ORIGINAL ARTICLE

PSYCHOPHYSIOLOGY SPR | WILEY

Semantic parafoveal processing in natural reading: Insight from fixation-related potentials & eye movements

Martín Antúnez[1] | **Sara Milligan[2]** | **Juan Andrés Hernández-Cabrera[1]** | **Horacio A. Barber[1,3,4]** | **Elizabeth R. Schotter[2]**

[1]Cognitive Psychology Department, Universidad de La Laguna, Santa Cruz de Tenerife, Spain

[2]Department of Psychology, University of South Florida, Tampa, Florida, USA

[3]Institute for Biomedical Technologies (ITB), Universidad de La Laguna, Santa Cruz de Tenerife, Spain

[4]Basque Center on Cognition, Brain and Language (BCBL), San Sebastián, Spain

Correspondence

Martín Antúnez, Cognitive Psychology Department, Universidad de La Laguna, Campus de Guajara s/n, S/C de Tenerife 38205, Spain.
Email: mantunez@ull.edu.es

Abstract

Prior research suggests that we may access the meaning of parafoveal words during reading. We explored how semantic-plausibility parafoveal processing takes place in natural reading through the co-registration of eye movements (EM) and fixation-related potentials (FRPs), using the boundary paradigm. We replicated previous evidence of semantic parafoveal processing from highly controlled reading situations, extending their findings to more ecologically valid reading scenarios. Additionally, and exploring the time-course of plausibility preview effects, we found distinct but complementary evidence from EM and FRPs measures. FRPs measures, showing a different trend than EM evidence, revealed that plausibility preview effects may be long-lasting. We highlight the importance of a co-registration set-up in ecologically valid scenarios to disentangle the mechanisms related to semantic-plausibility parafoveal processing.

KEYWORDS

co-registration, FRPs, N400, natural reading, parafovea, semantics

1 | INTRODUCTION

One remarkable characteristic of reading is the large amount of information that we can extract from a text in an extremely brief period of time. However, there are limitations to how fast we can scan strings of words. During the presentation of linguistic and orthographic stimuli, an accurate description of the constraints of the visual system is necessary to fully understand the nature of subsequent cognitive operations. For instance, readers can process words located not only in the foveal visual field, but also in the parafoveal region –located between 1 and 5 degrees away from the fixation point. However, information in the parafoveal region is of poorer quality, due to decreased visual acuity and visual attention (Schotter et al., 2012). Therefore, the orthographic input will depend on the perception of letters at different spatial locations in combination with a series of sequential eye movements and attentional shifts. This leads us to some relevant questions: How many letters can we perceive in the parafoveal

visual field? How deep do we process them? Is that information used only to guide our gaze or it is contributing to improve comprehension as well? At the core of all those issues is the debate on whether word meanings can be activated and integrated from parafoveal perception. In this study we have focused on parafoveal semantic processing during natural sentence reading combining two methodological approaches: The eye-tracking and the EEG-ERP research.

1.1 | Semantic parafoveal processing: Evidence from eye tracking research

Eye movement research has investigated parafoveal processing using the *gaze-contingent boundary paradigm* (Rayner, 1975), which allows making inferences about how information obtained from parafoveal perception modulates subsequent reading behavior. In the boundary paradigm, an invisible boundary is located before a previewed word. When the reader's gaze crosses the invisible boundary, the previewed word is replaced by a target word as the reader fixates it. Therefore, the previewed word could only have been perceived from the parafovea during the fixation of the previous word, and any difference in reading time of the target word when it is fixated must be due to that parafoveal processing (i.e., a parafoveal preview effect). The general conclusion from this paradigm is that readers regularly use orthographic and phonological features from parafoveal words, since fixated words need less time to be read after orthographically and phonologically related previews. On the other hand, evidence about the activation of semantic information was initially scarce (Hohenstein et al., 2010; Yan et al., 2009), leading to the conclusion that semantic information was not accessed parafoveally (Altarriba et al., 2001; Hyönä & Häikiö, 2005; Rayner et al., 1986; White et al., 2008; see Schotter et al., 2012).

Subsequent experiments have found that semantic information can be obtained from the parafovea (Rayner & Schotter, 2014; Schotter, 2013; Schotter et al., 2015), but they concluded that semantic preview effects are determined not by the relationship between the preview and target word, but rather by the semantic relationship between the parafoveal preview and the sentence context (i.e., plausibility preview effects; see Andrews & Veldre, 2019; Schotter, 2018). For example, Schotter and Jia (2016) used the boundary paradigm with identical, plausible and implausible unrelated previews (e.g., "Kevin's brother ate all their fresh/baked/place bread in the apartment"), in addition to synonyms and antonyms (e.g., "Harry bought a broken watch/clock to repair for fun" and "Jane will travel north/south on her trip to Los Angeles next week"

respectively); these words were read in low-constraint sentences, in order to ensure that predictability did not affect the processing of plausibility. They found that all plausible previews led to shorter durations compared to the implausible preview in first-pass reading measures on the target word, with no effects in later reading measures. Discrepancies between earlier and later eye movement reading measures could suggest that plausibility preview effects are short-lived; while implausible conditions had longer first-pass reading durations, total fixation durations were similar between plausible and implausible preview conditions (Schotter & Jia, 2016; Veldre & Andrews, 2016, 2017, 2018c). Andrews and Veldre (2019) suggested that a plausible preview may lead to later costs related to subsequent trans-saccadic integration processes between the preview and target words, which would lead to higher rates of regressions to the target word. This could explain the equivalence between plausible and implausible conditions in total fixation durations, since integrative processes of the preview with both the target word and the sentence context may influence this later processing measure. Since evidence suggest that plausibility preview effects are independent of trans-saccadic integration processes (Schotter & Leinenger, 2016; Veldre & Andrews, 2016; see Schotter, 2018), it is still possible that integrative processes of preview with the sentence context are still present in later processing after fixating the target word, but undetected by total fixation duration measures.

1.2 | Semantic parafoveal processing: Evidence from ERPs and FRPs

While studying the time-course of processing may be limited in eye tracking research, EEG has proved particularly useful in this regard. Word recognition is a multimodal and cumulative process that extends along time, determined by many lexical and contextual factors (Barber & Kutas, 2007). Early EM measures are very sensitive to the computations that determine eye movement control. However, considering the characteristics and speed of natural reading, eye movement control uses only the minimum amount of information necessary to maximize the efficiency of saccades. Word processing is not finished after our gaze leaves a given word. It continues until meanings are fully processed and involves the continuous updating of mental representations. Language-related Event-Related Potential (ERP) components like the N400 peak much later (around 400 ms) than the average fixation duration (250 ms), and therefore are crucial physiological markers that may help us to understand the discrepancies between early and late EM measures. For instance, some cognitive processes may not be detected by early EM

measures if they take place after saccades, but they may still modulate late ERP components and to have an impact on much later EM behavioral measures. Therefore, EEG and EM measures can be mutually complementary when describing the time course of parafoveal semantic processing during reading. Fixation-Related Potentials (FRPs) may be experimentally obtained through a co-registration set-up, allowing us to obtain ERPs time-locked to fixation onsets (similarly to EM fixation events) during natural sentence reading (Dimigen et al., 2011). By obtaining FRPs, semantic processing in the time-course of plausibility preview effects may be detected through the N400 amplitude modulation, which is an index of the ease of semantic access determined by sentence-level context information (Kutas & Federmeier, 2011).

Kretzschmar and colleagues (Kretzschmar et al., 2009) reported FRP effects compatible with parafoveal semantic processing in a natural reading task. They found modulations of the N400 component associated with semantically incongruent compared to congruent predictable words in highly constraining sentence constructions (e.g., "the opposite of black is white/yellow/nice"). These effects were found when Event-Related Potentials (ERPs) were time-locked to the last fixation before the target fixation, providing evidence that at least some semantic processing of the critical words took place parafoveally. However, considering the strong predictability manipulation used in that study, it is still an open question under which circumstances this kind of effect can be produced. In fact, a later study failed to replicate this parafoveal N400 effect in sentences with predictable targets but without extreme predictability (e.g., "The extremely skinny model looked like she suffered from *anorexia* and a lack of sleep") when compared with unpredictable targets (see Kretzschmar et al., 2015). Consequently, it is important to note that, especially in high-constraint sentences, semantic effects derived from predictability manipulations can be confounded with sub-lexical processing, as predictability effects may extend to the level of orthography (see Laszlo & Federmeier, 2009) by shaping expectations of orthographic word forms (see Schuster et al., 2021).

The time course of parafoveal semantic processing during reading has been also addressed with artificial reading tasks that allow for tight experimental control; the presentation of words in the sentence is controlled via Rapid Serial Visual Presentation with bilateral flankers (Flanker-RSVP) while the reader fixates the word at the center of the screen, which is flanked to the right by the next word of the sentence and to the left by the previous word of the sentence (Barber et al., 2010, 2011, 2013; Li et al., 2015). For example, Barber et al. (2010) used this paradigm to manipulate the parafoveal word presented in the right flanker, which could be congruous or incongruous with

the sentence context. Incongruent words in the parafoveally produced larger amplitudes in the N400 component time-locked to the presentation of the parafoveal word, showing that semantic processing of parafoveal words began before they were replaced by a new target word in the foveal region. In a later study, Barber et al. (2013) manipulated the contextual predictability of the critical words that were either congruent or incongruent within the sentential context. They again found larger N400 amplitudes for incongruent words when presented parafoveally while reading the previous word, both in high and low-constraint sentences. Interestingly, N400 modulations were greater under high contextual constraint, indicating that predictability can modulate the amount of parafoveal processing. In order to totally rule-out the possibility that predictions were primarily orthographic rather than semantic, Stites et al. (2017) used the same flanker-RSVP paradigm presenting a graded manipulation of the predictability of the target words, combining predictability and plausibility manipulations (high cloze probability, low cloze probability, unexpected but plausible, and anomalous words), which resulted in graded parafoveal N400 effects, with differences between unexpected plausible and anomalous words (i.e., a plausibility effect).

