Originally published in: Language, Cognition and Neuroscience vol. 38 (2023) no. 6, pp. 821-842.
DOI: *https://doi.org/10.1080/23273798.2022.2159990*

POSTPRINT

Jon Burnsky[a], Franziska Kretzschmar[b], Erika Mayer[c] and Adrian Staub[a]

[a]University of Massachusetts Amherst, Department of Psychological and Brain Sciences, Amherst, USA;
[b]Leibniz Institute for the German Language, Mannheim, Germany;
[c]University of Massachusetts Amherst, Department of Linguistics, Amherst, USA

The influence of predictability, visual contrast, and preview validity on eye movements and N400 amplitude: co-registration evidence that the N400 reflects late processes

Abstract: Two eye movement/EEG co-registration experiments investigated effects of predictability, visual contrast, and parafoveal preview in normal reading. Replicating previous studies, in Experiment 1 contrast and predictability additively influenced fixation durations, and in Experiment 2 invalid preview eliminated the predictability effect on early eye movement measures. In both experiments, predictability influenced the amplitude of the N400 component of the fixation-related potential. In Experiment 1, visual contrast did not influence the N400, and in Experiment 2, the effect of predictability on the N400 was larger with invalid preview, in opposition to the eye movement pattern. The N400 may reflect a late process of accessing conceptual representations while the duration of the eyes' fixation on a word is sensitive to the difficulty of perceptual encoding and early stages of word recognition. The effects of predictability on both fixation duration and the N400 suggest an influence of this variable at two distinct processing stages.

Keywords: Eye movements; N400; predictability; word recognition; reading

A word's predictability in its context, as measured by cloze probability, has reliable effects in the dominant experimental paradigms used to study incremental processing of written sentences. In eye movement studies, readers spend less time fixating a predictable word than an unpredictable one, and readers' eyes are also more likely to skip a predictable word (for review see Staub, 2015). In event-related potential (ERP) studies, there is a robust effect of predictability on the N400, a negative-going component of the waveform that peaks at about 400 ms post-stimulus onset, with the amplitude of this component being reduced for predictable compared to unpredictable words (for review see Kutas & Federmeier, 2011; Federmeier, 2022). Most ERP studies with visual presentation of stimuli have elicited this effect using word-by-word presentation of sentences, with each word appearing at central fixation, and with stimulus onset asynchrony (SOA) of at least several hundred milliseconds – sometimes up to 700 ms – between words; see Dambacher et al. (2012) for explicit investigation of how SOA modulates the predictability effect. However, a few previous studies (Dimigen et al., 2011; Kretzschmar et al., 2015) have investigated predictability effects by recording eye movements and EEG simultaneously, time-locking the EEG to the onset of fixation on the critical word (in which case the term *fixation-related potential*, or FRP, is used in place of ERP). These studies have successfully elicited predictability effects on both eye movements and the amplitude of the N400. The present study also adopts this co-registration paradigm.

Because eye fixation duration and N400 amplitude are both sensitive to a word's predictability, as well as to other variables such as a word's plausibility in context (e.g. Abbott & Staub, 2015; Nieuwland et al., 2020), it is intuitive that the two measures may be behavioral and electrophysiological indices, respectively, of a common set of underlying neural and cognitive processes. However, there are salient discrepancies between the effects of predictability on eye movements

Contact: Adrian Staub, astaub@psych.umass.edu, University of Massachusetts Amherst, Department of Psychological and Brain Sciences, Amherst, USA

Earlier versions of this work have been presented at the 2019 European Conference on Eye Movements, Alicante, Spain, and the 2021 CUNY Conference on Human Sentence Processing, University of Pennsylvania. Thanks to Serene Elbach, Megan Agro, Dawn Bissessar, Eunji Choi, Alex Green, Amber Hughes, Subin Jung, Colin McDermott, Leo Stern, Robin Striar, Tiffini Tam, Cora Vitale, Tiffini Vo, and Kuan-Jung Huang for assistance with data collection, and Lisa Sanders for the use of the NCAP lab.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/23273798.2022.2159990.

