AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

https://doi.org/10.1037/xge0001278

Are Eye Movements and EEG on the Same Page?: A Coregistration Study on Parafoveal Preview and Lexical Frequency

Sara Milligan[1], Martín Antúnez[2], Horacio A. Barber[2, 3, 4], and Elizabeth R. Schotter[1]

[1] Department of Psychology, University of South Florida
[2] Cognitive Department of Psychology, Universidad de La Laguna
[3] Institute for Biomedical Technologies (ITB), Universidad de La Laguna
[4] Basque Center on Cognition, Brain and Language (BCBL), San Sebastián, Spain

Readers extract visual and linguistic information not only from fixated words but also upcoming parafoveal words to introduce new input efficiently into the language processing pipeline. The lexical frequency of upcoming words and similarity with subsequent foveal information both influence the amount of time people spend once they fixate the word foveally. However, it is unclear from eye movements alone the extent to which parafoveal word processing, and the integration of that word with foveally obtained information, continues after saccade plans have been initiated. To investigate the underlying neural processes involved in word recognition after saccade planning, we coregistered electroencephalogram (EEG) and eye movements during a gaze-contingent display change paradigm. We orthogonally manipulated the frequency of the parafoveal and foveal words and measured *fixation related potentials* (*FRPs*) upon foveal fixation. Eye movements showed primarily an effect of preview frequency, suggesting that saccade planning is based on the familiarity of the parafoveal input. FRPs, on the other hand, demonstrated a disruption in downstream processing when parafoveal and foveal input differed, but only when the parafoveal word was high frequency. These findings demonstrate that lexical processing continues after the eyes have moved away from a word and that eye movements and FRPs provide distinct but complementary accounts about oculomotor behavior and neural processing that cannot be obtained from either method in isolation. Furthermore, these findings put constraints on models of reading by suggesting that lexical processes that occur before an eye movement program is initiated are qualitatively different from those that occur afterward.

Keywords: lexical frequency, parafoveal processing, FRPs, eye movements in reading

Taking in and making meaning of visual information is a complex process that involves recruiting knowledge about the world from previous experiences. This is particularly true when it comes to reading because the symbols that represent text are only meaningful in the context of a learned language system. The more often an individual encounters a particular visual word form and connects it with a particular meaning (i.e., when a word is high frequency), the stronger that connection becomes and the easier that word will be to recognize in the future. The speed at which visual word recognition during reading occurs is staggering and depends on efficient linguistic processing of the text across space and time, as well as optimal execution of eye movements to coordinate when each new word enters the processing pipeline. Therefore, a critical question is how, and at what stages of the reading process, word frequency guides oculomotor control and word recognition.

Recognition of a word can begin before the eyes even bring it into the center of vision (i.e., the *fovea*; see Schotter, 2018; Schotter et al., 2012), but the extent to which readers process a word based on *parafoveal* vision is not entirely clear. The effect of word frequency on parafoveal processing is an interesting test case because word frequency effects require recognition of the word, or at the very least processing far enough to perform a *familiarity check* (Reichle et al., 1998). Therefore, one question is whether a lexical property like word frequency can be processed parafoveally during natural reading and, if so, whether it impacts downstream processing of information obtained on the next fixation. Related is the question about whether the parafoveal lexical processing that occurs before saccade planning is qualitatively similar to processing on that word after the saccade plan has been initiated and may not be reflected in a reader's eye movement behavior.

Sara Milligan https://orcid.org/0000-0001-8600-2768

This study was partially funded by the Spanish government (FPI-MINECO Predoctoral Grant BES-2017-081797) and the society of Spanish scientists in United States (ECUSA; Fostering Grads mentorship program). Data from this study have been presented at the 2021 Psychonomic Society Annual Meeting and a departmental colloquium in the Department of Psychology at the University of South Florida. The data that support the findings of this study, the full sentence stimuli, and the code used for analyses, are available at https://osf.io/jkhvw/. The authors declare no conflicts of interest.

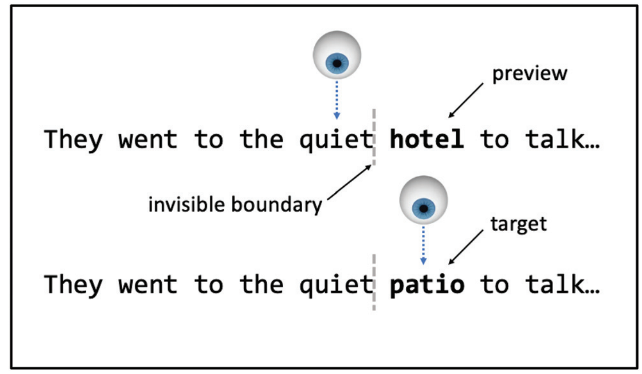
Correspondence concerning this article should be addressed to Sara Milligan, Department of Psychology, University of South Florida, 4202 East Fowler Avenue, PCD 4118G, Tampa, FL 33620, United States. Email: smilliga@usf.edu

The majority of empirical evidence about these questions comes from studies that track reader's eye movements. But while eye movements are ballistic (Schotter & Rayner, 2015), the cognitive processes underlying language comprehension unfold continuously over time (Barber & Kutas, 2007), so coordination of the two is not trivial. Although eye movements can reveal how the extraction of parafoveal information facilitates subsequent reading behavior (i.e., reduces reading times), using this measure alone may fail to capture further processing of the words that unfolds before or after an eye movement decision is initiated. Time-locking *fixation-related brain potentials* (*FRPs*) to these eye movements can further reveal how the neural processing of parafoveal information and integration of that information across saccades takes place in time. FRPs and their fixed-gaze equivalent, *event-related brain potentials* (*ERPs*), have exceptional temporal resolution and represent a time series rather than a single instantaneous event, making them particularly useful in mapping out the time course of cognitive processes that are made up of many subprocesses unfolding over time, as is the case for visual word recognition (Barber & Kutas, 2007). Therefore, FRPs provide insight into the unique stages of word recognition that occur after saccade plans are initiated that are much harder to investigate using eye movement measures. Coregistered measurements of neural activity during free reading can reveal whether (a) the language processing system can proceed based primarily on information that had been obtained parafoveally, (b) identification of words requires foveal input and begins only when a word is fixated, or (c) the language processing system retains information that had been obtained parafoveally to inform subsequent foveal processing.

The eyes can only fixate one word at a time, but while a reader looks directly at a given word, they are able to begin processing information about the upcoming word in parafoveal vision (see Rayner, 1998; Schotter et al., 2012). This preprocessing of the upcoming word contributes to reading efficiency, reducing the amount of time required to process the word once it is directly fixated, yielding a *parafoveal preview benefit* (see Schotter, 2018; Schotter et al., 2012).[1] Because parafoveal vision provides lower fidelity information about the text due to decreased acuity and attentional resources, a key question about the preview benefit is how much information readers are capable of extracting from parafoveal processing alone, and how that information is integrated as the reader's eyes move to actually land on the word.

The *gaze-contingent boundary paradigm* (Rayner, 1975; see Figure 1), is a flexible tool that has been used extensively to study parafoveal processing by experimentally manipulating what information is available to the reader in the parafovea and how the preview benefit varies as a consequence. In this eye tracking paradigm, an invisible boundary is placed before the target word of interest. While the eyes remain to the left of the boundary, the target word is replaced by a different preview. When the eyes cross to the right of the boundary to fixate the word, the display rapidly changes to reveal a target word in that location. Because one word is viewed only through parafoveal vision before fixating that location and a different word is viewed through foveal vision once it is fixated, the boundary paradigm is particularly useful for dissociating foveal and parafoveal processing of a word during reading. Traditionally, the parafoveal preview benefit is measured

Figure 1
Illustration of the Gaze-Contingent Boundary Display Change Paradigm (Rayner, 1975)



Note. This example demonstrates a display change condition. In traditional designs using this paradigm, the display change condition would be compared with a condition in which the preview and target are identical (i.e., no display change occurs when the eyes cross the boundary). See the online article for the color version of this figure.

as the reduction in reading time when the preview is identical to the target compared with when it is different. Within this framework, the characteristics of the preview have been manipulated to reveal preview benefits due to similarity between the preview and target based on orthography (Balota & Rayner, 1983; Balota et al., 1985; see Schotter et al., 2012), phonology (Pollatsek et al., 1992; see Leinenger, 2014), and semantics (Hohenstein & Kliegl, 2014; Schotter, 2013).

Based on initial work using the boundary paradigm, the idea of *trans-saccadic integration* between the preview and target was the prominent explanation for preview benefit effects for many decades (see Cutter et al., 2015). According to this account, information received across multiple fixations is merged or compared to form a singular representation, and the ability to form this representation is facilitated by similarity between the information obtained in these separate fixations (Pollatsek et al., 1992; Rayner, 1975). However, subsequent research revealed that the similarity between the preview and target may not be the sole factor determining how parafoveal information is used to facilitate reading because reductions in reading time are observed for completely unrelated preview words based on properties of the preview itself, such as word frequency (e.g., Risse & Kliegl, 2014; Schotter & Leinenger, 2016) and semantic plausibility (e.g., Schotter & Jia, 2016; Veldre & Andrews, 2016). Therefore, in these studies there is evidence for not only trans-saccadic integration effects, but also *preview difficulty effects* (e.g., based on preview frequency, see Schotter, 2018; and preview plausibility, see Andrews & Veldre, 2019), which suggest a more nuanced role for parafoveal processing, perhaps one that is qualitatively different for different stages of word identification. These distinctions may be better understood by integrating measurements of neuro-cognitive processing that

[1] Conversely, denying an accurate parafoveal preview by replacing the stimulus with something else can be considered a display change cost (Hutzler et al., 2019; Kliegl et al., 2013).

transpires between the initial fixation on a word and the saccade that ends that fixation.

Combining FRPs and eye movements can also provide insights about how lexical processing interacts with the oculomotor system during reading that have important implications for existing models of reading that separate stages of processing into pre- and postsaccade planning. For example, in the E-Z Reader model, the initial L_1 *stage* of processing completes after a familiarity check of a word that initiates the planning of the next saccade (Reichle et al., 2006; see Reichle & Sheridan, 2015). The subsequent stage, L_2, which has been conceptualized as the *lexical access* stage, is implemented as a proportion of L_1 processing that is influenced identically by lexical characteristics, like word frequency. So, if we find, in fact, that the processing that follows the familiarity check is qualitatively different, this would have important implications for a model like E-Z Reader, in which the relationship between the familiarity check and lexical access stages of processing would have to be reconceptualized. In contrast, OB1-Reader (Snell et al., 2018) proposes that the timing of saccade programming is only influenced by the lexical properties of a word if it is fully recognized while the eyes are fixating it. Therefore, it assumes that fixation durations and lexical access are influenced by only a singular stage of lexical processing. Once again, based on this model, we would expect to see the same effects of lexical frequency in fixation durations and downstream measures of word recognition after the eyes have moved forward. Therefore, if we find different patterns in early fixation durations and later effects in the FRPs, both of these models would need to adjust how they distinguish pre- and postsaccade lexical processing.

Frequency and Parafoveal Preview Effects in Eye Movements

Word frequency effects on fixation durations during reading have been consistently replicated (see Rayner, 2009). Words that are more frequent in the language are more frequently skipped rather than fixated (Rayner et al., 1996; White, 2008), and have shorter reading times when they are fixated (Inhoff & Rayner, 1986; Just & Carpenter, 1980; Rayner & Duffy, 1986). The skipping difference due to word frequency is observed even when the word does not make sense in the context, suggesting that eye movement decisions may be triggered based on initial word recognition processes, but not complete recognition or integration of that word (see Schotter, 2018). For example, Angele and Rayner (2013) used the boundary paradigm to replace three-letter content target words (e.g., *ace*) with the very high frequency preview word *the*; they found that the infelicitous *the* preview was skipped roughly half of the time (as frequently as a felicitous *the*), even when it did not make sense. This pattern extends to other three-letter content words, such that higher frequency words are skipped more often than low frequency words even when they are anomalous in the sentence context (Angele et al., 2014). By definition, skipping rates depend entirely on parafoveal processing because the skipping decision occurs before the word ever being fixated; therefore, frequency effects on skipping strongly suggest that the representational strength of a given word in the mental lexicon (Emmorey & Fromkin, 1988) influences how far the reader is able to get in word identification and this process can start when the word is perceived parafoveally. The fact that high frequency words were skipped at a higher rate than correct words, despite being anomalous, raises the question of whether complete recognition and integration processes are delayed until foveal information is obtained.

As noted before, an identical preview usually results in faster reading times compared with invalid preview conditions in which the display changes between the preview and target, presumably due to trans-saccadic integration. One exception to this pattern is when the different preview is a higher frequency word than the target word (Schotter et al., 2018, 2019; Schotter & Fennell, 2019; Schotter & Leinenger, 2016). Reading times on a low frequency target are longer after an identical parafoveal preview compared with a different preview condition in which the preview was a higher frequency word that makes sense in the sentence context. Schotter and Leinenger (2016) interpret this *reversed preview benefit* as the result of *forced fixations*, in which processing of the high frequency parafoveal preview reaches a threshold that triggers the preprogramming of eye movements forward from the upcoming word (in cases in which skipping of the word cannot be programmed because of a point-of-no-return in the saccade program toward that word). Therefore, it appears that eye movement planning *on a fixated target word* may depend more on the parafoveal processing that had occurred before that fixation rather than the foveal processing that occurs during that fixation.