In spite of this evidence, it has not been established yet if the previously described ERP parafoveal semantic effects can be replicated under conditions of natural reading. In relation to this question, Barber et al. (2013; experiment 2) showed that parafoveal N400 effects in low constraint sentences were observed only at a slow stimulus presentation rate (SOA = 450 ms) but not when words were presented to a faster speed, similar to that of natural reading (SOA = 250 ms). Therefore, it seems that semantic N400 modulations related to predictability can interact with other sources of cognitive load to determine the amount of semantic parafoveal processing at any time (see also Payne et al., 2016). FRPs seem to be a natural step forward to tackle the ecological validity of parafoveal ERP findings in complex natural reading situations.

For instance, FRPs have already been useful in testing the ecological validity of parafoveal ERP and EM effects unrelated to semantic processing (e.g., Degno et al., 2019a, 2019b; Hutzler et al., 2013; Niefind & Dimigen, 2016; for a review, see Degno & Liversedge, 2020). Experimental conditions where previews and targets are visually different show greater processing costs when compared to conditions where previews and targets are identical, a preview effect related to display change frequently reported in EM research (see Schotter et al., 2012). The display change preview effect could be a mixture of preview benefits and preview costs (Kliegl et al., 2013). The mechanisms behind the greater preview costs of dissimilar previews may be affected by visual and attentional processes, for they can

appear in the absence of conflicting orthography, phonology or semantics (Hutzler et al., 2013) and they can be increased by saliency (Hutzler et al., 2019). Therefore, these effects may be triggered by a perceptual mismatch that affects low level visuo-attentional processes, as well as by the identical preview facilitation of the subsequent target processing. These display change effects have also been reported in Flanker-RSVP-ERP paradigms during controlled reading (see Li et al., 2015), where valid previews elicited smaller N1 and N400 components than invalid preview when the target word was presented. More interestingly, in a situation more similar to natural reading, Dimigen et al. (2012) obtained FRPs while participants read word lists freely from left to right, and they used the boundary paradigm to manipulate parafoveal information. They presented an identical, semantically related or semantically unrelated word as a preview. They found that identical previews, compared to the other conditions where a display change was present, lead to facilitatory effects reflected in shorter fixation durations and a more positive amplitude that emerged from around 170 to 280 ms in the PO9 and PO10 electrodes. As they indicated, their findings in fixation durations and FRP amplitudes may support the idea that the display change effect is related to a pre-activation of orthographic codes before meaning activation. Additionally, they also reported a modulation of the N400 component such that the identical condition was less negative than the conditions with invalid previews. Dimigen et al. (2012) proposed that the N400 attenuation derived from a valid preview could be equivalent to the repetition priming effect described in other visual word recognition studies (see Holcomb & Grainger, 2006, 2007), which would suggest that similar mechanisms of trans-saccadic integration of low-level features in flanker paradigms could be involved in natural sentence reading.

The extraction of FRPs through a co-registration set-up provides some important advantages. For instance, both FRP and EM data together may discern between different types of processing that cause either distinct or comparable disruption to both data streams (for a review, see Degno & Liversedge, 2020). Additionally, FRPs have already been successfully combined with the boundary paradigm in word pair or word lists reading experiments exploring semantic parafoveal processing (Antúnez et al., 2021; Dimigen et al., 2012; López-Pérez et al., 2016). This combination allows a better interpretation of the ERP components that are highly overlapped in a situation of natural reading. Therefore, the ecological validity advantage of obtaining both FRPs and EM with the boundary paradigm over traditional ERP and EM approaches alone may provide a deeper understanding of how parafoveal processing may be affected by additional cognitive processes inherent to natural sentence reading, especially those related to reading speed and eye-movement control.

1.3 | The present study

In this experiment, we analyzed the relationship between EM and FRP measures of semantic parafoveal processing in natural reading scenarios, posing two questions: (1) do ERP semantic parafoveal effects that have been obtained under controlled situations (e.g., Flanker-RSVP) replicate in a natural reading task? (2) Do these FRP-based plausibility preview effects provide clarity on the discrepancies in earlier and later EM measures? We recruited a sample of native English speakers and obtained FRPs through the co-registration of EM and EEG during a natural sentence-reading task. As in EM research, we used the boundary paradigm to manipulate the relationship of the previewed word with the sentence context. Participants read sentences such as "Harry bought a broken watch to repair for fun." We manipulated the previewed word so that it was either identical to the target (e.g., Harry bought a broken watch...), an unrelated but plausible preview (e.g., Harry bought a broken chair...) or an unrelated and implausible preview (e.g., Harry bought a broken peace...; see Figure 1). The identical condition represented a situation where preview and target words share all features, allowing us to explore trans-saccadic integration effects when compared to the other two conditions where dissimilar previews may lead to preview costs. The comparison of plausible and implausible previews allowed us to explore integration processes of semantic preview information with the sentence context independent of any relationship between the preview and target because the previews in these conditions were both orthographically, phonologically, and semantically unrelated to the target word. The plausibility manipulation within low constraint sentences allowed us to confirm genuine parafoveal semantic processing in natural reading, ruling out alternative explanations such as orthographic prediction. Additionally, the comparison of the identical preview with the plausible and implausible previews was useful to separate out preview costs related to perceptual dissimilarity and to a mere pre-activation of orthographic codes before meaning activation. FRPs were time-locked to the pretarget and target words, in order to explore whether parafoveal information can be processed during the fixation of the pretarget word and if such semantic information may modulate the processing of the target word when fixated.

We expected to replicate previous FRPs findings of preview effects related to a display change (Dimigen et al., 2012) in a more ecologically valid reading situation in early and later components of the FRP time-locked to
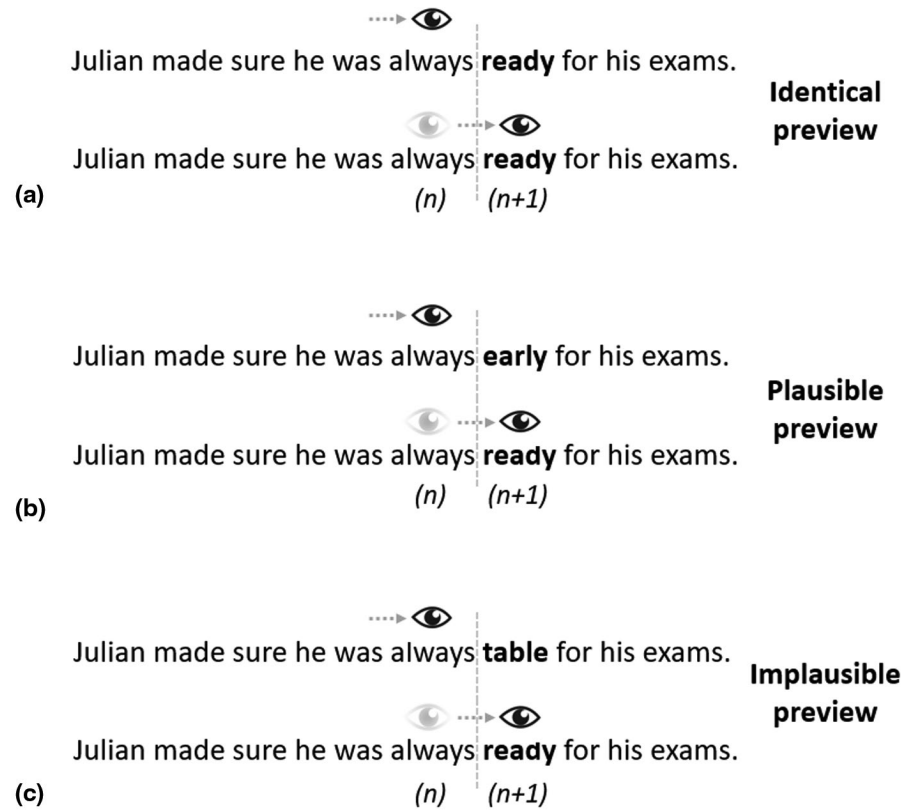
FIGURE 1 Illustration of the boundary paradigm. When readers crossed with their gaze an invisible boundary located between a pretarget (*n*) and a previewed word (*n* + 1), a target word replaced the preview word. The target word was always plausible to the sentence context. The previewed word could be identical to the target word (a), a different word but plausible to the sentence context (b) and a different word implausible to the sentence context (c)

the target word (e.g., N1 and N400). More importantly, considering previous electrophysiological evidence from controlled-reading paradigms where EMs were absent (Barber et al., 2013; Stites et al., 2017), we expected parafoveal word plausibility to modulate the N400 component time-locked to the fixation of the pretarget word (i.e., a greater negativity associated with the implausible preview). Such a finding would suggest that the N400 component involves semantic processing that is independent from EM behavior, as the effect would be found in both paradigms with and without the presence of eye movements, meaning that the semantic electrophysiological evidence is not disrupted or completely determined by the mechanisms related to eye movements. In addition to this, we expected plausibility preview effects in early reading measures on the target word and FRP components time-locked to fixation on the target word (i.e., around the 200 ms temporal window), consistent with previous EM evidence with similar experimental paradigms (Schotter & Jia, 2016; Veldre & Andrews, 2016, 2017, 2018c). Interestingly, and despite previous EM evidence showing that later reading measures are equivalent across plausible and implausible conditions, EEG measures (i.e., FRPs and the N400 component) may reveal types of processing undetected by fixation durations. Moreover, if plausibility effects are long-lasting, we would expect inconsistencies between EM and FRP

measures, finding modulations in the N400 component for the FRPs time-locked to the target word, with total fixation durations not showing plausibility effects. This would be our main guess related to the later time-course of plausibility effects for semantic experimental manipulations may be less disrupting compared to purely visual preview manipulations to EM measures, leading to less consistency between both data streams (Degno & Liversedge, 2020). On the other hand, if plausibility effects are short-lived and absent in later processing, we would not find modulations in the N400 component. Such consistency between both data streams would suggest that EM and FRPs measures share common cognitive mechanisms related to semantic-plausibility parafoveal processing.

