of context on word processing: Evidence from single-word event-related brain potentials. *Psychophysiology, 52,* 1456–1469. http://dx.doi.org/10.1111/psyp.12515

Payne, B. R., & Stine-Morrow, E. A. (2012). Aging, parafoveal preview, and semantic integration in sentence processing: Testing the cognitive workload of wrap-up. *Psychology and Aging, 27,* 638–649. http://dx.doi.org/10.1037/a0026540

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25. http://dx.doi.org/10.1080/00335558008248231

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. http://dx.doi.org/10.1080/09541440340000169

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81. http://dx.doi.org/10.1016/0010-0285(75)90005-5

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology (2006), 62,* 1457–1506. http://dx.doi.org/10.1080/17470210902816461

Rayner, K., Castelhano, M. S., & Yang, J. (2009). Eye movements and the perceptual span in older and younger readers. *Psychology and Aging, 24,* 755.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1290–1301. http://dx.doi.org/10.1037/0278-7393.30.6.1290

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3,* 504–509. http://dx.doi.org/10.3758/BF03214555

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. http://dx.doi.org/10.1037/0033-295X.105.1.125

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The EZ Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26,* 445–476. http://dx.doi.org/10.1017/S0140525X03000104

Reichle, E. D., & Reingold, E. M. (2013). Neurophysiological constraints on the eye-mind link. *Frontiers in Human Neuroscience, 7,* 1–6. http://dx.doi.org/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 & Review, 16,* 1–21. http://dx.doi.org/10.3758/PBR.16.1.1

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z Reader model. *Psychological Science, 17,* 742–746. http://dx.doi.org/10.1111/j.1467-9280.2006.01775.x

Risse, S., & Kliegl, R. (2014). Dissociating preview validity and preview difficulty in parafoveal processing of word n + 1 during reading. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 653–668. http://dx.doi.org/10.1037/a0034997

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics, 74,* 5–35. http://dx.doi.org/10.3758/s13414-011-0219-2

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 52,* 1021–1046. http://dx.doi.org/10.1080/713755859

Schwanenflugel, P. J., & LaCount, K. L. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 344–354. http://dx.doi.org/10.1037/0278-7393.14.2.344

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: Eye movements and event-related potentials. *Trends in Cognitive Sciences, 7,* 489–493. http://dx.doi.org/10.1016/j.tics.2003.09.010

Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition, 128,* 302–319. http://dx.doi.org/10.1016/j.cognition.2013.02.013

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass, 9,* 311–327. http://dx.doi.org/10.1111/lnc3.12151

Stites, M. C., & Federmeier, K. D. (2015). Subsequent to suppression: Downstream comprehension consequences of noun/verb ambiguity in natural reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1497–1515. http://dx.doi.org/10.1037/xlm0000119

Stites, M. C., Payne, B. R., & Federmeier, K. D. (2016). *Getting ahead of yourself: Parafoveal word expectancy modulates the N400 during sentence reading*. Manuscript submitted for publication.

Strimmer, K. (2008a). A unified approach to false discovery rate estimation. *BMC Bioinformatics, 9,* 303. http://dx.doi.org/10.1186/1471-2105-9-303

Strimmer, K. (2008b). fdrtool: A versatile R package for estimating local and tail area-based false discovery rates. *Bioinformatics (Oxford, England), 24,* 1461–1462. http://dx.doi.org/10.1093/bioinformatics/btn209

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

Thornhill, D. E., & Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology, 83,* 382–392. http://dx.doi.org/10.1016/j.ijpsycho.2011.12.007

Urbach, T. P., & Kutas, M. (2006) Interpreting event-related brain potential (ERP) distributions: Implications of baseline potentials and variability with application to amplitude normalization by vector scaling. *Biological Psychology, 72,* 333–343.

van den Brink, D., Brown, C., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience, 13,* 967–985. http://dx.doi.org/10.1162/089892901753165872

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology, 83,* 176–190. http://dx.doi.org/10.1016/j.ijpsycho.2011.09.015

Veldre, A., & Andrews, S. (2015). Parafoveal lexical activation depends on skilled reading proficiency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 586–595. http://dx.doi.org/10.1037/xlm0000039

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1656–1674. http://dx.doi.org/10.1037/0096-1523.24.6.1656

Warren, T. (2011). The influence of plausibility and anomaly on eye movements in reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford handbook on eye movements* (pp. 911–923). New York, NY: Oxford University Press.

Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review, 14,* 770–775. http://dx.doi.org/10.3758/BF03196835

White, S. J., & Liversedge, S. P. (2006). Foveal processing difficulty does not modulate non-foveal orthographic influences on fixation positions. *Vision Research, 46,* 426–437. http://dx.doi.org/10.1016/j.visres.2005.07.006

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 & Review, 12,* 891–896. http://dx.doi.org/10.3758/BF03196782

White, S. J., Warren, T., & Reichle, E. D. (2011). Parafoveal preview during reading: Effects of sentence position. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1221–1238. http://dx.doi.org/10.1037/a0022190

Wilson, M. (1988). MRC Psycholinguistic Database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments & Computers, 20,* 6–10. http://dx.doi.org/10.3758/BF03202594

Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *NeuroImage, 62,* 356–366. http://dx.doi.org/10.1016/j.neuroimage.2012.04.054

Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika, 70,* 203–212.

Received September 22, 2015
Revision received April 7, 2016
Accepted April 14, 2016 ■