Further evidence for an independent role of parafoveal processing on word recognition comes from findings that readers may fully identify and integrate the parafoveal word, even when they land on a different word. This is indicated by the readers' responses to comprehension questions about what word the sentence contained (Schotter et al., 2018) and regressions that they make out of subsequent sections of the sentence that render the preview word implausible (Schotter & Fennell, 2019; Schotter et al., 2019). For example, there are more regressions out of sentence regions after the target word in different preview conditions compared with identical preview conditions, and this effect is numerically larger for high frequency compared with low frequency previews (though the interaction was not statistically significant; Schotter et al., 2019). This pattern suggests that the transsaccadic integration failure (indicated by higher regression rates) after an invalid preview may be more likely when the preview was easy to identify and the reader had progressed further into processing it (Schotter et al., 2019). Together, these complex patterns suggest that initial reading time on a word may depend primarily on the ease of processing the preview whereas processing difficulty from a different preview may show up in the eye movement record after the reader has moved on from the target word. However, the fact that the initial ease of processing the preview modulates the magnitude of the cost of the different preview (Kliegl et al., 2013), suggests that word recognition does not completely start anew once new foveal information is encountered.

Schotter and colleagues argued that recognition and integration of the preview rather than the target results from attention having already shifted ahead in the sentence while the reader fixates the target during forced fixations. However, evidence of readers identifying the preview and ignoring the target only occurs on a subset of trials and it is unclear what the readers represent about the word when they encounter and attend to new, incompatible information upon fixating the target. Does identification of the target word begin anew only sometimes when it is fixated following an invalid preview? Or does the language processing system generally recognize the mismatch and attempt to process the new foveal linguistic information even though the oculomotor system has moved on?

Because fixation durations and regression rates are discrete measures derived from ballistic eye movements, eye tracking alone may not fully reveal how the integration of parafoveal and foveal information unfolds continuously over time and to what extent transsaccadic integration failure taxes the word recognition process as a whole. In the past, electroencephalography (EEG) and eye tracking have been used with great success to study the processes underlying visual word recognition and the overall reading process, but have been treated largely independently. The patterns of data produced by each approach reveal how various factors (e.g., sentence context, lexical frequency, and visual quality) influence the reading process, as measured by eye movement measures like fixation duration and skipping rates (see Rayner, 1998, 2009) and ERP components like the N400 (see Kutas & Federmeier, 2011). However, these measures have often been interpreted in isolation without considering how they align with one another. Therefore, the high temporal resolution of electrical brain responses from FRPs provides a useful tool for determining how parafoveal and foveal information are combined over time and how the relative difficulty of the parafoveal information influences this integration. Using FRPs not only allows us to measure electrical brain responses during natural reading in which the eyes are allowed to move freely, but also provides the opportunity to measure and analyze both eye tracking and EEG data simultaneously, allowing us to compare the patterns (and the resulting conclusions) between the two methodologies from exactly the same participants reading exactly the same sentences in exactly the same experiment. If the same patterns emerge from eye movements and EEG, we can conclude that the two measures are compatible, and perhaps even redundant, and that they measure the same aspect of the reading process (i.e., those that occur before vs. after a saccade is initiated). If on the other hand, we find that the eye movements and FRPs tell two different stories, it would suggest that each measure is tapping into separate, qualitatively different, components of that process. Therefore, FRPs have the potential to provide insights about processing that occurs after saccade planning that may not be readily apparent in the eye movement record if ocular behavior and lexical processing are decoupled further downstream.

Frequency Effects in EEG

ERPs provide a neural index of word identification difficulty that is well-suited to the questions raised above. For example, the N400 component has been proposed to reflect the process of accessing lexical information from long-term semantic memory (Kutas & Federmeier, 2011) and is modulated by both preprocessing of a target word in the parafovea (Antúnez et al., 2021; Barber et al., 2010, 2013, 2011; López-Pérez et al., 2016; Payne et al., 2019; Stites et al., 2017) and lexical frequency (Dambacher et al., 2006; Hauk & Pulvermüller, 2004; Rugg, 1990; Van Petten & Kutas, 1990). The N400 effect is a more negative-going deflection in the ERP waveform for difficult to identify words relative to easy to identify words, which occurs between 300 and 500 ms and peaks around 400 ms after the word is perceived. One conundrum with relying on the N400 as an index of lexical processing difficulty and word identification is that fixation durations on a word vary depending on word difficulty (e.g., lexical frequency), but the average fixation duration on a word is around 250 ms (Rayner, 1998), terminating before the canonical N400 time window even begins (Rayner & Clifton, 2009). Furthermore, for lexical frequency to influence fixation durations, the information must have been processed by the reading system before 250 ms; therefore, we might expect a lexical frequency effect to show up in the EEG record before 250 ms as well. Nevertheless, numerous studies have found the most robust frequency effects later, in the N400 component, in a variety of designs and tasks including single word reading and lexical decision tasks (Hauk & Pulvermüller, 2004; Rugg, 1990), and sentence reading with *rapid serial visual presentation* (*RSVP*; Dambacher et al., 2006; Van Petten & Kutas, 1990).

Although they have been somewhat less consistent than the frequency effects in the N400 time window, a number of studies have reported significant earlier effects of word frequency around 140–200 ms poststimulus onset, with some variability in scalp distribution (Dambacher et al., 2006; Hauk & Pulvermüller, 2004; Niefind & Dimigen, 2016; Sereno et al., 2003). Additionally, Laszlo and Federmeier (2014) performed a regression analysis on ERP responses to single words, predicting the amplitude across time by various lexical, orthographic, and semantic characteristics to identify the time course of different stages of lexical processing. They identified significant effects of word frequency, controlling for other characteristics, from 270 to 360 ms, slightly earlier than the typical peak of the N400.

Much of what we know about how word difficulty (i.e., frequency) manifests in ERP responses comes from foveal word processing, whether in single-word presentation paradigms (e.g., Hauk & Pulvermüller, 2004; Rugg, 1990) or RSVP sentence reading paradigms (Dambacher et al., 2006; Van Petten & Kutas, 1990). Therefore, less is known from ERPs about how lexical frequency influences parafoveal processing of an upcoming word before fixating it and how information processed during parafoveal preview influences foveal processing. Recently, however, the coregistration of eye movements and EEG has been used to isolate FRPs, effects time-locked to fixations on particular words during natural reading, in which parafoveal processing is possible. Niefind and Dimigen (2016) manipulated word frequency and whether a parafoveal preview was identical to or different from the fixated target word during word list reading. Their FRPs patterns demonstrated frequency effects in early time windows (140–200 and 200–300 ms), lining up with effects in single-word presentation. However, two coregistration studies that manipulated word frequency during sentence reading failed to find word frequency effects in FRPs (Degno et al., 2019; Kretzschmar et al., 2015). Degno et al. had a parafoveal manipulation in which they replaced the parafoveal word with X strings and illegal letter strings (as well as an identical preview condition) but then also manipulated the frequency of the target word. They found no effects of word frequency, even when timelocking to the foveal fixation in the identical preview condition. Kretzschmar et al. (2015) did not have a display change manipulation, but rather factorially manipulated word frequency and predictability in natural sentence reading, and they also did not find a main effect of word frequency time-locked to the foveal fixation. However, frequency effects are diminished as contextual support increases (Dambacher et al., 2006; Van Petten & Kutas, 1990), and Kretzschmar et al. (2015) did find a short-lived interaction between expectancy and frequency between 300 and 350 ms that indicated a frequency effect for low but not high expectancy words. Therefore, we should perhaps only expect frequency effects to appear in low constraint contexts when bottom-up word recognition processes may be more necessary.

Preview Effects in EEG

Coregistration studies have also allowed for the isolation of FRP effects related to parafoveal processing of lexical information in the boundary paradigm. A number of studies have found consistent effects of display changes (i.e., more positive amplitudes for identical compared with different previews) at 200–300 ms after fixation on the target at occipitotemporal scalp locations, as well as in the typical N400 time window and scalp location (Degno et al., 2019; Kornrumpf et al., 2016; Li et al., 2015; Niefind & Dimigen, 2016). Dimigen and colleagues (Dimigen et al., 2012) referred to this as a "preview positivity," but it could also be conceptualized as a *display change negativity* (i.e., more negative amplitudes for different compared with identical previews), demonstrating an increase in processing difficulty when processing that occurred before direct fixation is incompatible with foveal word recognition. It remains to be seen how these processing difficulty effects at the neural level (e.g., N400 effects in FRPs) relate to the patterns of effects observed in the eye movement record (e.g., fixation durations or the probability of making a regression after the word is read).

There are still many open questions about the relationship between reading processes that are reflected in eye movements and those that are reflected in neural measures such as FRPs. However, one study indicates that it may be possible to find links between these two measures. Metzner et al. (2017) recorded FRPs while participants read sentences with syntactic or semantic anomalies that usually generate P600 or N400 effects in RSVP-based ERP studies. They split trials into those in which the reader made a regression and those in which the reader did not make a regression and investigated the FRPs under these two scenarios. They found that these canonical ERP effects were more likely and larger in magnitude in cases in which readers made a regression, presumably because that meant the readers had noticed the linguistic anomalies as opposed to engaged in "good enough" processing (Ferreira & Patson, 2007). Therefore, in our study we may expect that the measure in the eye tracking record that would most closely align with the FRP effects reflecting processing difficulty (e.g., display change effects) would be regressive saccade probability rather than earlier fixation duration measures.

The Current Study

In the current study, we investigate whether the difficulty of processing a parafoveal preview changes the weighting of parafoveal and foveal information during the word recognition process in natural sentence reading. The degree of disruption when the parafoveal preview differs from the fixated target, compared with when it remains the same, would reveal how much relative influence the parafoveal preview has on word recognition. Furthermore, the extent to which differing target and preview words disrupts processing may be contingent on how easy or difficult a parafoveal word is to recognize. The interplay between parafoveal processing, saccade planning, and integration of parafoveal information with lexical information obtained during direct fixation has important implications for how word recognition unfolds over time and across space during natural reading. We make a novel contribution to this question by measuring eye movements and EEG simultaneously while participants read sentences in the boundary paradigm that factorially crossed the relationship between the preview and target (identical vs. different preview) and lexical frequency (high vs. low; see Schotter & Leinenger, 2016). In the eye movement record, we hypothesized to replicate prior findings of a *standard preview benefit* (i.e., longer fixation durations on targets following different previews compared with identical previews) for high frequency target words, but a *reversed preview benefit* (i.e., longer fixation durations on targets following different previews compared with identical previews) for low frequency targets (see Schotter & Leinenger, 2016). We also hypothesized there would be more regressions backward after the reader had left the target word in display change trials compared with identical preview trials, and an interaction whereby this effect would be more pronounced for high frequency previews (cf. Schotter et al., 2019).

We hypothesized several FRP effects based on those observed in the time windows and regions of interest (ROIs) described by Niefind and Dimigen (2016), who used the same design as the current study, except that they manipulated word frequency and preview validity on a target word in a word list, while we did so in full sentences. At occipital electrode sites at 140–200 ms after fixation on the target, we expected a preview frequency effect whereby amplitudes would be more negative for low compared with high frequency words. At occipital electrode sites at 200–300 ms after fixation on the target, we expected a display change effect whereby amplitudes would be more negative for the different preview conditions compared with the identical preview conditions. In addition, in the N400 ROI, at centro-parietal electrode sites at 300–500 ms after fixation on the target, we expected both a display change effect whereby amplitudes would be more negative for different previews compared with identical previews (Degno et al., 2019; Kornrumpf et al., 2016; Li et al., 2015; López-Perez et al., 2016), as well as a preview frequency effect whereby amplitudes would be more negative for low frequency compared with high frequency previews (Dambacher et al., 2006; Rugg, 1990; Van Petten & Kutas, 1990).

With respect to our question about the relationship between oculomotor control and neural processes reflected by FRPs, we assume that regressions triggered after a word is read may be related to processing difficulty that would be reflected in the N400 effect in FRPs. Therefore, we hypothesized to find an interaction between preview frequency and the display change effect in the N400 ROI, which would resemble the numerical interaction in regressions after the target word observed by Schotter et al. (2019); the display change effect (i.e., more negative amplitude for different compared with identical previews) would be larger when the preview was high frequency compared with when the preview was low frequency.

Method

Participants

Fifty-nine undergraduate students from the University of South Florida Psychology Department participated in the study for course credit. This study was approved by the University of South Florida Institutional Review Board under Pro00042067, "Psychophysiology, Eye Movements, and Cognition," and all participants provided electronic consent to participate. All participants were right-handed native English speakers with normal or corrected-to-normal vision and no history of reading, learning, or neurological disorders. Forty-five participants were included in the analyses; 14 were excluded because fewer than 30 total trials (across all conditions) remained after exclusions due to multiple first-pass fixations, early display changes, and EEG artifacts (see Data Processing for details). All of the remaining participants had at least 38 total trials. Data was collected over the course of a 3-month funded research stay and the equipment setup was temporary during that time. Therefore, it was determined ahead of time that we would collect data from as many participants as possible during that period. Of the previous FRP studies with display change and frequency manipulations, Degno et al. (2019) had the largest sample size at 42 participants, so we planned to meet or exceed this number for comparability and believed this to be possible in the 3-month period based on our lab's previous rates of data collection.