2 | METHOD

2.1 | Subjects

Fifty-nine Psychology students at University of South Florida (Florida, United States) volunteered to participate in the experiment in exchange for course credits. After excluding participants due to failure to follow instructions or stay awake (*N* = 5), problematic recording (e.g., inability to sufficiently reduce impedances in

time or electrodes disconnected during recording; $N = 6$), and excessive data loss (i.e., subjects with fewer than 20 trials in any condition were excluded; $N = 11$), thirty-seven participants (21 females and 16 males, age: $M = 20.7$, $SD = 4.19$) were included in the analyses. They all were monolingual native English speakers, had normal or corrected vision, were right-handed and had no history of neurological disorders.

2.2 | Materials and design

One hundred twenty-six sentences were taken from Schotter and Jia (2016) for the study. In each sentence, a preview of a specific target word could be either identical, an orthographically, phonologically, and semantically unrelated word that was plausible in the context of the sentence or an orthographically, phonologically, and semantically unrelated word that was implausible in the sentence context. All preview words shared the same length with the target word, were similar in lexical frequency, and had low orthographic similarity to the target word (for non-identical preview; see Table 1). Cloze probability norming was conducted with 30 volunteers who were not in the main experiment. This revealed that none of the preview words were predictable in the sentences (Table 1).

In the original study, plausibility norms were collected for the entire sentence containing each of the preview/target words. For this study, we conducted an additional plausibility norming task, which included the sentence fragment only up to the preview word to confirm the plausibility manipulation at the point where the preview word was encountered (i.e., the point where the FRPs were time locked). For the norming study, 30 participants indicated if sentences were well or poorly written using a 1–7 Likert scale. Sentence conditions were counterbalanced and randomly presented. From the norming procedure, the average plausibility rating was 4.6 ($SD = 0.98$), 4.6 ($SD = 0.9$), and 2.9 ($SD = 0.72$), in the identical (target), plausible, and implausible conditions, respectively.

2.3 | Task and procedure

Subjects were seated 60 cm away from a 20″ HP p1230 CRT monitor, with a refresh rate of 150 Hz and a screen resolution of 1024 × 768 pixels. After arriving, participants read and signed the informed consent. They were instructed to read sentences and to answer occasional yes-no comprehension questions. They answered the question by pressing the left or right button of a response controller, in order to answer affirmatively or negatively. After the EEG cap was set up and the eye tracker was calibrated, participants performed five practice trials before the real task, in order for them to get used to the experimental procedure.

During the task, a fixation point was presented in the center of the screen at the beginning of each trial in order to ensure that calibration of the eye tracker remained accurate. Then the experimenter started the trial, and a fixation box was presented on the left side of the screen, at the location of the beginning of the sentence. Once a fixation was detected in this box, the sentence was presented and stayed on the screen until the subject indicated that he had finished reading it by pressing a button on the response controller. They were also instructed to look at a target sticker located on the right side of the screen when they were done reading a sentence, to keep them from making additional eye movements that could have contaminated EM measures. When the reader's gaze crossed an invisible boundary located between the pretarget (n) and the previewed word ($n + 1$), a target word replaced the preview word, following the boundary paradigm (Rayner, 1975; see Figure 1). A "yes-no" question was presented after 30 of the sentences (23.8%). Accuracy on comprehension questions was high in all subjects ($M = 91.83\%$, $SD = 4.49\%$). After the experiment, participants were asked if they noticed any display or word change and, in case they noticed any change, they were asked if they recognized any previewed word. Participants reported little to no display or word changes after the experiment (below 5 trials) and no one reported recognizing a previewed word when there was a display change.

Condition	Log$_{10}$ freq		Orthographic similarity with target word		Cloze probability (%)	
	M	SD	M	SD	M	SD
Plausible preview word	1.5	0.7	0.13	0.15	1.58	3.32
Implausible preview word	1.4	0.8	0.14	0.17	0	0
Identical preview word (Target word)	1.4	0.8	–	–	2.93	2.89

TABLE 1 Descriptive statistics of the target and preview words used in the experiment

Stimuli from this experiment were intermixed with 144 sentences and 40 comprehension questions from another experiment (see M. Antúnez, S. Milligan, J. A. Hernández-Cabrera, H. A. Barber, & E. R. Schotter, in prep). 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.

2.4 | EEG and eye movements co-registration

EEG was recorded from 27 Ag/AgCl electrodes, following the 10/20 system (EasyCap, www.easycap.de). Four additional electrodes were placed in the external canthus of each eye and in the infra and supraorbital regions of the right eye. Electrodes were referenced online to the left mastoid and re-referenced offline to the algebraic mean of the right and left mastoids. The signal was amplified with a bandwidth of 0.01–100 Hz and a sampling rate of 500 Hz with the BrainVision system (www.brainproducts.com). Impedances were kept under 5 kΩ (electro-oculogram <10 kΩ).

EMs were recorded with a SR Research Ltd. Eyelink 1000 eye tracker in remote setup so that a target sticker was used to measure and control for head movements (Sampling rate = 500 Hz). Measures from the right eye were recorded, even though viewing was binocular. Calibration was performed on a standard five-point grid and eye position errors were less than 0.3° at each calibration point. Such calibration was performed not only at the beginning of the experiment but also during the task if calibration error was greater than 0.3°. Saccades crossing the invisible boundary activated the display change, which was completed almost immediately ($M = 5.38$ ms, $SD = 0.39$ ms).

2.5 | Processing

EMs were processed and inspected through SR Research DataViewer. On the first stage of pre-processing, fixation that were preceded or followed by blinks were discarded. Additionally, trials where a display change was triggered prior to the eye movement to the target word were removed from later analysis (5.8% of total data). Fixations on the pretarget and target interest areas were considered and exported for the analyses of interest. Only trials where readers fixated both the pretarget and target words during first-pass reading were kept and fixation durations shorter than 50 ms and greater than 800 ms

were excluded from analysis (retaining 82.14% of the total data trials).

The EEG data were pre-processed using the EEGLAB toolbox (Delorme & Makeig, 2004) for Matlab. The signal was filtered with a band-pass of 0.1–30 Hz and re-referenced offline to the average of the right and left mastoids. EMs were synchronized offline with the EEG signal with the EYE-EEG toolbox (Dimigen et al., 2011). Based on the trigger alignment, the mean synchronization error was below 1 ms. Independent components related to EMs were detected by using optimized ICA training data with overweighted spike potentials for better ocular artifact correction (Dimigen, 2020). Following Dimigen's (2020) guidelines, ICA was trained on band-pass filtered training data (at a passband edge of 2.5 Hz) and ocular components were removed with eye tracker-guided component identification (Plöchl et al., 2012), with a variance ratio threshold of 1.1. EEG data were segmented into two epochs of interest: −200 to 800 ms time-locked to the first fixation on the pretarget (n) and target ($n + 1$) words. Non-ocular artifacts were detected with a moving window peak to peak threshold of 100 μV and later visually inspected and rejected manually, in order to control for possible artifacts not detected automatically. After processing both EM and EEG data streams, only participants with at least 20 trials per condition were kept in the analyses to maximize signal to noise ratio.

2.6 | Analysis

For the EM data, we analyzed first fixation durations (duration of the first fixation made on a specific word during first-pass reading), single fixation durations (duration of the fixation made on a specific word, when there is only one fixation in first-pass reading), gaze durations (the sum of all fixations made on a specific word during first-pass reading before leaving it) and go-past time (the sum of all fixations on a specific word and subsequent fixations on words to the left of that word before fixating any word to the right of it) to assess early word processing. These measures were considered for both the pretarget (n) and target ($n + 1$) words, in order to study previous parafoveal processing and preview effects, respectively. Additionally, later word processing of the target word was assessed by analyzing total reading time (sum of all fixations on a word, including re-readings). Additionally, as in Schotter and Jia (2016), we analyzed fixation probability measures to better understand the effects of the preview word on the probability of fixating the target word during first-pass reading, the probability of making a regression out of the target word and re-reading words located to the left of it, and the probability of making a regression into the target

word from later words in the sentence. All the chosen measures are standard reading measures for the study of the time-course of word processing (Rayner, 1998).

For the electrophysiological measures, FRPs time-locked to the pretarget (n) and target ($n + 1$) words were also considered to study both previous semantic parafoveal processing while fixating the pretarget word and semantic preview effects when fixating the target word. Theoretically we expected to analyze time windows related to the N400 component, which should hold significant effects, based on our hypothesis. A mass univariate analysis was performed to select the specific time windows. More precisely, a point-by-point t-test analysis using the Guthrie-Buchwald approach (Guthrie & Buchwald, 1991) was performed for the whole epoch. The beginning and end of a time window would be defined by the beginning and end of, at least, 12 consecutive points with a significant t-test (Guthrie & Buchwald, 1991).