and ERPs, which we discuss below. Thus, in their influential review of the N400 literature, Kutas and Federmeier (2011) listed among the questions shaping the "future directions" of ERP research, "What is the relationship (associations and dissociations) between N400 measures and reaction time measures and eye movement measures?" The present study addresses this question by investigating whether and how manipulations of visual contrast and parafoveal preview validity influence N400 amplitude during natural reading; both of these manipulations are known to influence fixation duration, and in the case of the latter manipulation, to interact with the effect of predictability in influencing fixation duration. In the remainder of this Introduction, we first describe existing research suggesting that in contrast to eye fixation duration, the N400 may not be sensitive to manipulations that influence early stages of perceptual encoding and visual word recognition. We then describe and motivate the present experiments.

The effect of predictability on eye movements in reading has a fairly transparent interpretation as an effect on the difficulty of early stages of lexical processing. In English, a typical eye fixation duration in reading is in the range of 225–250 ms; the first (or in many cases, only) eye fixation on a highly predictable word will tend to be about 20 ms shorter than on an unpredictable one. Thus, predictability influences the amount of time that a reader spends on a word before moving on, leading to the inference that a predictable word requires less cognitive work to recognise and/or to integrate into a sentence-level syntactic and semantic representation. But because predictability influences the very earliest eye movement measures, including the probability that a word is skipped by the eyes on the basis of parafoveal pre-processing, both older (e.g. E-Z Reader and SWIFT; Reichle et al., 2003; Engbert et al., 2005) and more recent (e.g. Smith & Levy, 2013) models concur that the effect of predictability on eye movements should be understood as an effect on the difficulty of word recognition itself, as opposed to only an effect on later integrative stages. See Staub (2015) for a review of the literature on predictability effects on eye movements in reading, as well as discussion of the theoretical interpretation of these effects.

On the other hand, the interpretation of N400 effects has been highly contentious, with a long history of debate between interpretations of the reduction in N400 amplitude with predictable words as an effect on the difficulty of lexical access itself, and as an effect on the difficulty of post-lexical integration of a word into its sentence context (see, e.g. Kutas & Federmeier, 2011; Lau et al., 2008; Nieuwland et al., 2020; it is important to note that this simple dichotomy does not do justice to the nuance of the positions in the literature). With respect to the relationship of N400 effects to eye movement effects, it is notable, first, that it is generally the amplitude, not the latency, of the N400 that is modulated by predictability, and by other variables such as plausibility; the latency of the N400 is, in the words of Kutas and Federmeier (2011), "remarkably constant," with the peak generally appearing at about 400 ms post-stimulus-onset (cf. Dambacher et al., 2012). Thus, while eye movement models (e.g. Engbert et al., 2005; Reichle et al., 2003) have interpreted the effect of predictability as an effect on the difficulty of, and therefore the duration of, some lexical or pre-lexical processing stage(s), the fact that predictability influences the amplitude of the N400, rather than its latency, suggests an effect on the nature of the processing that occurs at a specific time, relative to when a word is first encountered, rather than an influence on how long some processing stage(s) require for completion.

Second, the N400 is very late relative to the timeline of normal reading. By 400 ms after the eyes have fixated a word, the reader's inspection of that word is usually long finished, with the eyes now fixated on a subsequent word (Rayner & Clifton Jr, 2009). This timing discrepancy is made even more striking when one considers the additional time that is required to plan and execute a saccade; the 225–250 ms mean duration of a reader's first fixation, which is influenced by predictability, includes the approximately 100 ms that is required to plan and execute the saccade that will end this fixation (Reichle et al., 2003). Thus, it appears that the effect of predictability on fixation durations must actually occur within the first 150 ms of fixation on a word. Moreover, as noted above, predictability also influences the even earlier measure of word skipping, which indicates that this variable influences processing that takes place parafoveally, i.e. while the eyes are still fixating the previous word. It is important to note that while the ERP literature does contain suggestions of an effect of predictability on components earlier than the N400, the reliability of these findings is far from clear; see Nieuwland (2019) for extensive discussion.