Stimuli and Design

One hundred forty-four stimuli were taken from Schotter and Leinenger (2016). Each stimulus contained a low constraint, high plausibility sentence frame with sentence-medial target word pairs that were either high or low frequency and matched on length (M = 5.88; range = 4–9 characters), part of speech (half of the pairs were nouns and half were verbs), orthographic neighborhood size, concreteness, and semantic diversity (see Table 1 for lexical characteristics retrieved from the English Lexicon Project; Balota et al., 2007). The design was a 2 × 2 factorial, crossing preview frequency (High vs. Low) and display type (Identical vs. Different preview and target words). For trials with different previews, the preview frequency was the opposite of the target frequency (see example stimuli below). If the target word was high frequency (1a), the different preview was low frequency, and if the target word was low frequency (1b), the different preview was high frequency. These words were embedded in low constraint sentences to maximize the opportunity to observe frequency-related FRPs (see Kretzschmar et al., 2015).

1a. They went to the quiet (hotel/patio) **hotel** to talk by themselves.

1b. They went to the quiet (patio/hotel) **patio** to talk by themselves.

Table 1
Summary Statistics of Target Word Lexical/Semantic Characteristics and Normative Data by Condition

Frequency condition	Log HAL frequency/ 400M	Orthographic neighborhood size	Concreteness rating	Semantic diversity
High	10.50 (1.05)	3.65 (4.05)	3.66 (1.02)	1.81 (0.24)
Low	6.81 (1.20)	2.88 (3.43)	3.79 (1.05)	1.51 (0.26)

Note. Values are the means of each variable with standard deviations in parentheses.

Thirteen sentences were modified to make the pretarget word at least four characters in length to increase the likelihood that it would be fixated (because this was a criterion for a trial to be included in the analysis). Each participant saw all items, counterbalanced across the four preview-target word frequency combination conditions, resulting in each participant seeing 36 items per condition.

Normative data on the plausibility of the sentences and predictability of the target words in the sentence contexts were collected on a separate sample of participants from the same participant pool (see Table 2). The plausibility of each sentence was computed as the mean rating on a Likert scale (1 = *very poorly written*, 7 = *very well written*) across 10 participants. The predictability of each target word was calculated as the proportion of participants out of 10 who provided the word as a continuation in a cloze task (i.e., fill-in-the-blank; Taylor, 1953). These data were collected to confirm that the target words were equally plausible in both frequency conditions and that the target word was not predictable. The sample size is common for norming tasks like these and is justifiable because these tasks are used as a manipulation check and this is not a primary analysis for the study.

Apparatus and Recording

EEG was recorded from 27 Ag/AgCl passive electrodes embedded in an Easycap extended 10/20-system (see Figure 2) and amplified using BrainVision BrainAmp with a 500 Hz sampling rate and a .01–100 Hz band pass filter. Horizontal and vertical electrooculogram (EOG) was recorded from two additional pairs of electrodes placed on the outer canthi of each eye and above and below the left eye. The signal was referenced online to the left mastoid and rereferenced offline to the algebraic mean of the right and left mastoids. Impedance values were reduced to 5 kΩ or lower at all electrode sites before recording.

Eye movements were recorded using an SR Research Ltd. Eyelink 1000 Plus eye tracking camera (sampling rate of 1000 Hz). Viewing was binocular, but eye movements were recorded only from the right eye. A 5-point calibration was performed at the beginning of the experiment and calibration accuracy had to fall within .3° of visual angle at each point to be accepted. Recalibration was also performed periodically throughout the experiment if accuracy dropped below this level. In the Eyelink recording settings, saccade detection was set to *Normal* (recommended for cognitive tasks like reading). The sample filter was set to *Extra* and the link analog filter was set to *STD* (both the recommended default settings; SR Research Ltd., 2009).

Procedure

Participants were seated at a viewing distance of 60 cm from a HP p1230 CRT monitor, with a refresh rate of 150 Hz and a screen resolution of 1024 × 768 pixels. Text was displayed in black Courier New 14 font on a white background on one line in the vertical center of the screen so that 2.52 characters subtended 1° of visual angle. At the beginning of the experiment, participants were instructed to read sentences normally for content to answer comprehension questions about them. They were given five practice trials to acclimate them to the task. Stimuli from this experiment were intermixed with 126 sentences and 30 comprehension questions from another experiment (see Antúnez et al., 2022).

Table 2
Cloze Probability and Plausibility Normative Data

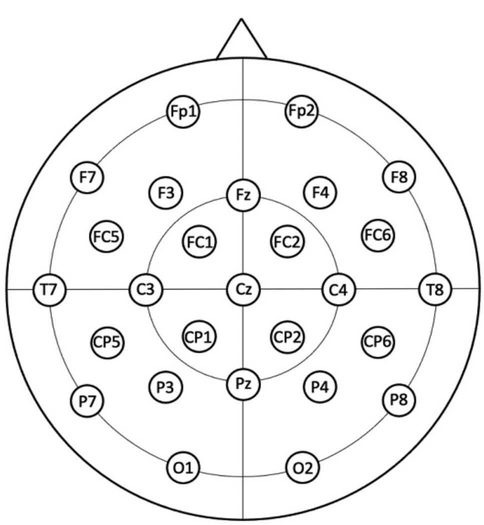
Frequency condition	Cloze probability (proportion)	Plausibility rating (1−7 Likert scale)
High	0.03 (0.08)	4.44 (0.83)
Low	0.01 (0.02)	4.51 (0.87)

Note. Values are the means of scores for each measure with standard deviations in parentheses.

Following this experimental procedure, another reading task was performed and measures of spelling ability were collected. Those data were not analyzed for the purpose of this study and are not reported here. The entire experimental session took 90 min, including setup.

At the beginning of each trial, a fixation point appeared in the center of the screen and calibration accuracy was checked using the single centered drift check point. If the calibration was accurate (error within 0.3° of visual angle), the trial was initiated by the experimenter. To trigger the presentation of the sentence, the participant had to make a fixation in a black box on the left side of the screen at the location of the beginning of the sentence. The participant silently read the sentence at their own pace and looked at a bullseye off to the right of the screen when they were done, pressing a button to move forward. An invisible boundary was located at the beginning of the space before the preview or target word. While the participant fixated to the left of the boundary, the preview word was visible, and once their eyes crossed the boundary the target word was revealed. The average time delay between the boundary being triggered and the screen changing was 4.31 ms (*SD* = 1.90). Participants were asked after the experiment whether they noticed anything unusual about the sentence display. If they reported noticing words flickering or anything odd about the display we followed up and asked if they noticed words changing. No participants were able to report specific instances of seeing a word change or recognizing that the word they landed on was different. Comprehension questions were presented after 40 (27.8%) of the experimental trials and participants responded yes or no by pressing one of two buttons on a response pad. Comprehension accuracy was high (*M* = 92%) and all participants produced accuracy scores of at least 80%.

Figure 2
Easycap 10-20 System 27-Electrode Montage



Data Processing

Eye tracking data processing was performed in SR Research Data Viewer. Fixations shorter than 80 ms within one character of another fixation were combined; fixations shorter than 80 ms that were not combined and fixations longer than 800 ms were excluded. The EEG and eye tracking data streams were synchronized online using parallel port triggers sent at the start of every fixation. The triggers had a consistent 50 ms delay from the onset of a fixation, so they were corrected by shifting them back 50 ms in offline processing. All analyses were based on trials with exactly one first-pass fixation on the target word and at least one fixation on the pretarget word. Therefore, trials in which the pretarget word was skipped, the target word was skipped, or the target word was fixated multiple times on the first pass were excluded, leaving 3,041 trials (46.9%) available for analysis. We chose to only include single fixations because of the potential for variability between cognitive processes time-locked to a single fixation and the first (or subsequent) of multiple fixations. For example, it has been suggested that refixations may be due to incomplete lexical processing (Reingold et al., 2010) or suboptimal saccade targeting (Schotter & Leinenger, 2016). Therefore, the nature and time course of the underlying word processing could conceivably be quite different between single and first of multiple fixations. Although selecting single fixation trials includes less data, the cognitive processes are likely to be more consistent so power can be maintained through smaller variance despite the smaller number of observations.

We also excluded 41 trials (1.3%, leaving 3,000 retained) when there was an early display change (i.e., the display change was triggered by blinking that occurred before the display change, even when a fixation was not actually made to the right of the boundary); these trials were detected by manual inspection of the data.

EEG data processing was performed in the EEGLAB (v2019.0; Delorme & Makeig, 2004), ERPLAB (v8.02; Lopez-Calderon & Luck, 2014), and EYE-EEG (v0.85; Dimigen et al., 2011) toolboxes in Matlab. The data were rereferenced offline to the average of the right and left mastoids and band-pass filtered from .1–50 Hz (−6 dB), with .2–32.8 Hz half-power (−3 dB) cutoffs, using an IIR Butterworth filter. Ocular artifacts were removed from the EEG using optimized independent components analysis (OPTICAT, Version 2020-01-28), following the procedures and recommendations described in Dimigen (2020). The ICA was trained using band-pass filtered (with a passband edge of 3 Hz) training data that overweighted spike potentials by a factor of 1. Ocular artifact components were automatically flagged and removed using eye tracker-guided eye artifact component identification (Plöchl et al., 2012), using a variance ratio threshold of 1.1. EEG was epoched into segments from 200 ms before to 1,000 ms after the start of fixations on the target word and baseline corrected by

Table 3
Result of Linear Mixed Effects Regression Predicting Single Fixation Duration by Preview Frequency and Preview Validity

	Primary analysis				Follow-up (high frequency)				Follow-up (low frequency)			
Predictors	Est.	*SE*	*t*	*p*	Est.	*SE*	*t*	*p*	Est.	*SE*	*t*	*p*
(Intercept)	261.8	7.47	35.02	**<.001**	248.89	6.97	35.73	**<.001**	274.22	8.75	31.33	**<.001**
Preview frequency (identical)	26.42	5.12	5.16	**<.001**								
Display change	3.31	3.25	1.02	.309	14.22	4.67	3.04	**.002**	−6.57	4.86	−1.35	.177
Display Change × Preview Frequency	−21.92	7.00	−3.13	**.002**								

Note. Both the primary and secondary analyses produced singular fits. Random effects variance and correlation estimates are presented in Table B1 of Appendix B. Bold font indicates statistical significance at the .05 α-level.

subtracting the mean voltage from −200 to 0 ms for each channel.[2]

Epochs containing artifacts were flagged for removal using a moving window peak-to-peak threshold automatic artifact detection algorithm, rejecting epochs with voltage changes of greater than 100 µV within a 200 ms time span. The epoched data was also inspected manually to confirm that artifact-contaminated epochs were removed, resulting in 68 total trials (2.3% of 3000) being excluded due to EEG artifacts.

The resulting dependent variables from the eye tracking data and FRP data were exported from their respective processing software and were merged on a trial-level for confirmatory analyses in R. After all exclusions based on fixation behavior, early display changes, and EEG artifacts, 2,932 trials (45% out of 6480 total trials) were included in the analyses. By condition, this left 777 trials for High Frequency Identical Preview, 705 trials for High Frequency Different Preview, 720 trials for the Low Frequency Identical Preview, and 730 trials for the Low Frequency Different Preview. On average, participants had 16.29 trials retained per condition.[3]

Transparency and Openness

The processed data on which these analyses were performed, the R code for waveform plots and confirmatory analyses, MATLAB code for exploratory analyses, and the experimental sentence stimuli can be found on Open Science Framework (OSF) at https://osf.io/jkhvw/. We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study (Simmons et al., 2012). No components of this study were preregistered.

Results

Eye Movements

We focused our analyses on trials with single fixations on the target word to allow for direct comparison between patterns in eye movements and FRPs. We analyzed single fixation durations (SFD) on the target word and the probability of regressions out of posttarget regions to provide snapshots of both early and downstream effects of word frequency and display changes on eye movements. Supplementary analyses of later measures can be found in Appendix A (Table A1), as well as figures showing fixation duration patterns in gaze durations (Figure A1) and total time (Figure A2). Each dependent measure was analyzed using a separate linear mixed effects regression model using the lme4 package (Version 1.1–17; Bates, 2010), and *p*-values were estimated using the Satterthwaite approximation via the lmerTest package (Kuznetsova et al., 2017). Preview frequency (low vs. high) was entered as a treatment contrast with the identical condition as the baseline and display change (different vs. identical) was coded with centered (i.e., sum-to-zero) contrasts. For both the SFD and regression analyses, we used the maximal random effects structure with intercepts and random slopes for frequency, display type, and their interaction for both items and participants.