Electrodes were grouped into three clustered factors in the final mixed model analyses, in order to estimate the topographic distribution of effects. We followed the same topographic design as Barber et al. (2013), with the anteriority, laterality and hemisphere factors. The anteriority factor had 5 levels: frontal (Fz, F7, F3, F4, F8), frontocentral (FC5, FC1, FC2, FC6), central (Cz, T7, C3, C4, T8), centroparietal (CP5, CP1, CP2, CP6), parietal (Pz, P7, P3, P4, P8). The laterality factor had 2 levels: medial (F3, F4, FC1, FC2, C3, C4, P3, P4, CP1, CP2), lateral (F7, F8, FC5, FC6, CP5, CP6, P7, P8, T7, T8). Finally, the hemisphere factor had 2 levels: left (F7, F3, FC5, FC1, T7, C3, CP5, CP1, P7, P3) and right (F8, F4, FC2, FC6, T8, C4, CP2, CP6, P8, P4).

In order to more accurately observe the display change effect in the FRP signal, an additional analysis was performed in parieto-occipital electrodes (P7, O1, O2, P8). We based our analysis on the FRP study of Dimigen et al. (2012), where he found a display change preview effect from 170 to 252 ms in PO9 and PO10 electrodes in free reading of lists of words. The temporal window of choice was guided by the point-by-point t-test analysis, although we expected the effect to be present at a similar time-window as in the mentioned study.

All analyses were performed with R software (http://www.rproject.org), by using the ULLRToolbox (https://sites.google.com/site/ullrtoolbox/home). All EM measures and mean voltage from the selected time windows were analyzed using linear mixed effects models with the lme4 and lmerTest R packages (Bates et al., 2011, 2015; Kunzetsova et al., 2017). If a preferable maximal random effects model (Barr et al., 2013) did not converge, we reduced the random effects structure to include random intercepts for subjects and items and a random slope for the preview condition for subjects, followed by random

intercept for items and a random slope for subjects model. If none of these models converged, we reduced the structure to an only intercept for subjects and items random effects model. We used Satterthwaite's method to calculate the pooled degrees of freedom of the variances (Khuri et al., 1998; Satterthwaite, 1941). In case of the non-normality of the residuals of the estimated models, a scaled power (box-cox) transformation was performed with the estimated lambda of the model (Box & Cox, 1964; Fox & Weisberg, 2018). For fixation probability measures, the mixed model was conducted using a logistic link function.

For the eye movements analysis, orthogonal Helmert contrast comparisons were included in the mixed model, which were decided a priori based on the hypothesis described in the introduction. We compared the identical preview condition to the combination of plausible and implausible preview conditions, in order to look for display change effects. Additionally, we compared plausible and implausible conditions to each other, in order to look for pure preview plausibility effects. For the FRP analysis, we included the three clustered topographic factors ($5 \times 2 \times 2$) to explore the interaction of the main manipulation with scalp topography. We used the *anova* output of lmer and emeans (Lenth et al., 2018) packages to look at the contrasts at relevant topographical levels. Contrasts were performed with the emeans package and p values were adjusted with Hochberg's method (Hochberg, 1988). We report significant F and p values for the *anova* output for the topographical factors and b, t and p values of the fixed effects table.

3 | RESULTS

3.1 | Eye movements

For the EM analysis, we ran mixed models with random intercepts for items and subjects because the maximal model did not converge (Barr et al., 2013). Early reading time measures for fixation on the target words revealed that, compared to the implausible and plausible conditions, the identical condition led to shorter first fixation durations, single fixation durations, gaze durations, and go-past times (all $ps < .001$). Additionally, compared to the plausible condition, the implausible condition led to longer first fixation durations ($p < .05$), single fixation durations ($p < .01$), gaze durations ($p < .01$) and go-past times ($p < .001$). For fixations on the pretarget word, there were no differences in reading times between the different preview conditions.

Contrasts of total reading time spent on the target word revealed that the total time spent on the target word was shorter in the identical condition, compared to the

conditions where a display change took place ($p < .001$). Contrary to earlier reading measures, time spent on the target word was longer for the plausible condition than for the implausible condition, but the difference was not significant ($p = .38$; see Figure 2).

Participants had similar target fixation probability across conditions (both $ps > .05$). However, participants regressed out of the target word more often in the implausible compared to the plausible preview condition ($p < .01$). Additionally, they regressed back to the target word more often in the plausible condition than in the implausible condition ($p < .001$), and less often when there was no display change compared to the other two conditions ($p < .01$; see Tables 2 and 3).

3.2 | Fixation-related potentials

Pre-analysis stage. After performing the mass univariate analysis, we considered at least 12 consecutive compared points with significant p values to consider a temporal window for the mixed model analysis, following the Guthrie and Buchwald (1991) method. For the FRPs time-locked to the pretarget word, there was an effect of the preview between the 350 and 450 ms, reflecting the N400 component. For the FRPs time-locked to the target word, there were significant preview effects, which started at 100 ms and lasted until 400 ms. Because we were interested in studying the N400 component and to test the effects in earlier and later semantic and plausibility processing, we split the window and selected the 100–250 and 250–400 time windows for the mixed effects analyses (see Figures 3 and 4). For the analysis of the parieto-occipital electrodes, the preview effect was located between the 150 and 300 ms for the FRPs time-locked to the target word, so we chose that temporal window for the mixed model analysis (see Figure 5).

Mixed effects analysis. Because a maximal random model did not converge (Barr et al., 2013), we selected a model with random intercepts for subjects and items and
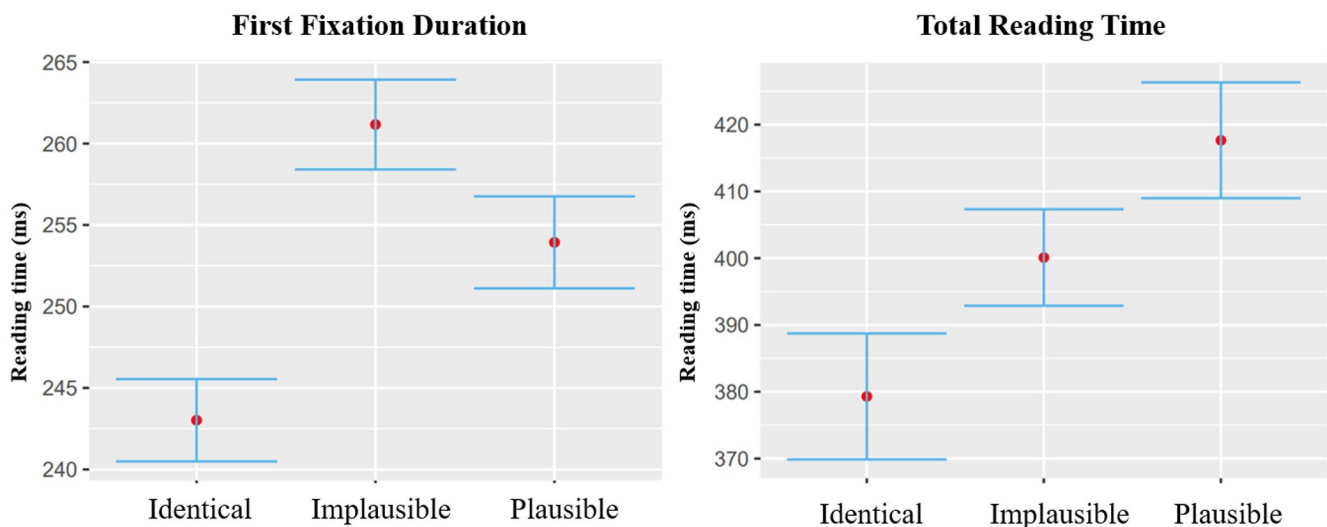

FIGURE 2 Early (left) and late (right) reading measures on the target word for the identical, plausible and implausible conditions

TABLE 2 Reading measures on the target word

Measures	Preview condition		
	Identical	Plausible	Implausible
First fixation duration	243.01 (2.73)	253.93 (2.68)	261.16 (2.73)
Single fixation duration	243.67 (3.3)	252.92 (3.51)	263.35 (3.69)
Gaze duration	262.5 (3.35)	277.87 (3.48)	288.4 (3.76)
Go-past time	321.76 (6.79)	330.02 (7.22)	354.58 (6.25)
Total time	379.29 (8.47)	417.64 (8.72)	400.09 (8.12)
Fixation probability	0.88 (0.01)	0.90 (0.01)	0.90 (0.01)
Regressions out of the target	0.12 (0.01)	0.11 (0.01)	0.14 (0.01)
Regressions into the target	0.24 (0.01)	0.31 (0.01)	0.25 (0.01)

Note: Mean and standard errors.