As noted above, at least two previous studies have elicited a predictability-based N400 effect in a natural reading paradigm, by time-locking the EEG to the start of the first eye fixation on a critical word, rather than by presenting words one at a time at central fixation (Dimigen et al., 2011; Kretzschmar et al., 2015). Both of these studies have reinforced questions about whether the predictability-based modulations of N400 amplitude and early eye movement measures reflect the same underlying processes. Dimigen et al. (2011), who had subjects read naturalistic texts in which the predictability of individual words varied across a wide range, established that in normal reading the N400 is very late relative to the timeline of eye movements, as had been suggested by comparison of previous results from separate eye movement studies and ERP studies. Dimigen et al. found that in the substantial majority of instances, the reader's eyes have already left a word by the time the N400 effect reaches its peak, and indeed, that in the majority of instances the reader's eyes have left the word by the time of the statistically-defined onset of the effect. On the basis of these results, they concluded:

> [T]he present data make it hard to conceive [of] the measurable neural effects of predictability as being causal in some way for the behavioral effects, because the bulk of the predictability effects in ERPs only followed those in behavior. This raises questions about the functional interpretation of the N400 peak, whose latency does not seem to correspond to the maximum processing difficulty as reflected in the [eye movement] record. (p. 565)

Kretzschmar et al. (2015) used co-registration to explore a different discrepancy between the eye movement and ERP effects. A word's context-independent frequency, like its predictability, has a robust effect on early eye movement measures (Rayner & Duffy, 1986; Staub et al., 2010), with many studies finding frequency and predictability to have additive effects on fixation duration (e.g. Rayner et al., 2004; Staub, 2020; see Staub, 2015, for detailed review of the evidence for additivity as opposed to interaction). However, previous ERP studies had found only inconsistent effects of word frequency on the N400; frequency effects tend to appear when a word is presented in isolation, but often do not appear when a word is presented with substantial preceding context (Dambacher et al., 2006; Van Petten & Kutas, 1990). Kretzschmar et al. factorially manipulated the frequency and predictability of a target word occurring relatively late in sentences, recording both eye movements and FRPs. They replicated, in a single experiment, the two distinct patterns present in the previous literature: Fixation duration measures demonstrated additive effects of predictability and frequency, while the N400 demonstrated only an effect of predictability, with no hint of an effect of frequency. Degno et al. (2019) have recently confirmed the lack of frequency effect on the N400 in another co-registration experiment.

Thus, both the timing of the N400 in natural reading and its insensitivity to word frequency, in sentence context, suggest that it may reflect processes downstream of the processes to which early fixation duration measures are sensitive. The amplitude of the N400 may reflect the difficulty of access to conceptual representations in semantic memory (Federmeier, 2022), or even the difficulty of post-lexical integration. In the present study, we test further predictions that arise from the view of the N400 as reflecting a relatively late process, by assessing other potential dissociations between eye movement and N400 effects. We explore the question of how two manipulations that target relatively early stages of letter or word processing affect N400 amplitude, or interact with predictability in modulating N400 amplitude, when implemented in sentence context. These are manipulations of the visual contrast of the text (Experiment 1) and parafoveal preview validity of the target word (Experiment 2).

In Experiment 1, we manipulate the predictability of a target word as well as visual contrast between the text and background. Several previous studies (e.g. Reingold & Rayner, 2006; White & Staub, 2012) have shown that like predictability, visual contrast affects early fixation duration measures, with faint text inducing longer mean first fixation duration and gaze duration (i.e. the sum of all fixation durations before leaving a word). Recently, Staub (2020, Experiment 1) conducted a high-powered experiment assessing the interaction between the effects of visual contrast and predictability, and did not find convincing evidence of such an interaction; while the main effects of both manipulations were substantial in size and statistically reliable, the effect of predictability was only a few milliseconds larger with faint text, and this interaction effect did not reach significance even in a very large experiment.