Single Fixation Duration

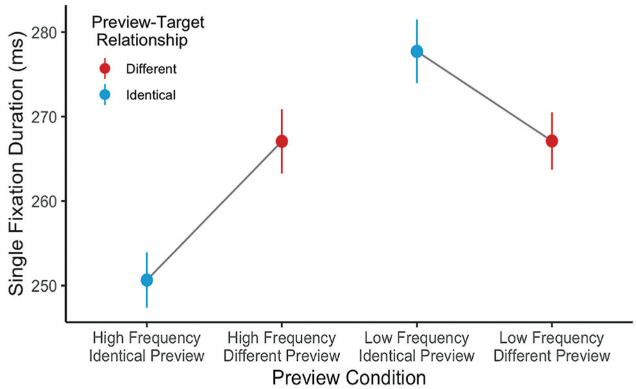
The SFD analysis replicated previous findings from the eye tracking literature. There was a significant main effect of preview frequency, replicating established frequency effects in eye movements, such that high frequency previews resulted in shorter fixations than low frequency previews (Rayner, 1998). We found no main effect of display change, but rather a significant interaction (Schotter & Leinenger, 2016), such that the display change effect was in the opposite direction between the two preview frequency conditions (see Table 3). To tease apart this interaction, we performed follow-up analyses predicting SFD by display type separately for high and low frequency previews (using the same contrasts as the primary analysis). When the preview was high frequency, the different preview condition led to significantly longer fixations (M = 266.87 ms, SD = 100.62) on the target than the identical previews (M = 250.26

[2] One drawback to the FRP technique is that the experimenter does not control how long the participant looks at the pretarget word. This raises the concern that any variability between conditions before the target fixation could result in condition differences during the baseline period, introducing spurious effects in the FRPs time locked to the target word. To investigate whether this concern might affect our results of interest, we conducted mixed effects regression analyses of the gaze durations on the pretarget word as well as on the pretarget FRPs from 60–260 ms postfixation (the time window in which the target baseline period would fall, on average). We found no significant effects (all $ps > .05$) of frequency or display change and no interactions in either gaze durations or FRP amplitudes. This suggests that the effects we report below are not caused by differences in fixation behavior that could change the baseline period for our FRP effects.

[3] Results from a post hoc power analysis conducted using PANGEA (v0.2; Westfall, 2015) indicated that with 16 items per condition and 45 participants, the current study design would be capable of detecting a medium effect size (d = 0.49) for the main effects at a power level of 0.8.

Figure 3

Single Fixation Duration on Target Word by Preview and Target Frequency



Note. Error bars represent standard error. See the online article for the color version of this figure.

ms, *SD* = 91.05), demonstrating a standard preview benefit, but when the preview was low frequency, the different preview condition led to numerically shorter fixations (*M* = 277.42 ms, *SD* = 102.64; not significant) on the target than the identical previews (*M* = 266.13 ms, *SD* = 91.00), demonstrating a reversed preview benefit (see Figure 3).

Regressions Out of Posttarget Words

There was a numerical increase in regressions for low frequency previews, but this difference was not statistically significant (Table 4). There was a significant display change effect, which did not interact with preview frequency; readers made more regressions when the preview was different from the target than when it was identical, but this did not differ between high and low frequency previews (see Figure 4). This did not align with our hypotheses and does not replicate the pattern reported by Schotter et al. (2019) in which the display change effect on regressions was numerically larger for high frequency previews, although it does replicate the statistical patterns they reported because the interaction was not significant. We address the comparison between our data and those data in the Discussion section.

Table 4

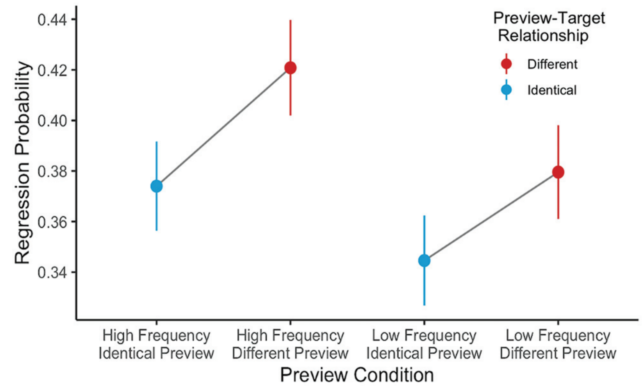
Result of Binomial Generalized Linear Mixed Effects Regression Predicting Regression Probability by Preview Frequency and Preview Validity

	Regression rate			
Predictors	Est.	*SE*	*z*	*p*
(Intercept)	−0.64	0.15	−4.36	**<.001**
Preview frequency (identical)	−0.06	0.13	−0.46	.642
Display change	0.22	0.11	1.99	**.045**
Display Change × Preview Frequency	−0.12	0.19	−0.63	.529

Note. The model produced a singular fit. Random effects variance and correlation estimates are presented in Table B2 of Appendix B. Bold font indicates statistical significance at the .05 α-level.

Figure 4

Probability of Regressing Out of the Posttarget Region After Leaving the Target Region by Preview and Target Frequency



Note. Error bars represent standard error. See the online article for the color version of this figure.

FRPs

In addition to confirmatory analyses (described below), we conducted exploratory analyses on the FRP data because there have been relatively few studies investigating frequency and parafoveal processing using FRPs during reading. Therefore, exploratory analyses allow us to fully understand the time course and distribution of these effects during natural sentence reading, which may differ from patterns found previously in ERPs to serially presented words.

Based on previous findings, we had reason to expect an early frequency effect, followed by a display change effect at occipitotemporal electrode sites, as well as an effect of both frequency and display change on the N400. Our selection of electrodes and time windows were guided by Niefind and Dimigen (2016) for the pre-N400 effects and by the canonical scalp distribution and latency for the N400. We chose to replicate these time windows and scalp locations from Niefind and Dimigen (2016) because their design was almost identical to ours aside from using word lists instead of sentences and they provide extensive justification based on the prior literature for why they would expect display change effects and frequency effects to arise in these time windows. In the 140–200 ms time window, we expected an effect of preview frequency, with more negative amplitudes for low compared with high frequency previews. In the 200–300 ms time window we expected a display change effect, such that the different preview would lead to more negative amplitudes than the identical preview. We also expected an interaction whereby there is a larger effect of display changes for high frequency previews than low frequency previews in the 200–300 ms time window, based on the patterns of data reported (but not analyzed) by Niefind and Dimigen (2016). In the N400 time window, we expected both a display change effect (more negative for invalid preview compared with identical) and a preview frequency effect (more negative for low compared with high frequency).

Confirmatory Analyses

We constructed separate lmer models for average amplitudes within each of the time windows and spatial ROIs: 140–200 ms and 200–300 ms post fixation on the target at occipitotemporal

Table 5

Result of Linear Mixed Effects Regression Predicting FRP Amplitudes for the Effects of Preview Frequency and Display Change

	140−200 ms (occipital)				200−300 ms (occipital)				300−500 ms (centro-parietal)			
Predictors	Est.	*SE*	*t*	*p*	Est.	*SE*	*t*	*p*	Est.	*SE*	*t*	*p*
(Intercept)	−1.00	0.28	−3.5	**<.001**	−0.29	0.22	−1.33	.188	−0.92	0.25	−3.66	**<.001**
Preview frequency (identical)	−0.39	0.32	−1.24	.219	−1.06	0.33	−3.16	**.002**	−1.28	0.42	−3.02	**.003**
Display change	−0.13	0.24	−0.55	.586	−0.6	0.24	−2.52	**.014**	−1.41	0.31	−4.55	**<.001**
Display Change × Preview Frequency	0.44	0.48	0.92	.365	1.75	0.53	3.28	**.001**	2.21	0.68	3.25	**.002**

Note. FRP = fixation related potentials. All significant effects at $\alpha = .05$ are still significant when controlling for the multiple comparisons in each time window using a Bonferroni correction ($\alpha = .05/3 = .017$). All analyses produced singular fits. Random effects variance and correlation estimates are presented in Table B3 of Appendix B. Bold font indicates statistical significance at the .05 α-level.

electrode sites (P7, P8, O1, O2), and 300–500 ms post fixation on the target at centroparietal sites (Cz, Pz, CP1, CP2; Table 5). All models had the same fixed effects structure, which included main effects of preview frequency (entered as a centered, sum-to-zero contrast), and display change (entered as a treatment contrast with the identical condition as the baseline), and an interaction between them with maximal random effects structure (intercepts and slopes for all fixed effects) for participants and items.

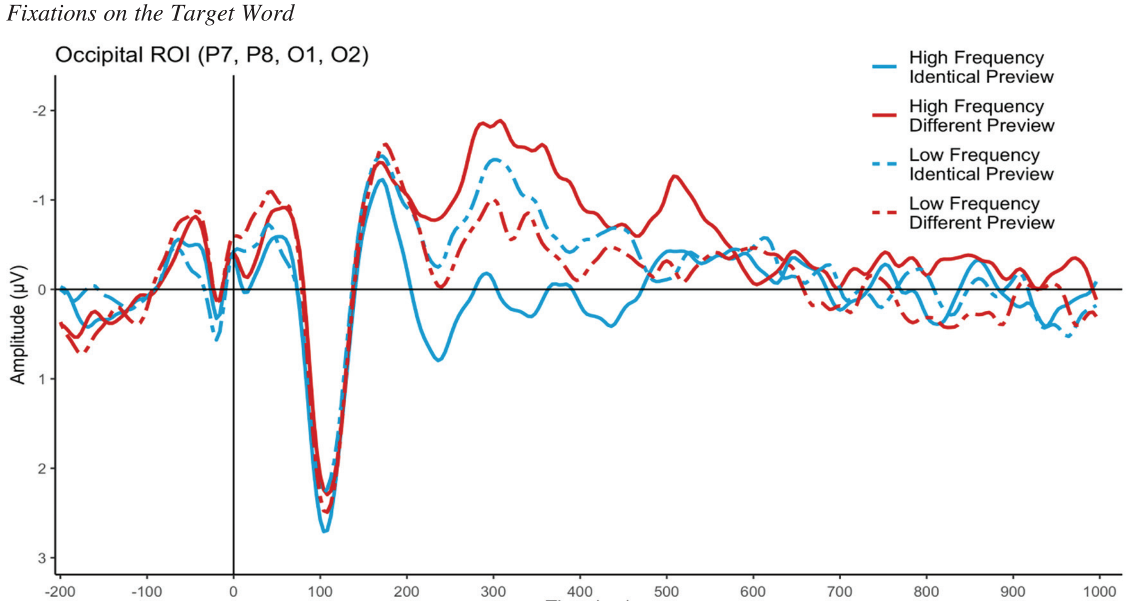
In the 140–200 ms occipital ROI, there were no significant effects. In the 200–300 ms occipital ROI, there was a significant effect of display change such that the amplitude was more negative for the different compared with identical preview conditions, a significant frequency effect (for identical display) and there was a significant interaction between display change and preview frequency such that there was a larger negativity for the different compared with identical preview condition for high frequency previews, but no difference for low frequency previews (see Figure 5). These patterns were confirmed by follow-up analyses predicting the display change effects for high and low frequency previews separately. There was a significant display change effect for high frequency, but not low frequency, previews in both the 200–300 ms time window at occipital sites and the N400 time window, centroparietally (see Table 6). The same pattern is observed in the N400 ROI: there was a significant display change effect and a significant interaction, with the display change effect only occurring in the high frequency preview condition (see Figure 6). Plots of the scalp topographies for the effect of word frequency in the identical preview condition (see Figure 7) and the display change effects calculated separately for the high frequency and low frequency previews (see Figure 8) confirm the timing and location of these effects.

Exploratory Analyses

Cluster-based permutation tests were performed using the Mass Univariate ERP Toolbox (Groppe et al., 2011) in Matlab to detect any reliable effects of word frequency and display changes that may not have been represented in our a priori hypotheses. The FRPs were submitted to repeated-measures pairwise *t* tests based on the cluster mass statistic using the

Figure 5

FRP Waveforms by Preview Type Averaged Across Occipital Scalp Electrodes Time-Locked to Initiation of Single Fixations on the Target Word



Note. FRP = fixation related potentials. See the online article for the color version of this figure.

Table 6
Follow-Up Linear Mixed Effects Regression Analyses of Display Change Effect in FRPs Separately for High and Low Frequency Preview Conditions

	200−300 ms (high frequency)				200−300 ms (low frequency)				300−500 ms (high frequency)				300−500 ms (low frequency)			
Predictors	Est.	*SE*	*t*	*p*	Est.	*SE*	*t*	*p*	Est.	*SE*	*t*	*p*	Est.	*SE*	*t*	*p*
(Intercept)	0.24	0.27	0.90	.37	−0.81	0.27	−2.97	**.004**	−0.30	0.31	−0.97	.338	−1.56	0.33	−4.75	**<.001**
Display change	−1.48	0.35	−4.23	**<.001**	0.28	0.36	0.77	.45	−2.49	0.43	−5.72	**<.001**	−0.33	0.47	−0.70	.487

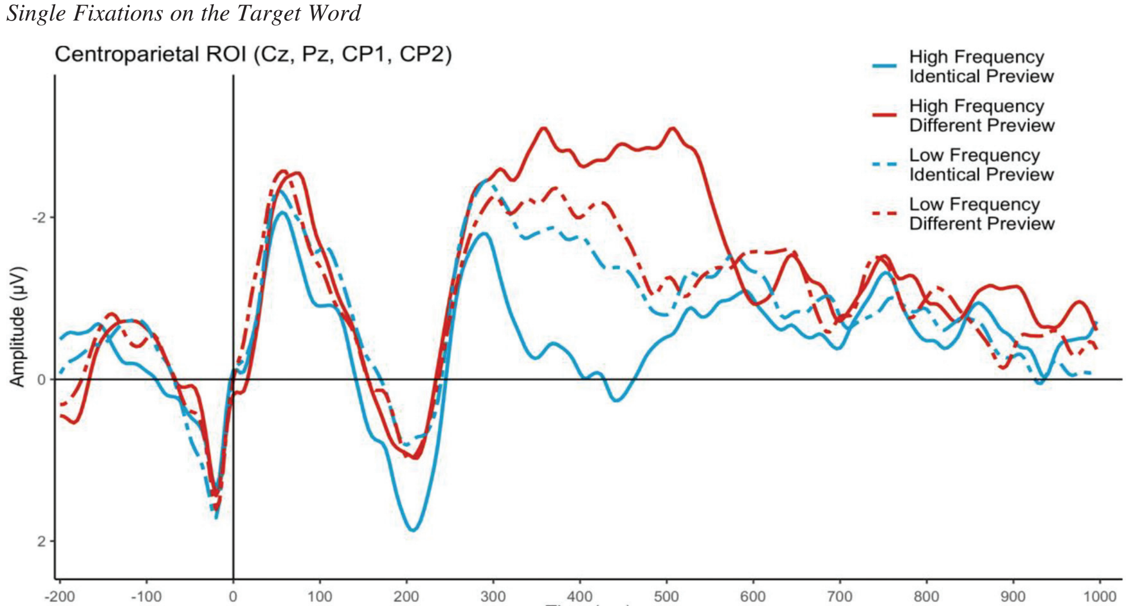
Note. FRP = fixation related potentials. These follow-up analyses were conducted to aid interpretability of the significant interactions from analyses in Table 5. Therefore, only the time windows with significant interactions were analyzed. Random effects variance and correlation estimates are presented in Table B4 of Appendix B. Bold font indicates statistical significance at the .05 α-level.

original data and 2,500 random within-participant permutations of the data (Bullmore et al., 1999) with a family-wise alpha level of .05. Any electrodes within approximately 5.44 cm of one another were considered spatial neighbors and adjacent time points were considered temporal neighbors. Clusters were formed for each permutation that included all neighboring *t*-scores corresponding to uncorrected *p*-values of .05 or less. This method has been shown to have relatively good power to detect broadly distributed effects like the N400 (Groppe et al., 2011; Maris & Oostenveld, 2007). Although the earlier frequency and display change effects were likely to be more focal, we believed they would not be so short-lived or spatially localized that this test would be insensitive to them. Additionally, Niefind and Dimigen (2016), testing similar comparisons, expressed concern about the excessively conservative nature of the t-max permutation test (Blair & Karniski, 1993) used in their exploratory analyses, so we chose a slightly more liberal method to avoid Type II errors.