TABLE 3 Fixed effects of the contrasts of the linear mixed effects models for eye movements measures on the target word

Measures	Contrast	*b*	SE	|*t*|	*p*
First fixation duration	Identical versus Plausible + Implausible	−5.220	1.066	−5.059	<.001
	Plausible versus Implausible	−3.028	1.843	−2.008	<.05
Single fixation duration	Identical versus Plausible + Implausible	−4.876	0.988	−4.618	<.001
	Plausible versus Implausible	−4.041	1.686	−2.697	<.01
Gaze duration	Identical versus Plausible + Implausible	−7.300	1.320	−5.636	<.001
	Plausible versus Implausible	−4.702	2.253	−2.839	<.01
Go-past time	Identical versus Plausible + Implausible	−7.687	2.631	−4.459	<.001
	Plausible versus Implausible	−11.606	4.491	−4.138	<.001
Total time	Identical versus Plausible + Implausible	−10.913	3.278	−6.503	<.001
	Plausible versus Implausible	8.696	5.595	0.877	0.38

Measures	Contrast	*b*	SE	|*z*|	*p*
Fixation probability	Identical versus Plausible + Implausible	−0.05	0.033	−1.53	.124
	Plausible versus Implausible	0.007	0.059	0.13	.894
Regressions out of the target	Identical versus Plausible + Implausible	−0.007	0.036	−0.19	.8476
	Plausible versus Implausible	−0.172	0.062	−2.75	<.01
Regressions into the target	Identical versus Plausible + Implausible	−0.086	0.029	−2.97	<.01
	Plausible versus Implausible	0.175	0.048	3.624	<.001
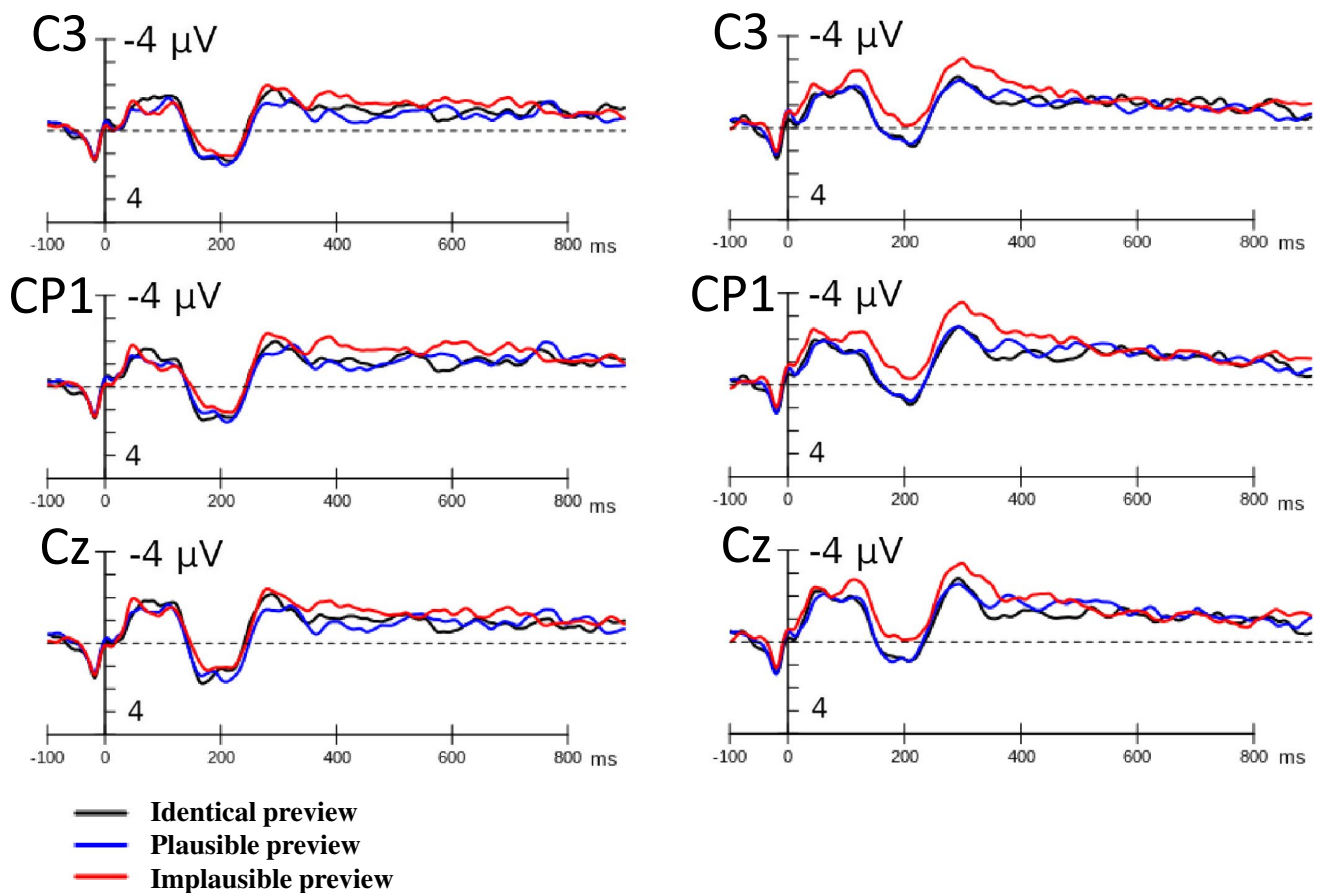
FIGURE 3 Grand average fixation-related potentials at the C3, CP1 and Cz electrodes for the fixation onset on the pretarget word (left) and on the target word (right) for the identical, plausible and implausible preview conditions
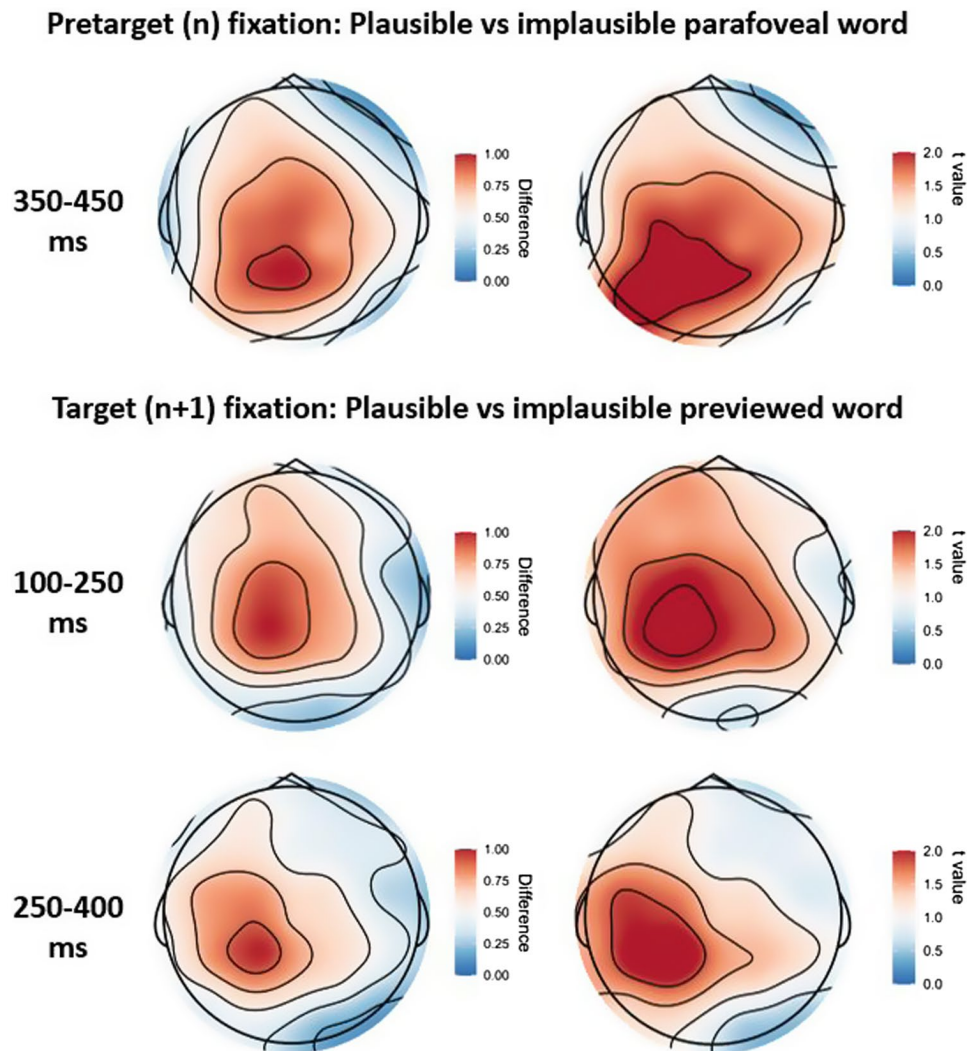
Pretarget (n) fixation: Plausible vs implausible parafoveal word

Target (n+1) fixation: Plausible vs implausible previewed word

FIGURE 4 Topographic maps for the fixation-related potentials time-locked to the target word, showing semantic parafoveal processing for the 350–450 ms temporal window during the fixation of the pretarget (n) word (up) and semantic preview effects for the 100–250 ms and 250–400 ms temporal windows during the fixation of the target (n + 1) word (down). The maps display the mean differences and t values of the comparison of plausible and implausible preview conditions

random slopes for the preview factor for subjects in both the FRP time-locked the pretarget and target words. The mixed effects analysis for the FRP time-locked to the pretarget word in the 350–450 temporal window revealed that the implausible condition showed a greater negativity when compared to the plausible condition ($t = 2.26$, $p < .05$), with no differences between the identical and the combination of plausible and implausible conditions ($t = -0.42$, $p = .66$) and no influence of the topographical factors.