The present Experiment 1 provides an opportunity to replicate this additive pattern in eye movements, while at the same time addressing the previously unaddressed question of whether visual contrast influences the amplitude of the N400 in sentence reading. If the N400 reflects relatively late processing operations involving semantic access and/or integration of a word with its context, as suggested both by its timing and by the lack of effect of word frequency, visual contrast should not influence N400 amplitude. On this view, N400 amplitude reflects processing that takes place after the encoding stages that are affected by contrast are likely to be complete. Thus, we predict a dissociation between the eye movement and EEG results that is similar to the dissociation observed by Kretzschmar et al. (2015) with respect to predictability and frequency: Predictability and visual contrast should both influence fixation duration, and do so additively, but only predictability should influence N400 amplitude.

An early, classic ERP study (Holcomb, 1993) did address the influence of stimulus quality on the N400, though with a different task and somewhat different manipulations. Holcomb had subjects perform lexical decisions to targets that were preceded by either a semantic prime or by an unrelated word; semantic priming is known to reduce N400 amplitude. In addition, the target words were either intact or visually degraded (stimulus degradation was accomplished by different means in two experiments; in neither was this a contrast manipulation as in the present experiment). Holcomb replicated previous findings (e.g. Becker & Killion, 1977) that stimulus degradation increases lexical decision RT, and that it interacts with the priming manipulation, such that the effect of stimulus quality on RT is larger with unprimed targets. Priming also influenced N400 amplitude, as expected. The critical result was that stimulus quality did not influence N400 amplitude, nor was there a priming-by-stimulus quality interaction in this measure. Stimulus degradation was found to increase the latency of the N400, though this analysis relied on a peak latency analysis that may be susceptible to statistical artifacts (Luck, 2014); later evidence using mean amplitude measures supports the view that stimulus properties that are relevant during encoding do not impact the N400. Specifically, manipulations that target text characteristics such as font size or type and letter case have been found to differentially influence early EEG components in single-word processing studies, while having no influence on later N400 amplitude (e.g. Chauncey et al., 2008; Vergara-Martínez et al., 2015). Thus Experiment 1 of the present study, which tests a manipulation of stimulus quality while simultaneously recording eye movements, may be seen as potentially reinforcing a conclusion that is already suggested by Holcomb (1993), namely that N400 amplitude is not influenced by manipulations of stimulus quality even when behavioral measures are. Consistent with the hypothesis that we have outlined above, Holcomb speculated on the basis of the N400's insensitivity to stimulus quality either that "the mechanism underlying the N400 is a relatively late (post-word-level) integrative process," (p. 59) or that "N400 might index activation of meaning in a postlexical semantic memory system" (p. 60).

In Experiment 2, we test how parafoveal preview validity may modulate the predictability effect on the N400. In two experiments, Staub and Goddard (2019) manipulated the validity of the reader's parafoveal preview of a target word, using the boundary paradigm (Rayner, 1975), while also manipulating the target's predictability. In the invalid preview conditions, an unrelated word or a nonword letter string was presented in place of the target until the eyes crossed an invisible boundary between the preceding word and the target word. As expected based on a large previous literature (see Schotter et al., 2012, for review), invalid preview resulted in longer fixation durations on the target. The critical result, however, was that while normal predictability effects were observed in the valid preview conditions, invalid preview eliminated the effect of the predictability of the target on the duration of fixations on this word, at least in early measures (see also Luke, 2018). Staub and Goddard (2019) interpreted this pattern as suggesting that when the reader's first encounter with the target word is in foveal vision, as is the case in the invalid preview conditions, the bottom-up input is too clear for predictability to have much facilitatory effect on early stages of letter or word recognition.