Three comparisons were tested: the frequency effect (low–high) when the preview was identical, the display change effect (different–identical) when the preview was high frequency, and the display change effect (different–identical) when the preview was low frequency. The sum of the *t*-scores in each cluster is the "mass" of that cluster and the most extreme cluster mass in each of the 2,501 sets of tests was recorded and used to estimate the distribution of the null hypothesis. A null hypothesis distribution was estimated using the most extreme cluster mass (sum of the *t*-scores in that cluster) in each of the 2501 sets of tests. Each test included all 27 scalp electrodes and all time points between 110 and 550 ms (the time window of the effects of interest), resulting in 5,967 total comparisons. The results of this analysis can be visualized with raster plots of significant *t*-score clusters for each of the comparisons tested.

This analysis revealed significant preview frequency effects, with more negative amplitudes for low compared with high frequency words, beginning at 200 ms and lasting until 360 ms at left occipitotemporal sites (P7 and O1) as well as a significant cluster beginning at 284 ms and lasting until 468 ms at centroparietal sites (CP2, Pz, P4; Figure 9A). For the display change effect in the high frequency preview conditions, significant

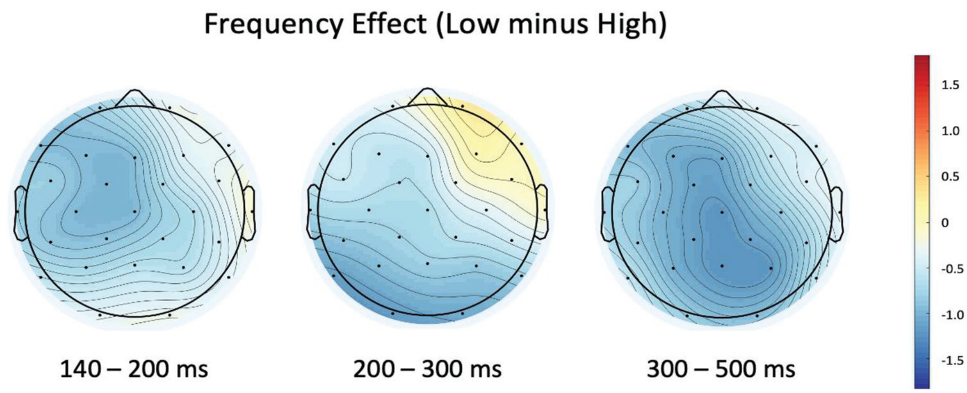
Figure 6
FRP Waveforms by Preview Type Averaged Across Centroparietal Scalp Electrodes Time-Locked to Initiation of Single Fixations on the Target Word



Note. FRP = fixation related potentials. See the online article for the color version of this figure.

Figure 7

Scalp Topographies of Amplitude (μV) Differences for the Frequency Effect (Low Minus High Frequency for Identical Conditions) Averaged Across the Time Windows of the Confirmatory Analysis (140–200 ms, 200–300 ms, and 300–500 ms)



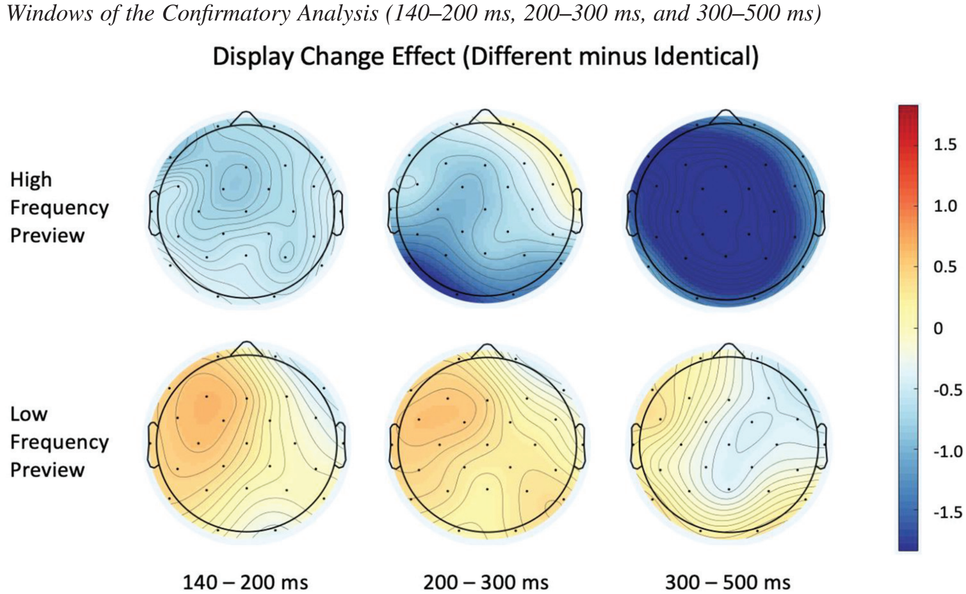
Note. See the online article for the color version of this figure.

clusters began at 186 ms at occipital, central, and parietal sites, which expanded to encompass the entire scalp by 320 ms and lasted across most of the scalp locations up through the end of the analysis time window at 550 ms (Figure 9B). In contrast to these broadly distributed and robust effects of the display change for high frequency previews, there were no significant display change effects when the preview was low frequency (Figure 9C). These findings support what we observed in the confirmatory analysis, that display change effects are robust, but are driven by scenarios in which the preview is high frequency and is processed more deeply before fixation on the target.

The timing of the onset of the frequency effect and the high frequency preview display change effect are strikingly similar, as demonstrated by the difference waves of these effects plotted at the occipital (Figure 10A) and centroparietal sites (Figure 10B). However, the high frequency display change effect was larger in amplitude than the pure frequency effect and persisted longer in time than the frequency effect. This slightly smaller and shorter effect of frequency time-locked to the identical target word makes sense because some preprocessing of the words could already have begun parafoveally, mitigating processing difficulty on a low frequency word once it is fixated. In the high frequency preview display change condition, on the other hand, any processing

Figure 8

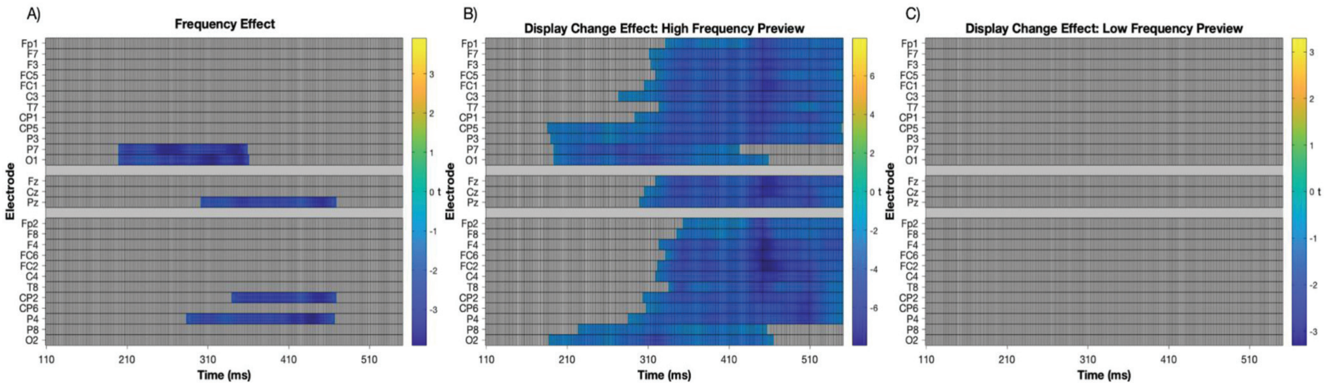
Scalp Topographies of Amplitude (μV) Differences for the Display Change Effect (Different Minus Identical Preview) for High and Low Frequency Preview Conditions Averaged Across the Time Windows of the Confirmatory Analysis (140–200 ms, 200–300 ms, and 300–500 ms)



Note. See the online article for the color version of this figure.

Figure 9

Results of Exploratory Cluster-Based Permutation Tests for the Frequency Effect in Identical Preview Conditions (A), and the Display Change Effect for High Frequency Previews (B), and Low Frequency Previews (C)



Note. See the online article for the color version of this figure.

completed parafoveally conflicts with the low frequency foveal word, so this conflict must be resolved and recognition of the newly fixated word must begin again.

Discussion

We investigated the influence of word frequency and trans-saccadic integration difficulty during sentence reading using two indices: the programming of eye movements and the amplitude of neural activity unfolding over time after a word is fixated (i.e., FRPs). In terms of oculomotor behavior, single fixation durations showed an effect of word frequency (i.e., shorter fixation durations for high frequency compared with low frequency previews) and an interaction between preview frequency and the display change effect (i.e., a standard preview benefit for high-frequency previews, longer fixations for different compared with identical previews, and a reversed preview benefit for low frequency previews, shorter fixations for different compared with identical previews). Regressive eye movements initiated downstream of the target word showed only an effect of display changes (i.e., there were more regressions when the preview was different than when it was identical). In terms of neural activity, we found an effect of word frequency (i.e., more negative amplitudes for low frequency compared with high frequency parafoveal preview words) at occipital sites starting around 200 ms and lasting to around 450 ms post fixation on the target word. We also found an interaction between preview frequency and the effect of display changes (i.e., more negative amplitudes for different compared with identical previews only when the preview was high frequency) that started occipitally around 200 ms and became broadly distributed from around 300 to 550 ms postfixation on the target. We discuss each of these findings in turn before discussing the fact that they suggest a partial alignment between oculomotor behavior and underlying neural activity.

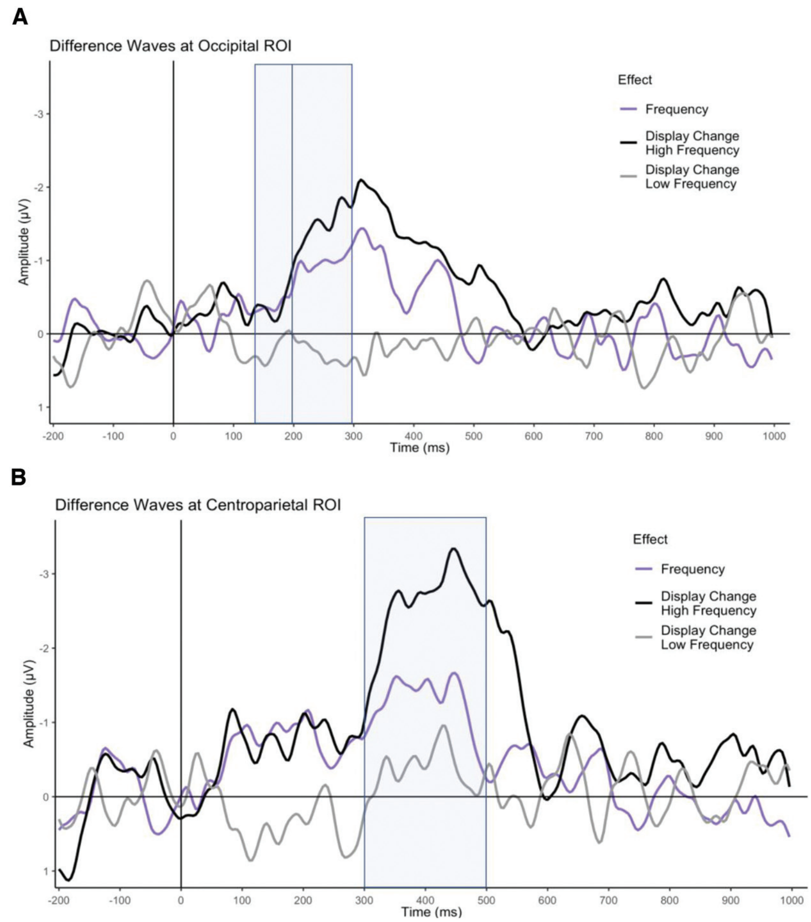
Eye Movement Patterns

We replicate many past findings from the eye tracking literature. First, our clear frequency effect when no display change occurred shows that reading behavior is sensitive to the ease of word recognition (Rayner & Duffy, 1986; see Rayner, 1998). However, we also find that word frequency information is obtained both foveally and parafoveally, and that properties of the word perceived from both of these locations influence oculomotor programming. Although this may suggest that trans-saccadic integration may drive oculomotor decisions (see Cutter et al., 2015), the fact that the display change effect interacted with preview frequency suggests that the benefit of the identical parafoveal preview does not outweigh the benefit of having a higher frequency preview. Although the magnitude of the fixation durations in the two display change conditions appears comparable, there could be two very different things going on in these two scenarios. When the preview is high frequency, the reader may get enough information from the upcoming word to complete a substantial amount of the word recognition process and plan eye movements away from the target word before landing on it (i.e., forced fixations), resulting in a reversed preview benefit (Schotter & Leinenger, 2016). On the other hand, when a preview is low frequency, minimal parafoveal processing may occur because the parafoveal information is difficult to recognize. Consequently, when the reader lands on a different target word, the disruption from the display change might be minimal and processing of the high-frequency target could proceed without much display change cost (see Schotter, 2018). Compared with the identical low frequency condition, we would then expect shorter fixations on the target because word recognition would depend largely on foveal processing, which is exactly what we find.