Looking at the analysis of the FRP time-locked to the target word, in the 100–250 ms temporal window there was a greater negativity for the implausible condition compared to the plausible condition ($t = 2.6$, $p < .01$), with no differences between the identical and the combination of plausible and implausible conditions

($t = 1.5$, $p = .12$). In terms of interactions with the topographical factors, in the *anova* output of the lmerTest, the preview factor interacted with the laterality topographic factor $F(2, 52064) = 3.83$, $p < .05$ in the 100–250 ms temporal window, revealing that the difference between implausible and plausible conditions was mainly present in medial electrodes ($t = 3.2$, $p < .01$) but not in lateral electrodes ($t = 1.8$, $p = .14$). The analysis of the 250–400 temporal window showed that the implausible condition had a greater negativity compared to the plausible condition ($t = 2.1$, $p < .05$). Additionally, the identical condition was marginally different from the combination of implausible and plausible condition ($t = 1.8$, $p = .07$), with a reduced negativity in the 250–400 ms temporal window. In addition, for the analysis of the parieto-occipital electrodes in the 150–300 temporal
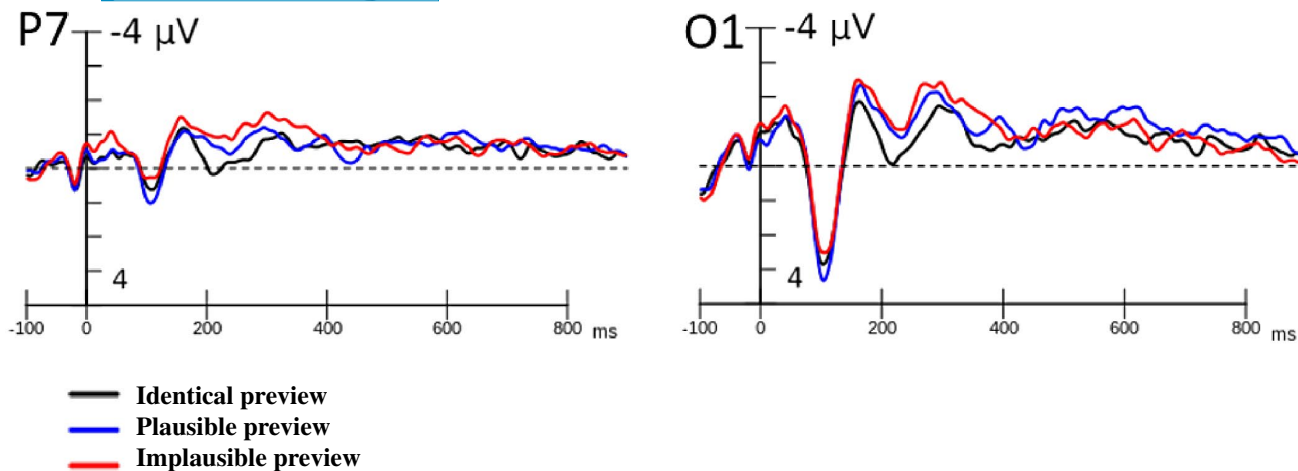
FIGURE 5 Grand average fixation-related potentials at the parieto-occipital (P7 and O1) electrodes for the fixation onset on the target word for the identical, plausible, and implausible preview conditions

window, we found that the identical condition had a greater positivity when compared to the combination of plausible and implausible conditions ($t = 3.35$, $p < .01$) while there was no difference between the plausible and implausible conditions ($t = 1.39$, $p = .17$).

4 | DISCUSSION

In this study, we explored the processing of semantic information perceived in the parafovea during sentence reading. We investigated whether semantic parafoveal information can be accessed and integrated during natural reading and aimed to describe the time-course of the parafoveal semantic effects. In order to do so, we used an EEG-EM co-registration set-up, which brought us two main benefits: (1) We were able to extract ERPs associated with fixations and therefore to study comprehension processes during a free reading task. In this way, we were able to test if previous EEG effects obtained in highly controlled RSVP paradigms are also observed in a more ecologically valid situation. (2) We simultaneously obtained two online measures of the processes related to reading that previously have resulted in slightly different interpretations. This allowed us to track the time course of sentence comprehension integrating two complementary approaches. Furthermore, we used the invisible boundary paradigm to manipulate parafoveally previewed words with varying plausibility. The plausibility manipulation allowed us to study the semantic processing in the parafovea independently of orthographic predictions.

Exploring the time-course of plausibility preview effects, we expected first-pass reading durations of the target words being affected by the semantic manipulation of the previews. This result would fit with amplitude modulations on early time windows in FRPs time-locked to the target ($n + 1$) word. However, based on previous EEG studies (Barber et al., 2013; Kretzschmar et al., 2009; Stites et al., 2017) we also expected effects in later time windows (e.g., N400 effects) in spite of the fact that previous EM evidence has suggested that plausibility preview effects are short-lived (Schotter & Jia, 2016; Veldre & Andrews, 2016, 2017, 2018c). Additionally, we expected to find a modulation in the N400 component in FRPs time-locked to the pretarget (n) word, replicating flanker-RSVP-ERP evidence of semantic parafoveal processing in the absence of EMs (Barber et al., 2013; Stites et al., 2017).

4.1 | Eye-tracking results

Eye Movements measures related to the target word show a display change effect, consistent with trans-saccadic integration processes. Readers obtained a benefit from the identical preview compared to the display change conditions in both first-pass reading duration and total reading duration measures, reflecting the cost of integrating previews unrelated to the target word (for a review, see Schotter et al., 2012). As previously mentioned, the display change preview effect may be a mixture of preview benefits and preview costs (Kliegl et al., 2013), which may be related to cost of a perceptual dissimilarity between previews and targets (see Hutzler et al., 2013, 2019). More interestingly, first-pass reading measures (but not total reading duration) revealed longer fixation durations for the implausible condition compared to the plausible condition, consistent with recent evidence of preview plausibility effects (Schotter & Jia, 2016; Veldre & Andrews, 2016, 2017, 2018c; see Andrews & Veldre, 2019). Both plausible and implausible unrelated previews shared little to no orthographic or semantic features with the target

word, suggesting this effect is not due to trans-saccadic integration. Instead, such effects could only be explained by a contextual fit account, in which integrating the implausible previews to the sentence context would have a greater cost. This supports the idea that both contextual and trans-saccadic integration processes are independent of each other, probably operating at different levels (Veldre & Andrews, 2017). It is important to clarify that such preview plausibility effects were not influenced by anticipatory predictions of upcoming words, since all previews and targets had extremely low cloze probability values (i.e., below 2%–3%). Thus, these findings were not enhanced by a facilitatory effect of predictability (Staub, 2015).

4.2 | Fixation-related potentials: display change effects

The analysis of the FRP time-locked to the target word for the parieto-occipital electrodes revealed a preview effect related to the display change between the 150 and 300 ms. More specifically, a greater positivity in this temporal window was found when the previewed word was identical, compared to when the preview word was different from the target word. Our findings replicate previous evidence from Dimigen et al. (2012), who found that identical previews lead to facilitatory effects reflected in shorter fixation durations and more positive amplitude that emerged from around 170 ms to 280 ms in the PO9 and PO10 electrodes, compared to the other conditions where a display change was present. As they also indicated, both their and our findings in fixation durations and FRPs amplitudes may support the classic idea that the display change effect is related to a pre-activation of orthographic codes before lexical access, an idea established from both EM (Rayner, 1998) and ERP research in visual word recognition paradigms (see Barber & Kutas, 2007). Additionally, the analysis of all electrodes revealed modulations of the N400 component, being the identical condition marginally less negative than the average of the implausible and plausible unrelated conditions, consistent with a later facilitation effect in FRPs for valid previews with the boundary paradigm (Li et al., 2015; López-Pérez et al., 2016). Following Dimigen et al. (2012), the N400 attenuation derived from a valid preview could be equivalent to the repetition priming effect derived from visual word recognition studies (see Holcomb, & Grainger, 2006, 2007), which could suggest that similar mechanisms of trans-saccadic integration of low-level features are involved in both word recognition and natural sentence reading paradigms. The full repetition priming effect involves the activation of words at multiple levels since the activation of form and

orthography leads to higher levels like phonology, morphology and semantics (see Barber & Kutas, 2007). This would explain the early activation of orthographic codes and the late activation reflected in the N400, related to a semantic access of words (Kutas & Federmeier, 2011), which would be consistent with the behavior of readers reflected in both, ours and previous, EM data in natural sentence reading (see Schotter et al., 2012). Having said that, there could be additional processes related to word activation taking place during the facilitation (or costs) derived from display change effects. For instance, word-identification processing in the conditions where previews and targets are dissimilar could be restarted after fixating the target word, whereas this processing would start without change as the fixation of the pretarget word in the condition where the preview never changes. Thus, the N400 could also be related to changes in the time-course of the word-identification process. More refined experimental designs, in combination with electrophysiological measures, may be used in the future in natural sentence reading to disentangle the processes related to perceptually dissimilar previews and the word identification process.

4.3 | Fixation-related potentials: semantic and plausibility effects

Moving on to semantic and plausibility effects in the electrophysiological record, FRPs time-locked to the pretarget word revealed semantic parafoveal processing reflected in the modulation of the N400 component. Specifically, the implausible condition showed a greater negativity when compared to the plausible condition, meaning that a contextually implausible preview had greater processing costs. Hence, during natural reading, words located in the parafoveal region are semantically accessed and their meaning interacts with sentence-level context information. This finding could be considered complementary to EM literature. Despite the fact that semantic parafoveal effects during the fixation of the pretarget words had been mostly absent in EM experiments with the boundary paradigm (for a review, see Schotter et al., 2012), more recent studies have reported preview plausibility effects on skipping the target word (Veldre & Andrews, 2017, 2018a, 2018b, 2018c; Veldre et al., 2020), which also demonstrates some access to the meaning of parafoveal words. Furthermore, the semantic processing of parafoveal words during the fixation of the pretarget word replicates previous evidence of semantic parafoveal processing from ERPs in more artificial reading situations and in visual word recognition paradigms (Barber et al., 2010, 2011; Li et al., 2015; López-Pérez et al., 2016; Snell et al., 2019).

Our data suggest that these prior findings can be extended to more naturalistic reading situations, supporting the idea that evidence from artificial reading situations (e.g., flankers-RSVP or word-pairs paradigms) may be valid for drawing conclusions about what may be happening in sentence reading. Even though Kretzschmar et al. (2009) previously reported semantic modulations in the N400 component during the fixation of a pretarget word, they were associated with semantically incongruent parafoveal words compared to congruent predictable words in highly constraining sentence constructions, which may suggest, as they also pointed out, that their findings may be better described as orthographic in nature, rather than semantic. Additionally, the use of the boundary paradigm in this experiment ensures that any semantic effects are derived from the parafoveal word and not from the meaning of the target word, isolating parafoveal effects. More interestingly, these findings also extend pure plausibility effects from flanker-RSVP-ERPs paradigms without the presence of EM (Stites et al., 2017) to natural sentence reading. Both their and our findings of plausibility parafoveal processing confirm that semantic access of parafoveal words is independent from orthographically related predictions of the upcoming words, in both artificial and natural reading scenarios. In addition, such consistency between EEG measures from different paradigms would support the idea that semantic and plausibility parafoveal processing during the fixation of the pretarget word involves cognitive mechanisms that are independent from saccade programming. It should be mentioned that in our sentences, pretarget words and previews were not formally or semantically related. Consequently, effects at the time of the processing of the pretarget cannot be explained as a facilitation of the parafoveal information (preview word) over the foveal processing. Instead, our results show that the parafoveal meaning activation is modulated by the previous sentence context. Previous ERP studies did not find evidence of independent processing of foveal and parafoveal words, especially attending to the morphology and latencies of the foveal and parafoveal N400 effects (Barber et al., 2013). Keeping that in mind, the pretarget effect in our study is compatible with the idea that the parafoveal word is pre-processed when previewed to a stage where its meaning is activated and integrated with the context resulting in a cost if the preview is implausible.

Looking at the FRPs time-locked to the target word in the 100–250 ms temporal window, analyses revealed more negative amplitudes in the implausible condition when compared to the plausible condition. These results align with our findings in first-pass reading measures, revealing a greater processing cost when a contextually implausible word was previewed before fixating the target word. These results replicate previous FRPs findings

of López-Pérez et al. (2016) of early semantic preview effects, but contrast with Antúnez et al. (2021), who only found such effects in the N400 component in FRPs time-locked to the target word. In both studies, readers had to read word-pairs in Spanish and had to indicate if they were semantically related or not. With the invisible boundary paradigm, López-Pérez et al. (2016) manipulated the previewed word, so it was semantically related or not to the pretarget word. In Antúnez et al. (2021), readers were Basque-Spanish bilinguals and the preview was either a Basque non-cognate translation of the Spanish target word or a totally unrelated Basque word. A possible explanation of the replicability between these FRPs studies in word-pair paradigms and our results may be related to the presence of N400-related semantic parafoveal effects during the fixation of the pretarget word. In our study, similarly to the semantic parafoveal-on-foveal effects reported by López-Pérez et al. (2016), we found a previous activation of the meaning of the parafoveal words during the fixation of the pretarget word, followed by the early semantic preview effect during the fixation of the target word. In contrast, Antúnez et al. (2021) failed to find either parafoveal-on-foveal or early preview benefit effects. A potential explanation is that previous activation of semantic parafoveal information during the fixation of the pretarget word could have boosted early plausibility preview effects during the fixation of the target word. It is important to mention that, while exploring sequences of ERPs and FRPs that are closely related in time, we have to assume some overlapping activity between two temporally adjacent electrophysiological effects (i.e., the N400 during the fixation in the pretarget word and the early semantic activation during the fixation in the target word). In fact, even though both effects are displaying partially independent semantic-plausibility processing, both semantic effects may be highly interactive and it is difficult at first to confirm which one of the two effects contributes the most to the final waveform. However, since the semantic access of words happens between 300 and 500 ms, and our semantic effects begin as early as 100 ms after fixating the target word, it is highly likely that the semantic access was triggered by the previewed parafoveal word during the fixation of the pretarget word, which would have lasted approximately 250 ms. Therefore, it would make sense that the early semantic effects when fixating the target word actually reflect a modulation of the N400 initiated during the fixation of the pretarget word. Interestingly enough, the discrepancies between Antúnez et al. (2021) and both ours and López-Perez et al.'s (2016) findings may be owed to the fact that they used a more subtle semantic manipulation (i.e., preview non-cognate translations of

target words in a bilingual sample). Moreover, while we manipulated the contextual fit of parafoveal words and López-Pérez et al. (2016) manipulated the semantic relationship between preview and pretarget words, Antúnez et al. (2021) focused on manipulating the preview-target semantic relationship across languages, which may have reduced the semantic parafoveal effects during the fixation of the pretarget word. Future research may explore different semantic parafoveal manipulations in FRP studies in natural sentence-reading paradigms.

In line with our hypothesis, modulations of the N400 time-locked to the target word also revealed a greater processing cost for the implausible condition compared to the plausible condition. This replicates the N400 findings of semantic preview effects in FRP studies (Antúnez et al., 2021; López-Perez et al., 2016), indicating that the meaning of parafoveal words is accessed and used to facilitate the consequent processing of target words. We present evidence for the first time of semantic preview effects in the electrophysiological record in natural sentence reading. Because our results are consistent with the word-pair paradigms of Antúnez et al. (2021) and López-Perez et al. (2016), FRP studies with controlled-reading paradigms may be valid for drawing conclusions about how the meaning of parafoveal words are accessed and integrated in natural sentence reading. Having said that, it would be important to point out that the semantic modulation of the N400 found here involved contextual integration processes, rather than a semantic integration between preview and target words like in word-pair paradigms, so future research should add further control to better understand the role of high-order semantic parafoveal processing in word recognition and natural sentence-reading paradigms.

4.4 | Time-course of parafoveal semantic processing

The findings related to semantic N400 modulations in the FRPs time-locked to the target word are of great assistance when delving into the time-course of plausibility preview effects. The modulation in the temporal window of the N400 component adds an important breakthrough when combined with our EM measures and with previous EM evidence suggesting that processing difficulties of the target word after previewing a contextually implausible unrelated word could be short-lived (see Andrews & Veldre, 2019). More specifically, our results from EM alone, in line with previous evidence, are compatible with the idea that plausibility effects are limited to early processing, as revealed by first-pass reading measures (Schotter & Jia, 2016; Veldre

& Andrews, 2016, 2017, 2018c). No significant differences were found between plausible and implausible previews in total reading time, which could suggest that plausibility may not affect later processing. However, even though not significant, the patterns in EMs differ from early to late processing, since participants had longer total reading times in the plausible condition. This could be explained by the greater probability of regressing into the target in the plausible condition, as our results show, which could have diluted the greater cost derived from the implausible condition. Another interpretation is that early saccadic planning may be benefited by plausible previews, having an impact across the full distribution of early measures on the target ($n + 1$) word, such as single fixation duration (Veldre et al., 2020). Having said that, the display change in the plausible preview may have caused an interference detected after fixating the target word, causing later disruption. Nevertheless, readers were still more prone to regress to the region before the target word in the implausible condition, suggesting that later costs linked with difficulties in contextual integration processes. More importantly, the modulation of the N400 component confirms that plausibility effects are still present during later processing, following a different trend than our findings in EMs. One possible explanation is that, as the temporal window of the N400 is not affected by the greater probability for readers to regress into the target in the plausible condition, the electrophysiological measure would be more suitable to isolate the effects derived from greater integration costs after an implausible preview. Therefore, as opposed to previously suggested in the literature, the high-order integration of the previewed word with the sentence context affects both early and later processing of the target word and reading behavior. Finally, the different (but complementary) trends between FRPs and EM measures would also suggest that later plausibility preview effects are linked to different cognitive processing mechanisms related to semantic processing and oculomotor behavior, making it necessary to use a co-registration set-up to better understand the time-course of parafoveal effects.

4.5 | Methodological considerations

In line with our question of interest, our results also raise new concerns and questions about the study of semantic parafoveal processing and the experimental paradigm of choice. As previously discussed, most neural evidence for semantic parafoveal processing has come from electrophysiological studies using more artificial reading paradigms (derived from ERP designs), where

EMs are absent. Our results show that artificial tasks that allow the adequate experimental control of multiple variables are necessary and recommended for isolating questions and consolidating hypotheses. However, the co-registration technique in natural reading situations offers a number of advantages to be considered. First, and probably the most obvious one, is the ecological validity associated with recording electrophysiological brain activity during a natural sentence reading situation where oculomotor behavior is present. Second, the patterns of the FRP signal in free-viewing reading vary from more classical ERP paradigms (for a review, see Degno & Liversedge, 2020). For instance, the modulation of the N400 component in the fixation of the target word here was found between the 250–400 ms, reflecting an earlier onset of the effect than in semantic preview studies with more artificial paradigms, in which it arises between 300–500 ms (Antúnez et al., 2021; Dimigen et al., 2012, López-Perez et al., 2016). Our earlier onset of semantic parafoveal effects in natural sentence-reading scenarios is not completely surprising, though. Previously, Kornrumpf et al. (2016) manipulated the preview of the upcoming word by changing the number of visible letters in the parafoveal region in two different scenarios: a flanker-RSVP and a free-reading of word-lists. They found a preview effect starting at 230 ms under flanker-RSVP reading, but the same effect started at 160 ms in natural reading of lists of words, which, in combination with our findings, raises new questions about the nature and timing of processing during reading. For example, our earlier N400 onset may reflect a more accurate estimation of word recognition processes during natural reading. Finally, co-registration of EMs and EEG provides additional insight into natural reading studies where only EMs are recorded. For instance, EM research uses direct measures of reading behavior to infer what kind of cognitive processing is taking place. However, as the evidence provided here suggests, not all cognitive activity modulates behavior in the same way, so electrophysiological measures can capture some information that goes undetected by behavioral EM measures. Exploring the consistency between FRPs and EM measures may offer a clearer explanation of which processes are linked to oculomotor behavior and which ones may be independent from them.