Our regression data do not align with the patterns we hypothesized, and do not entirely replicate past findings. We find only a main effect of display changes, which increases the likelihood that readers make a regression once they have read past the target word. Although we replicate the statistical findings of Schotter et al. (2019; i.e., a main effect of display changes and no interaction between display changes and preview frequency), numerically their data showed an interaction such that the display change effect was much larger for high frequency previews than low frequency previews. We may have not found this pattern in our study because our participants may have been less sensitive to the

Figure 10

Difference Waves of the Frequency Effect (Low - High Frequency for the Identical Preview Conditions) and Display Change Effects (Different - Identical Preview) for High and Low Frequency Preview Conditions at the Occipital ROI (A) and Centroparietal ROI (B)



Note. Confirmatory analysis time windows are highlighted for each ROI (140–200 ms and 200–300 ms in the occipital ROI; 300–500 ms in the centroparietal ROI). ROI = region of interest. See the online article for the color version of this figure.

display changes because the preview and target words were both plausible in the following sentence context, whereas in Schotter et al. (2019) readers sometimes encountered anomalous sentences due to a plausibility manipulation.

FRP Patterns

We find a robust frequency effect in the neural activity during natural sentence reading (more negative amplitudes for low compared with high frequency words when the preview is identical, lasting from approximately 200–450 ms in the exploratory analysis), which contrasts with null effects that have been reported elsewhere (Degno et al., 2019; Kretzschmar et al., 2015). Kretzschmar et al. (2015) suggested that frequency effects may not be observed in FRPs because information comes in so rapidly and that the effects of word difficulty may be reduced in time or magnitude, making them harder to detect, or differences in fixation durations may cause smearing with overlap in processing from one fixation to another. However, we did observe word frequency effects, so this cannot be the case. It is worth noting that Kretzschmar et al. (2015) did report an interaction between word frequency and predictability from 300 to 350 ms at lateral electrode sites such that the effect of word frequency was only apparent in the FRPs when the word was not predictable, which aligns with findings from more traditional single word presentation ERP studies (Dambacher et al., 2006; Van Petten & Kutas, 1990). In our study, we specifically used sentence frames in which the target words were not predictable (but were highly plausible) for this reason. Theoretically, it makes sense that word frequency effects might be amplified when the sentence context does not provide information about lexical identity and the majority of the word recognition process must be performed in a bottom-up manner.

The time course of our effects (i.e., starting around 200 ms post fixation on the target) suggests that readers may have initiated the word recognition process in parafoveal vision, before fixation on the word. In contrast, Degno et al. (2019) may have failed to find frequency effects in their coregistration study because their manipulation of the parafoveal preview involved masks that inhibited word recognition (e.g., x-strings or random letters). Indeed, this explanation aligns with the single fixation duration data, which suggest that preview frequency plays a role in oculomotor planning processes that are initiated before the reader leaving the target word (i.e., within 250 ms, which is the shortest condition mean for single fixation duration in our study). This conclusion is also supported by the display change effects that we observed in the FRP record, which suggest that neural recognition of display changes depends on preview frequency.

We observed more negative amplitudes for different compared with identical previews, but this display change effect was only present when the preview was high frequency and disappeared when the preview was low frequency. The display change effect for high-frequency previews is fairly robust in that it arises at occipital electrodes around 200 ms after fixation on the target word, expands to all electrodes around 350 ms after fixation on the target word and remains broadly distributed (except for occipital electrodes after about 450 ms). This interaction suggests that when a word is difficult to process parafoveally, the preview has very little impact on the processing of the word once it has been directly fixated. On the other hand, quite a bit of processing can occur when the preview is easy to recognize and this information is carried over for integration with the foveal input. When the foveal information does not match the parafoveal preview, a large disruption occurs, as demonstrated by the large negativities. The idea that more processing can occur parafoveally when the input is easier to recognize aligns with previous studies showing modulation of parafoveal processing depth depending on the available cognitive resources and contextual facilitation. For example, Barber et al. (2013) showed that the parafoveal N400 to semantic anomalies is dependent on the amount of sentence constraint (i.e., it is larger when the cloze probability of the word is .7 or above compared with when it is .4 or below) and the availability of sufficient processing time (i.e., it is larger when there is a 450 ms compared with 250 ms interval between word onsets). These findings together suggest that the limitations of parafoveal processing are not uniform across all circumstances and that there are a number of factors that appear to influence the extent to which parafoveal information is integrated during word recognition in reading.

It is difficult to compare our statistical interaction to the data reported by Niefind and Dimigen (2016) because their analyses were not conducted in a similar way. However, their raw means suggest the same interaction between preview frequency and display changes such that there is an effect for high-frequency previews (i.e., mean amplitudes that are 1 μV more negative for different previews compared with identical previews) but not for low frequency previews (i.e., only a .04 μV difference between different and identical previews) observed at occipital sites between 200 and 300 ms. They do not report data after this time window so it is unclear whether this effect would have extended into the N400 time window and scalp distribution that we observed in our study. Nevertheless, the fact that we find similar patterns of FRP data as did Niefind and Dimigen (2016) during natural sentence reading suggests that these effects of preview frequency and display changes are robust to experimental task (i.e., sentence reading for comprehension vs. "word list" reading while making semantic judgements).

One potential limitation of the current study, and an issue that researchers using the FRP technique will continue to have to contend with, is the fact that there are systematic differences in the timing of fixations on the target word based on the condition. Therefore, the FRP effects related to the manipulation and time locked to the target word overlap with subsequent fixations on the posttarget word. The average fixation duration on the target words was approximately 260 ms, which means that the N400 time window falls, most of the time during a fixation on the posttarget word. Therefore, the magnitude could be influenced by condition-specific jitter in the timing of the next fixation. In fact, Dimigen and Ehinger (2021) demonstrated a situation in a similar reading study with a display change during sentence reading in which overlapping potentials produced a spurious significant effect around 400 ms. They showed that this small, short-lived display change effect in the FRPs went away after applying a devonvolution model that accounted for saccade and fixation overlap. This deconvolution method (Ehinger & Dimigen, 2019) shows great promise for dealing with the methodological hurdle of overlapping potentials that vary systematically in latency by condition. Unfortunately, at the current time, this method cannot be used in conjunction with mixed effects modeling to account subject- and item-level variability, so we were unable to implement this approach with the analysis methods used here. However, in the current study, conditions were counterbalanced and rotated across sentences, so there should be no systematic differences in the posttarget word by condition. Additionally, the largest mean difference in fixation duration between conditions was ∼27 ms, which is over 10 ms smaller than the fixation duration differences in the data presented in Dimigen and Ehinger (2021). Furthermore, the N400 analyses were on a time window from 300–500 ms and the significant main effects and interactions in our study were quite robust. Therefore, it seems quite unlikely that the effects we see in the FRPs are due simply to a 27 ms (at most) difference in the fixation durations on the target word introducing spurious effects in the FRP effects across the N400 time window.

Comparisons Between Oculomotor Behavior and Neural Activity

Although both oculomotor behavior (i.e., single fixation durations) and neural activity (i.e., the N400 effects) show an interaction between preview frequency and display change effects, the nature of those interactions is different, suggesting that these measures reflect different ways in which the reading system uses parafoveal information. The eye movements suggest that display changes are processed to a smaller degree for high frequency previews and that oculomotor decisions can be preprogrammed based on the preview frequency, leading to apparently no cost, and even a benefit when information extracted from the parafovea clearly did not match the foveal input. In contrast, the benefit of the easy-to-process preview is not reflected in the FRP activity, but rather the cost of the mismatched foveal information is exacerbated or only present when a substantial amount of lexical processing occurred parafoveally (i.e., for high frequency previews).

When comparing these two measures, it is important to keep in mind that the time course differs. For a fixation duration of 250 ms, the decision to initiate an eye movement had to have occurred by approximately 125 ms after fixating the target word, because saccades take approximately 125 ms to program and execute (Becker & Jürgens, 1979; Rayner, 1998, 2009; Rayner et al., 1983; Schotter, 2018). The FRP effects we see do not begin until around 190 ms and continue past 500 ms after the fixation was initiated. Therefore, the processing cost for display change conditions may not register in the brain until the eyes have already been programmed to move on from the target word. Although these findings appear to be contradictory, the pattern in the FRPs does mirror previous reports of larger display change effects for high frequency previews in a different eye tracking measure, regressions from words after the target word (Schotter et al., 2019).

We suggest that a resolution to this contradiction may lie in an assumption of one of the most prominent models of eye movement control in reading (E-Z Reader; Reichle et al., 1998), which suggests that initial eye movement decisions are based on partial, not complete, word recognition, which can be initiated parafoveally. The actual stages of word recognition that are indexed by the commonly studied measures we focused on in this experiment (e.g., first pass fixation durations, the N400, regressive saccades) remains an open question. However, the fact that we found different patterns in fixation durations and the N400 suggests that these measures do in fact reflect different stages of processing and provides further evidence that first pass fixation durations are driven, at least some of the time, by early, cursory lexical processing and not by complete recognition of the word. Therefore, although preview frequency may influence not only reading time directly, but also the degree to which display changes disrupt processing, it may do so in different ways for early reading measures (e.g., single fixation duration) that likely reflect partial word recognition and later reading measures (e.g., the N400) that likely reflect a more complete stage of word recognition (see Schotter, 2018).

Our data also suggest that a core component of the architecture of E-Z Reader needs to be revised because it cannot accommodate the fact that fixation durations and FRPs show qualitatively different interactions between frequency and display changes. According to E-Z Reader, the durations of fixations are determined by the L_1 stage of lexical processing (that is influenced by the length, frequency, and predictability of the word) whereas the completion of word recognition is determined by the L_2 stage, which follows the L_1 stage, progresses independently of saccade planning, and the duration of which is a proportion of the duration of L_1 (Reichle et al., 1998). This latter feature is what needs to be revisited in light of our findings because this assumption implies that the outcome of L_1 (i.e., fixation duration) and the outcome of L_2 (for our purposes we assume this is reflected by the N400 FRP) should reflect the same computations and should show the same patterns of influences of lexical properties such as word frequency. However, our data show qualitatively different patterns between fixation durations and FRPs, suggesting that this assumption of E-Z Reader is not valid.

Likewise, OB1-Reader operates on the assumption that lexical processing only plays a role in eye movement timing when the word is fully recognized. Our data patterns show that early familiarity check processes, reflected in parafoveal frequency effects, can influence eye movements even when subsequent processing (i.e., after the eyes have moved on) is required for full recognition. Therefore, our data also suggest that this aspect of the model should be reconsidered to account for the fact that eye movements appear to be influenced by word frequency even before the completion of word recognition.

Based on a consideration of the differences between eye movement and FRP patterns, we also analyzed the likelihood of making a regression after leaving the target word. While we did find a display change effect in regressions, this effect did not interact with preview frequency as it appeared to do in prior work (cf. Schotter et al., 2019). We interpreted the lack of an interaction in our regression data as a consequence of our sentences making sense with both the preview and target word (contra the manipulation used by Schotter et al., 2019). However, because this is a null finding, we do not want to over interpret it as it may be due simply to an odd sample. Nevertheless, the FRP data suggest that the cost of display changes differed between the two preview frequency conditions.

One point to be made about this study is that, although this is one of the first experiments to use FRPs in natural sentence reading to investigate parafoveal frequency effects, the reader's experience is not truly that of a natural reading scenario. Display changes create a mismatch between parafoveal and foveal information, which does not normally happen during reading. Furthermore, display change effects are likely due to both the benefit of extracting parafoveal information in identical conditions as well as the cost of having an orthographic, semantic, and visual change occur in an invalid preview condition (Kliegl et al., 2013). Display changes have been shown to disrupt processing of the foveal word even when the preview is linguistically meaningless (Hutzler et al., 2013). However, the insights provided by the current study stand regardless of whether preview effects are conceptualized as benefits or costs. Our FRP patterns, in particular, demonstrate that the display change effect is different depending on how easy the parafoveal word is to recognize. Therefore, we leverage the existence of display change effects to demonstrate that parafoveal processing is much more extensive, and leads to larger effects, when the different preview is high frequency.

Further work should be done to investigate the relationship between effects observed in eye movement behavior and those observed in neural activity. Metzner et al. (2017) found that neural activity (e.g., the N400 and P600 effects elicited by semantic and syntactic anomalies, respectively) differed dramatically depending on the reader's behavior. They suggest that readers use different strategies to make sense of what they read: sometimes they use eye movements to reread and resolve misunderstandings when the neural signal indicates a problem. However, at other times they may not notice a comprehension problem (e.g., during "good enough" processing; Ferreira & Patson, 2007) and may not reread and will exhibit different neural responses.

Conclusion

The data presented here demonstrate the complementary nature of eye movement and EEG recordings in revealing a more complete account of the word recognition process during natural reading. A tension has existed between these two methodologies in that the patterns and time courses of word processing do not seem to fully align, especially under the assumptions that the decision to move the eyes away from a word *and* the peak of the N400 are both indices of *word recognition.* One way to resolve this tension is to recognize that the decision to move the eyes away from a word does not mean that word recognition has completed (see Reichle et al., 1998; Schotter, 2018). Furthermore, the N400 may be better characterized as an index of accumulating lexical activation rather than a "magical moment of recognition." Therefore, it may not be necessary that the time course and patterns of fixation durations and the N400 should align. We did expect that the regressions might reflect the same later stage of processing difficulty demonstrated in the N400, but these patterns did not fully align either.

Although we were unable to draw direct connections between the time course or patterns of processing reflected in the eye movements and EEG, we find that each data stream reveals aspects of the integration of foveal and parafoveal information that the other does not. The eye movements demonstrate that oculomotor planning can be driven primarily by parafoveal processing when the parafoveal word is easy to recognize and that mismatched foveal input does not interfere with the eyes progressing forward. This lack of a display change cost and the basing of oculomotor behavior on parafoveal processing does not appear to be reflected in the EEG record during the eye movement planning time window. However, the FRPs reveal a downstream cost of the different preview when the preview is high frequency, and easily recognized parafoveally, that is not apparent in the eye movements because it occurs after the fixation has already terminated. Together these complementary data sources demonstrate that the eye movement decisions can be determined by parafoveal processing alone but that the full recognition and integration of a word still relies on foveal processing. The extent to which parafoveal information is integrated with foveal input depends on the frequency of the parafoveal word. Therefore, the accumulated knowledge we gain by repeated encounters with a given word can fundamentally change how the oculomotor system and language processing system interact with the visual word form during reading.

These insights from reading also suggest an interesting characteristic of the interface between the eyes and the brain. Although eye movements are guided by higher level cognitive processing and by prior experience, the results from our study suggest that the eyes engage in a certain level of *good enough* processing. When and where the eyes move can be determined by an early familiarity check that may not align perfectly with later, deeper processing of the incoming visual input. Although eye movement execution may not flawlessly align with downstream processing demands, having them be determined by a cursory first-pass intake of information may be crucial for efficient processing in the long run.

Context

After receiving a grant to support collaboration across labs, the authors convened at a conference in 2019 to discuss our mutual interests in parafoveal semantic processing during reading and comparisons between eye tracking (Schotter & Jia, 2016), ERPs (Barber et al., 2010), and FRPs (Kretzschmar et al., 2009). During this conversation, the topic of lexical frequency effects on eye movements arose and it became clear that data patterns and inferences derived from the eye tracking and ERP literatures did not fully align. For example, frequency exhibits a robust effect on eye movements, and can even cause reversed preview benefit effects when orthogonally crossed with display changes (Schotter & Leinenger, 2016). However, the frequency effect on ERPs is less consistent, especially when semantic constraint from context is high (see Barber & Kutas, 2007), and is sometimes absent in coregistration studies (e.g., Degno et al., 2019; Kretzschmar et al., 2015). Therefore, we decided to also investigate how FRPs align (or do not) with the reversed preview benefit patterns in eye movements. We hypothesized that the preprogramming of eye movements would be modulated more strongly by the frequency of the parafoveal preview than by a mismatched foveal word, but the ongoing mental processing and downstream cost of the mismatched foveal word might be revealed in the FRP signal.

References

Andrews, S., & Veldre, A. (2019). What is the most plausible account of the role of parafoveal processing in reading? *Language and Linguistics Compass*, *13*(7), Article e12344. https://doi.org/10.1111/lnc3.12344

Angele, B., & Rayner, K. (2013). Processing the in the parafovea: Are articles skipped automatically? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 649–662. https://doi.org/10.1037/a0029294

Angele, B., Laishley, A. E., Rayner, K., & Liversedge, S. P. (2014). The effect of high- and low-frequency previews and sentential fit on word skipping during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 1181–1203. https://doi.org/10.1037/a0036396

Antúnez, M., Mancini, S., Hernández-Cabrera, J. A., Hoversten, L. J., Barber, H. A., & Carreiras, M. (2021). Cross-linguistic semantic preview benefit in Basque-Spanish bilingual readers: Evidence from fixation-related potentials. *Brain and Language*, *214*, Article 104905. https://doi.org/10.1016/j.bandl.2020.104905

Antúnez, M., Milligan, S., Hernández-Cabrera, J. A., Barber, H. A., & Schotter, E. R. (2022). Semantic parafoveal processing in natural reading: Insight from fixation-related potentials & eye movements. *Psychophysiology*, *59*(4), Article e13986. https://doi.org/10.1111/psyp.13986

Balota, D. A., & Rayner, K. (1983). Parafoveal visual information and semantic contextual constraints. *Journal of Experimental Psychology: Human Perception and Performance*, *9*(5), 726–738. https://doi.org/10.1037/0096-1523.9.5.726

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*(3), 364–390. https://doi.org/10.1016/0010-0285(85)90013-1

Balota, D. A., Yap, M. J., Hutchison, K. A., Kessler, B., Cortese, M. J., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459. https://doi.org/10.3758/BF03193014

Barber, H. A., & Kutas, M. (2007). Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Research Reviews*, *53*(1), 98–123. https://doi.org/10.1016/j.brainresrev.2006.07.002

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*(4), 523–531. https://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*(2), 152–156. https://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*(1), 48–59. https://doi.org/10.1111/j.1469-8986.2012.01489.x

Bates, D. (2010). lme4: Linear mixed-effects models using S4 classes. http://cran.r-project.org/web/packages/lme4/index.html

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research*, *19*(9), 967–983.

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, *30*(5), 518–524. https://doi.org/10.1111/j.1469-8986.1993.tb02075.x

Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., & Brammer, M. J. (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Transactions on Medical Imaging*, *18*(1), 32–42. https://doi.org/10.1109/42.750253

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2015). How is information integrated across fixations in reading. In A. Pollatsek & R. Treiman (Eds.), *Oxford handbook of reading* (pp. 245–260). Oxford University Press.

Dambacher, M., Kliegl, R., Hofmann, M., & Jacobs, A. M. (2006). Frequency and predictability effects on event-related potentials during reading. *Brain Research*, *1084*(1), 89–103. https://doi.org/10.1016/j.brainres.2006.02.010

Degno, F., Loberg, O., Zang, C., Zhang, M., Donnelly, N., & Liversedge, S. P. (2019). Parafoveal previews and lexical frequency in natural reading: Evidence from eye movements and fixation-related potentials. *Journal of Experimental Psychology: General*, *148*(3), 453–474. https://doi.org/10.1037/xge0000494

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dimigen, O. (2020). Optimizing the ICA-based removal of ocular EEG artifacts from free viewing experiments. *NeuroImage*, *207*, Article 116117. https://doi.org/10.1016/j.neuroimage.2019.116117

Dimigen, O., & Ehinger, B. V. (2021). Regression-based analysis of combined EEG and eye-tracking data: Theory and applications. *Journal of Vision*, *21*(1), 3–3. https://doi.org/10.1167/jov.21.1.3

Dimigen, O., Kliegl, R., & Sommer, W. (2012). Trans-saccadic parafoveal preview benefits in fluent reading: A study with fixation-related brain potentials. *NeuroImage*, *62*(1), 381–393. https://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*(4), 552–572. https://doi.org/10.1037/a0023885

Ehinger, B. V., & Dimigen, O. (2019). Unfold: An integrated toolbox for overlap correction, non-linear modeling, and regression-based EEG analysis. *PeerJ*, *7*, Article e7838. https://doi.org/10.7717/peerj.7838

Emmorey, K. D., & Fromkin, V. A. (1988). The mental lexicon. *Linguistics: The Cambridge Survey*, *3*, 124–149.

Ferreira, F., & Patson, N. D. (2007). The "good enough" approach to language comprehension. *Language and Linguistics Compass*, *1*(1–2), 71–83. https://doi.org/10.1111/j.1749-818X.2007.00007.x

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, *48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Hauk, O., & Pulvermüller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology*, *115*(5), 1090–1103. https://doi.org/10.1016/j.clinph.2003.12.020

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 166–190. https://doi.org/10.1037/a0033670

Hutzler, F., Fuchs, I., Gagl, B., Schuster, S., Richlan, F., Braun, M., & Hawelka, S. (2013). Parafoveal X-masks interfere with foveal word recognition: Evidence from fixation-related brain potentials. *Frontiers in Systems Neuroscience*, *7*, Article 33. https://doi.org/10.3389/fnsys.2013.00033

Hutzler, F., Schuster, S., Marx, C., & Hawelka, S. (2019). An investigation of parafoveal masks with the incremental boundary paradigm. *PLoS ONE*, *14*(2), Article e0203013. https://doi.org/10.1371/journal.pone.0203013

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, *40*(6), 431–439. https://doi.org/10.3758/BF03208203

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354. https://doi.org/10.1037/0033-295X.87.4.329

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*, *66*(3), 581–600. https://doi.org/10.1080/17470218.2012.658073

Kornrumpf, B., Niefind, F., Sommer, W., & Dimigen, O. (2016). Neural correlates of word recognition: A systematic comparison of natural reading and rapid serial visual presentation. *Journal of Cognitive Neuroscience*, *28*(9), 1374–1391. https://doi.org/10.1162/jocn_a_00977

Kretzschmar, F., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). Parafoveal versus foveal N400s dissociate spreading activation from contextual fit. *Neuroreport*, *20*(18), 1613–1618. https://doi.org/10.1097/WNR.0b013e328332c4f4

Kretzschmar, F., Schlesewsky, M., & Staub, A. (2015). Dissociating word frequency and predictability effects in reading: Evidence from coregistration of eye movements and EEG. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1648–1662. https://doi.org/10.1037/xlm0000128

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*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

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 event-related potentials. *Language, Cognition and Neuroscience*, *29*(5), 642–661. https://doi.org/10.1080/01690965.2013.866259

Leinenger, M. (2014). Phonological coding during reading. *Psychological Bulletin*, *140*(6), 1534–1555. https://doi.org/10.1037/a0037830

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*(10), 1361–1374. https://doi.org/10.1111/psyp.12502

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in*

Human Neuroscience, *8*, Article 213. https://doi.org/10.3389/fnhum.2014.00213

López-Peréz, P. J., Dampuré, J., Hernández-Cabrera, J. A., & Barber, H. A. (2016). Semantic parafoveal-on-foveal effects and preview benefits in reading: Evidence from fixation related potentials. *Brain and Language*, *162*, 29–34. https://doi.org/10.1016/j.bandl.2016.07.009

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Metzner, P., von der Malsburg, T., Vasishth, S., & Rösler, F. (2017). The importance of reading naturally: Evidence from combined recordings of eye movements and electric brain potentials. *Cognitive Science*, *41*(Suppl. 6), 1232–1263. https://doi.org/10.1111/cogs.12384

Niefind, F., & Dimigen, O. (2016). Dissociating parafoveal preview benefit and parafovea-on-fovea effects during reading: A combined eye tracking and EEG study. *Psychophysiology*, *53*(12), 1784–1798. https://doi.org/10.1111/psyp.12765

Payne, B. R., Stites, M. C., & Federmeier, K. D. (2019). Event-related brain potentials reveal how multiple aspects of semantic processing unfold across parafoveal and foveal vision during sentence reading. *Psychophysiology*, *56*(10), Article e13432.

Plöchl, M., Ossandón, J. P., & König, P. (2012). Combining EEG and eye tracking: Identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Frontiers in Human Neuroscience*, *6*, Article 278. https://doi.org/10.3389/fnhum.2012.00278

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), Article 148.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*(1), 65–81. https://doi.org/10.1016/0010-0285(75)90005-5

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. https://doi.org/10.1080/17470210902816461

Rayner, K., & Clifton, C., Jr. (2009). Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biological Psychology*, *80*(1), 4–9. https://doi.org/10.1016/j.biopsycho.2008.05.002

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*(3), 191–201. https://doi.org/10.3758/BF03197692

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(5), 1188–1200. https://doi.org/10.1037/0096-1523.22.5.1188

Rayner, K., Slowiaczek, M. L., Clifton, C., & Bertera, J. H. (1983). Latency of sequential eye movements: Implications for reading. *Journal of Experimental Psychology: Human Perception and Performance*, *9*(6), Article 912.