Having said that, all the advantages of the co-registration mentioned here come at some cost. First, we may face a greater loss of trials owing to the filtering and processing criteria of both EM and FRPs data streams, which may lead to more difficulties in detecting smaller effects owing to less statistical power. Additionally, we may face an overlapping activity from multiple fixations that may contaminate the effects of interest. A recent alternative approach is related to non-linear deconvolution models, where regression-ERPs can be extracted before entering them into a second-level group analysis (see Dimigen & Ehinger, 2021). Here, we applied linear mixed-effects models as they have important advantages, such as including crossed random effects for subjects and items, which allows us to analyze trial-level data rather than average across participants first. Having said that, computational and statistical research may focus on integrating both deconvolution and mixed-effects models to obtain the advantages of both approaches. Future lines of research may take into account these considerations when co-registering EM and FRPs in natural sentence reading, so they may use a tool with several advantages that provides complementary evidence of the multiple cognitive processes taking place during reading.

5 | CONCLUSIONS

In summary, we provided evidence for the first time of semantic-plausibility parafoveal processing in natural sentence reading in the electrophysiological record during the fixation of both the pretarget and target words. By using an EEG-EM co-registration set-up in a more ecological reading situation, our findings support the validity of previous highly controlled reading paradigms. Importantly, both complementary data-streams allowed us to disentangle the time-course of parafoveal semantic access determined by sentence-level context information. The co-registration technique during natural reading may be of great assistance in the study of the cognitive mechanisms involved in semantic parafoveal processing.

AUTHOR CONTRIBUTIONS
Martín Antúnez: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Resources; Visualization; Writing—original draft; Writing—review & editing. **Sara Milligan:** Conceptualization; Investigation; Writing—review & editing. **Juan Andrés Hernández-Cabrera:** Formal analysis; Methodology; Software. **Horacio A. Barber:** Conceptualization; Funding acquisition; Investigation; Project administration; Resources; Supervision; Writing—review & editing. **Elizabeth R. Schotter:** Conceptualization; Funding acquisition;

Investigation; Project administration; Resources; Supervision; Writing—review & editing.

ORCID

Martín Antúnez https://orcid.org/0000-0001-7131-1461

REFERENCES

Altarriba, J., Kambe, G., Pollatsek, A., & Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish-English bilinguals. *Perception & Psychophysics*, *63*(5), 875–890. https://doi.org/10.3758/BF03194444

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), e12344. https://doi.org/10.1111/lnc3.12344

Antúnez, M., Mancini, S., Hernández-Cabrera, J., Hoversten, L., Barber, H., & Carreiras, M. (2021). Cross-linguistic semantic preview benefit in basque-spanish bilingual readers: Evidence from fixation-related potentials. *Brain and Language*, *214*, e104905. https://doi.org/10.1016/j.bandl.2020.104905

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., & 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., 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

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). *Fitting linear mixed effects models using, lme4*. 67, 48. arXiv preprint arXiv:1406.5823. http://hdl.handle.net/10.18637/jss.v067.i01

Bates, D., Maechler, M., & Bolker, B. (2011). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999375-42.

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*(2), 211–243. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Degno, F., & Liversedge, S. P. (2020). Eye movements and fixation-related potentials in reading: A review. *Vision*, *4*(1), 11. https://doi.org/10.3390/vision4010011

Degno, F., Loberg, O., Zang, C., Zhang, M., Donnelly, N., & Liversedge, S. P. (2019a). A co-registration investigation of inter-word spacing and parafoveal preview: Eye movements and fixation-related potentials. *PLoS One*, *14*(12), e0225819. https://doi.org/10.1371/journal.pone.0225819

Degno, F., Loberg, O., Zang, C., Zhang, M., Donnelly, N., & Liversedge, S. P. (2019b). 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*, e116117. 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. 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. https://doi.org/10.1037/a0023885

Fox, J., & Weisberg, S. (2018). Visualizing fit and lack of fit in complex regression models with predictor effect plots and partial residuals. *Journal of Statistical Software*, *87*(1), 1–27. https://doi.org/10.18637/jss.v087.i09

Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology*, *28*(2), 240–244. https://doi.org/10.1111/j.1469-8986.1991.tb00417.x

Hochberg, Y. (1988). A sharper bonferroni procedure for multiple tests of significance. *Biometrika*, *75*(4), 800–802. https://doi.org/10.1093/biomet/75.4.800

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1150. https://doi.org/10.1037/a0020233

Holcomb, P. J., & Grainger, J. (2006). On the time course of visual word recognition: An event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience*, *18*(10), 1631–1643. https://doi.org/10.1162/jocn.2006.18.10.1631

Holcomb, P. J., & Grainger, J. (2007). Exploring the temporal dynamics of visual word recognition in the masked repetition priming paradigm using event-related potentials. *Brain Research*, *1180*, 39–58. https://doi.org/10.1016/j.brainres.2007.06.110

Hutzler, F., Fuchs, I., Gagl, B., Schuster, S., Richlan, F., Braun, M., & Hawelka, S. (2013, July). Parafoveal X-masks interfere with foveal word recognition: Evidence from fixation-related brain potentials. *Frontiers in Systems Neuroscience*, *23*(7), 33. https://doi.org/10.3389/fnsys.2013.00033

Hutzler, F., Schuster, S., Marx, C., & Hawelka, S. (2019, February 28). An investigation of parafoveal masks with the incremental boundary paradigm. *PLoS One*, *14*(2), e0203013. https://doi.org/10.1371/journal.pone.0203013

Hyönä, J., & Häikiö, T. (2005). Is emotional content obtained from parafoveal words during reading? An eye movement analysis.

Scandinavian Journal of Psychology, *46*(6), 475–483. https://doi.org/10.1111/j.1467-9450.2005.00479.x

Khuri, A., Mathew, T., & Sinha, B. (1998). *Statistical tests for linear mixed models*. Wiley.

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*, 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. https://doi.org/10.1037/xlm0000128

Kunzetsova, A., Brockhoff, P., & Christensen, R. (2017). lmertest package: Tests in linear mixed effect models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the n400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Laszlo, S., & Federmeier, K. D. (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language*, *61*(3), 326–338. https://doi.org/10.1016/j.jml.2009.06.004

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R Package Version*, *1*(1), 3.

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

López-Pérez, P., Dampuré, J., Hernández-Cabrera, J., & Barber, H. (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

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. (2016). Out of the corner of my eye: Foveal semantic load modulates parafoveal processing in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(11), 1839. https://doi.org/10.1037/xhp0000253

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*, 278. https://doi.org/10.3389/fnhum.2012.00278

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. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology/Revue Canadienne De Psychologie*, *40*(4), 473. https://doi.org/10.1037/h0080111

Rayner, K., & Schotter, E. R. (2014). Semantic preview benefit in reading English: The effect of initial letter capitalization. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1617. https://doi.org/10.1037/a0036763

Satterthwaite, F. E. (1941). Synthesis of variance. *Psychometrika*, *6*(5), 309–316. https://doi.org/10.1007/BF02288586

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. In K. Federmeier & D. Watson (Eds.), *Psychology of learning and motivation* (Vol. *68*, pp. 263–298). Elsevier. https://doi.org/10.1016/bs.plm.2018.08.011

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., & 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. https://doi.org/10.1037/xlm0000281

Schotter, E. R., Lee, M., Reiderman, M., & Rayner, K. (2015). The effect of contextual constraint on parafoveal processing in reading. *Journal of Memory and Language*, *83*, 118–139. https://doi.org/10.1016/j.jml.2015.04.005

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. https://doi.org/10.1037/xhp0000270

Schuster, S., Himmelstoss, N. A., Hutzler, F., Richlan, F., Kronbichler, M., & Hawelka, S. (2021). Cloze enough? Hemodynamic effects of predictive processing during natural reading. *NeuroImage*, *228*, e117687. https://doi.org/10.1016/j.neuroimage.2020.117687

Snell, J., Meade, G., Meeter, M., Holcomb, P., & Grainger, J. (2019). An electrophysiological investigation of orthographic spatial integration in reading. *Neuropsychologia*, *129*, 276–283. https://doi.org/10.1016/j.neuropsychologia.2019.04.009

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass*, *9*(8), 311–327. https://doi.org/10.1111/lnc3.12151

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/s1341 5-016-0492-6

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. https://doi. org/10.1037/xhp0000200

Veldre, A., & Andrews, S. (2017). Parafoveal preview benefit in sentence reading: Independent effects of plausibility and orthographic relatedness. *Psychonomic Bulletin & Review*, *24*(2), 519–528. https://doi.org/10.3758/s13423-016-1120-8

Veldre, A., & Andrews, S. (2018a). Beyond cloze probability: Parafoveal processing of semantic and syntactic information during reading. *Journal of Memory and Language*, *100*, 1–17. https://doi.org/10.1016/j.jml.2017.12.002

Veldre, A., & Andrews, S. (2018b). How does foveal processing difficulty affect parafoveal processing during reading? *Journal of Memory and Language*, *103*, 74–90. https://doi.org/10.1016/j. jml.2018.08.001

Veldre, A., & Andrews, S. (2018c). Parafoveal preview effects depend on both preview plausibility and target predictability. *Quarterly Journal of Experimental Psychology*, *71*(1), 64–74. https://doi. org/10.1080/17470218.2016.1247894

Veldre, A., Reichle, E. D., Wong, R., & Andrews, S. (2020). The effect of contextual plausibility on word skipping during reading. *Cognition*, *197*, 104184. https://doi.org/10.1016/j.cognition.2020.104184

White, S. J., Bertram, R., & Hyönä, J. (2008). Semantic processing of previews within compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 988. https://doi.org/10.1037/0278-7393.34.4.988

Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Readers of Chinese extract semantic information from parafoveal words. *Psychonomic Bulletin & Review*, *16*(3), 561–566. https://doi. org/10.3758/PBR.16.3.561