Reichle, E. D., & Sheridan, H. (2015). E-Z Reader: An overview of the model and two recent applications. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 277–290). Oxford University Press.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*(1), 125–157. https://doi.org/10.1037/0033-295X.105.1.125

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E–Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*, *7*(1), 4–22. https://doi.org/10.1016/j.cogsys.2005.07.002

Reingold, E. M., Yang, J., & Rayner, K. (2010). The time course of word frequency and case alternation effects on fixation times in reading: Evidence for lexical control of eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1677–1683. https://doi.org/10.1037/a0019959

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*(2), 653–668. https://doi.org/10.1037/a0034997

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory & Cognition*, *18*(4), 367–379. https://doi.org/10.3758/BF03197126

Schotter, E. R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language*, *69*(4), 619–633. https://doi.org/10.1016/j.jml.2013.09.002

Schotter, E. R. (2018). Reading ahead by hedging our bets on seeing the future: Eye tracking and electrophysiology evidence for parafoveal lexical processing and saccadic control by partial word recognition. *Psychology of Learning and Motivation*, *68*, 263–298. https://doi.org/10.1016/bs.plm.2018.08.011

Schotter, E. R., & Fennell, A. M. (2019). Readers can identify the meanings of words without looking at them: Evidence from regressive eye movements. *Psychonomic Bulletin & Review*, *26*(5), 1697–1704. https://doi.org/10.3758/s13423-019-01662-1

Schotter, E. R., & Jia, A. (2016). Semantic and plausibility preview benefit effects in English: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(12), 1839–1866. https://doi.org/10.1037/xlm0000281

Schotter, E. R., & Leinenger, M. (2016). Reversed preview benefit effects: Forced fixations emphasize the importance of parafoveal vision for efficient reading. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(12), 2039–2067. https://doi.org/10.1037/xhp0000270

Schotter, E. R., & Rayner, K. (2015). The work of the eyes during reading. In A. Pollatsek & R. Treiman (Eds.), *Oxford handbook of reading* (pp. 44–62). Oxford University Press.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, *74*(1), 5–35. https://doi.org/10.3758/s13414-011-0219-2

Schotter, E. R., Leinenger, M., & von der Malsburg, T. (2018). When your mind skips what your eyes fixate: How forced fixations lead to comprehension illusions in reading. *Psychonomic Bulletin & Review*, *25*(5), 1884–1890. https://doi.org/10.3758/s13423-017-1356-y

Schotter, E. R., von der Malsburg, T., & Leinenger, M. (2019). Forced fixations, Trans.-saccadic integration, and word recognition: Evidence for a hybrid mechanism of saccade triggering in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(4), 677–688. https://doi.org/10.1037/xlm0000617

Sereno, S. C., Brewer, C. C., & O'Donnell, P. J. (2003). Context effects in word recognition: Evidence for early interactive processing. *Psychological Science*, *14*(4), 328–333. https://doi.org/10.1111/1467-9280.14471

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). *A 21 word solution*. https://doi.org/10.2139/SSRN.2160588

SR Research Ltd. (2009). EyeLink® 1000 user manual (Version 1.5.0). https://www.sr-research.com/

Snell, J., van Leipsig, S., Grainger, J., & Meeter, M. (2018). OB1-reader: A model of word recognition and eye movements in text reading. *Psychological Review*, *125*(6), 969.

Stites, M. C., Payne, B. R., & Federmeier, K. D. (2017). Getting ahead of yourself: Parafoveal word expectancy modulates the N400 during sentence reading. *Cognitive, Affective & Behavioral Neuroscience*, *17*(3), 475–490. https://doi.org/10.3758/s13415-016-0492-6

Taylor, W. L. (1953). Cloze procedure": A new tool for measuring readability. *Journalism & Mass Communication Quarterly*, *30*(4), 415–433. https://doi.org/10.1177/107769905303000401

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in eventrelated brain potentials. *Memory & Cognition*, *18*(4), 380–393.

Veldre, A., & Andrews, S. (2016). Is semantic preview benefit due to relatedness or plausibility? *Journal of Experimental Psychology: Human Perception and Performance*, *42*(7), 939–952. https://doi.org/10.1037/xhp0000200

Westfall, J. (2015). *PANGEA: Power analysis for general ANOVA designs* (Unpublished manuscript). http://jakewestfall.org/publications/pangea.pdf

White, S. J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 205–223. https://doi.org/10.1037/0096-1523.34.1.205

Appendix A

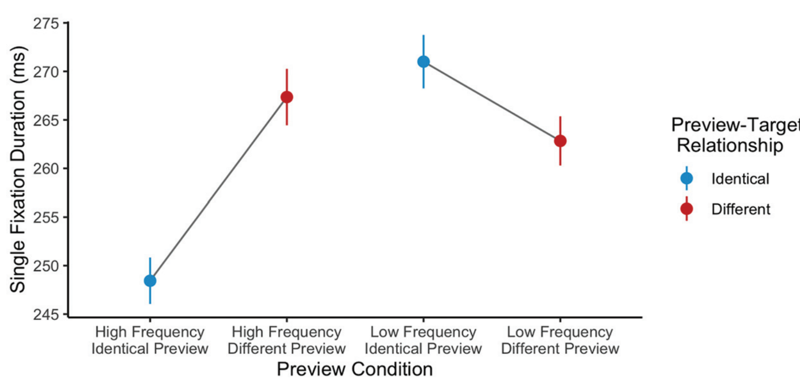
Supplementary Analyses of Later Fixation Duration Measures

Analyses were conducted following methods identical to those reported for single fixation duration in the main article. Each dependent measure was analyzed using a separate linear mixed effects regression model using the lme4 package (Version 1.1–17; Bates, 2010), and *p*-values were estimated using the Satterthwaite approximation via the lmerTest package (Kuznetsova et al., 2017). Preview frequency (low vs. high) was entered as a treatment contrast with the identical condition as the baseline and display change (different vs. identical) was coded with centered (i.e., sum-to-zero) contrasts.

The trials entered into these analyses were slightly different than those that entered in the SFD analyses because for those analyses only trials and participants that were included in the FRP analyses were retained for cleaner comparison between the two measures. In the following analyses, all trials and participants that had data for the given eye tracking measure were included.

Figure A1

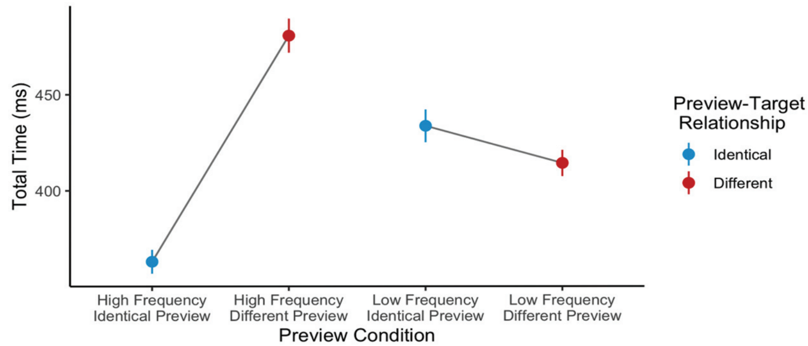
Gaze Duration on Target Word by Preview and Target Frequency



Note. See the online article for the color version of this figure.

(*Appendices continue*)

Figure A2
Total Time on Target Word by Preview and Target Frequency



Note. See the online article for the color version of this figure.

Table A1
Result of Linear Mixed Effects Regression Predicting Gaze Duration and Total Time by Preview Frequency and Preview Validity

	Gaze duration				Total time			
Predictors	Est.	*SE*	*t*	*p*	Est.	*SE*	*t*	*p*
(Intercept)	280.25	6.99	40.09	**<.001**	396.27	13.78	28.75	**<.001**
Preview frequency (identical)	31.97	5.62	5.69	**<.001**	69.43	14.26	4.87	**<.001**
Display change	12.02	3.84	3.13	**.002**	48.20	8.75	5.51	**<.001**
Display Change × Preview Frequency	−43.62	8.23	−5.30	**<.001**	−132.91	23.22	−5.72	**<.001**
Random effects								
σ^2	14,109.30				77,111.92			
τ_{00}	662.71_{trigitem}				$4{,}664.94_{\text{trigitem}}$			
	$2{,}329.09_{\text{subject}}$				$7{,}749.74_{\text{subject}}$			
τ_{11}	$1{,}063.77_{\text{trigitem.Pfreq1}}$				$9{,}576.61_{\text{trigitem.Pfreq1}}$			
	$339.81_{\text{trigitem.display1}}$				$275.07_{\text{trigitem.display1}}$			
	$1{,}922.15_{\text{trigitem.Pfreq1:display1}}$				$33{,}872.59_{\text{trigitem.Pfreq1:display1}}$			
	$350.83_{\text{subject.Pfreq1}}$				$2{,}190.03_{\text{subject.Pfreq1}}$			
	$189.40_{\text{subject.display1}}$				$1{,}415.56_{\text{subject.display1}}$			
	$1{,}047.36_{\text{subject.Pfreq1:display1}}$				$6{,}128.43_{\text{subject.Pfreq1:display1}}$			
ϱ_{01}	0.74				0.65			
	−0.26				−0.49			
	−0.87				−0.66			
	0.95				1.00			
	0.42				0.18			
	−0.87				−0.99			
N	144_{trigitem}				144_{trigitem}			
	59_{subject}				59_{subject}			
Observations	6,289				6,288			

Note. Bold font indicates statistical significance at the .05 α-level.

(*Appendices continue*)

Appendix B

Random Effects Variance and Correlations for Linear Mixed Effects Regression Analyses Reported in Results Section

Table B1

Random Effects for Linear Mixed Effects Regressions Predicting Single Fixation Durations (See Table 3)

	Participants				Items			
		Correlations				Correlations		
Analysis	Variance	Int.	Freq.	DC	Variance	Int.	Freq.	DC
Primary analysis								
Intercept	2,119.75	—	—	—	579.0	—	—	—
Frequency	143.41	0.99	—	—	653.62	0.45	—	—
Display change	3.84	0.80	0.69	—	152.64	−0.70	−0.95	—
Frequency × Display Change	381.37	−1.00	−0.99	−0.80	446.98	−0.99	−0.58	0.80
Follow-up (high frequency)								
Intercept	1,604.0	—	—	—	540.8	—	—	—
Display change	127.2	1.00	—	—	—	—	—	—
Follow-up (low frequency)								
Intercept	2,708.73	—	—	—	916.49	—	—	—
Display change	76.36	−1.00	—	—	398.79	−1.00	—	—

Table B2

Random Effects for Linear Mixed Effects Regressions Predicting Regression Rate (See Table 4)

	Participants				Items			
		Correlations				Correlations		
Analysis	Variance	Int.	Freq.	DC	Variance	Int.	Freq.	DC
Intercept	0.69	—	—	—	0.35	—	—	—
Frequency	0.07	−1.00	—	—	0.074	−0.86	—	—
Display change	0.08	−0.35	0.35	—	0.26	−0.51	0.81	—
Preview Frequency × Display Change	0.03	0.89	−0.89	−0.74	0.33	0.44	−0.35	−0.57

(*Appendices continue*)

Table B3
Random Effects for Linear Mixed Effects Regressions Predicting FRP Amplitudes (See Table 5)

	Participants				Items			
		Correlations				Correlations		
Analysis	Variance	Int.	Freq.	DC	Variance	Int.	Freq.	DC
140−200 ms (occipital)								
Intercept	2.33	—	—	—	1.33	—	—	—
Frequency	0.37	0.94	—	—	2.05	−0.32	—	—
Display change	0.42	−0.99	−0.89	—	0.99	−0.91	0.46	—
Frequency × Display Change	2.27	−0.49	−0.69	0.46	2.83	−0.23	−0.84	0.08
200−300 ms (occipital)								
Intercept	0.97	—	—	—	0.97	—	—	—
Frequency	0.62	−0.33	—	—	3.14	0.47	—	—
Display change	0.35	−0.60	0.67	—	1.35	−0.88	−0.58	—
Frequency × Display Change	4.62	−0.07	−0.92	−0.45	3.47	−0.73	−0.84	0.92
300−500 ma (centro-parietal)								
Intercept	0.95	—	—	—	0.49	—	—	—
Frequency	0.23	0.20	—	—	2.62	0.05	—	—
Display change	<0.001	0.16	1.00	—	2.17	−0.41	−0.46	—
Frequency × Display Change	3.55	0.42	−0.80	−0.83	8.16	0.15	−0.96	0.54

Note. FRP = fixation related potentials.

Table B4
Random Effects for Follow-Up Linear Mixed Effects Regressions Predicting FRP Amplitudes (See Table 6)

	Participants				Items			
		Correlations				Correlations		
Analysis	Variance	Int.	Freq.	DC	Variance	Int.	Freq.	DC
200−300 ms (high frequency)								
Intercept	1.30	—	—	—	0.94	—	—	—
Display change	2.00	−0.48	—	—	0.19	−1.00	—	—
200−300 ms (low frequency)								
Intercept	0.71	—	—	—	2.45	—	—	—
Display change	0.97	−0.74	—	—	4.12	−0.90	—	—
300−500 ms (high frequency)								
Intercept	0.72	—	—	—	0.73	—	—	—
Display change	0.92	−0.66	—	—	1.39	−1.00	—	—
300−500 ms (low frequency)								
Intercept	0.86	—	—	—	1.09	—	—	—
Display change	0.67	0.56	—	—	5.93	−0.67	—	—

Note. FRP = fixation related potentials.

Received June 16, 2021
Revision received May 23, 2022
Accepted June 17, 2022 ■