As noted above, the predictability effect on the N400 is reliably obtained in ERP experiments with sequential presentation of each word at central fixation. Thus, this effect occurs even in the complete absence of parafoveal preview of the target; clearly, it is not eliminated when a word is first encountered in foveal vision. This previous evidence from separate eye movement and ERP experiments leads to the prediction that in a co-registration experiment employing the boundary paradigm, invalid parafoveal preview should eliminate the influence of predictability on target word fixation duration measures, as in Staub and Goddard (2019), while the predictability effect on the N400 should persist. This pattern would again support an interpretation of the N400 whereby it reflects a relatively late process that is unaffected by manipulations influencing early encoding stages.

Experiment 1

Methods

Participants

Thirty-three undergraduate students at the University of Massachusetts Amherst participated in exchange for monetary compensation. Participants provided written informed consent prior to the experiment, and the experimental protocol was approved by the IRB of the University of Massachusetts, Protocol Number 2018-4925. All participants reported being right-handed, native speakers of American English, with no history of neurological or language impairments. Two participants' data could not be used because of experimenter error, two were excluded based on extreme EEG voltages likely reflecting poor electrode contact, and three were excluded based on excessive electrode bridging. No subjects were excluded based on a pre-established criterion of having fewer than 10 trials in any condition after removing trials in which there were artifacts or for which the participant skipped the target word. Thus, the data from 26 participants (12 female; mean age:

20.32 years old; standard deviation: 3.79 years) were included in the analysis.

Materials

The 180 experimental sentences were taken from Staub and Goddard (2019, Experiment 2), and are provided in our data repository. There were 90 target words, with each embedded in two separate sentence contexts, one rendering the word highly predictable and the other rendering it unpredictable. The word immediately preceding the target word was the same in the two contexts. We crossed the predictability of the target word and the stimulus quality of the sentence resulting in a 2×2 design. The conditions are exemplified in Table 1. In the predictable contexts, target words had a mean cloze probability of 0.93 (sd = 0.03, max = 0.99, min = 0.87). In the unpredictable contexts, the target words had a mean cloze of 0.004 (sd = 0.01, max = 0.05, min = 0). The target words were an average of 4.2 characters long (sd = 0.74, max = 6, min = 3). All target words were of moderate to high frequency, with a mean Zipf frequency of 4.74 (sd = 0.61, max = 6.29, min = 3.36). The stimulus quality manipulation approximated the luminance values used by White and Staub (2012), by maintaining a constant grey background, but altering the contrast of the text in the foreground. The RGB values of the background in the experiment were (RGB = 245 245 245). The clear text was generated with the text colored black (RGB = 0 0 0); the faint text was generated with the text just slightly darker than the background (RGB = 235 235 235). The text was presented using 11 pt Courier New font.

Each participant saw all 180 sentences, with 45 trials in each condition. Each sentence was displayed on a single line. There were two experimental lists, which were counterbalanced between subjects. Participants saw each target word once in the clear text condition and once in the faint text condition; if a participant saw, e.g. a given target word in the predictable context and in the clear text condition (upper left cell of Table 1), they would also see the same target word in the unpredictable context and in the faint text condition (lower right cell in Table 1).

Table 1. Example stimuli in Experiment 1, with target word *rose* (not underlined in experiment). The luminance of the text and background in the table are for demonstration purposes only; they are not identical to the values in the experiment.

	Clear Text	Faint Text
Predictable Target	On Valentine's Day the woman received a single red rose from her secret admirer.	On Valentine's Day the woman received a single red rose from her secret admirer.
Unpredictable Target	The traffic cop finally admitted that the red rose that fell out of the car wasn't meant for him.	The traffic cop finally admitted that the red rose that fell out of the car wasn't meant for him.

Procedure

Eye movement data were collected using a desk-mounted Eyelink 1000 (SR Research, Ontario), with a sampling rate of 1000 Hz. Participants sat and rested their chin on a chin rest located 60 cm from the display computer. At that distance, 3 characters subtended approximately 1 degree of visual angle. Viewing was binocular, but only the right eye was monitored. EEG data were collected using a 128-channel HydroCel Geodesic Sensor Net (EGI, Eugene OR). Continuous EEG data were recorded at a sampling rate of 500 Hz with an online bandpass filter of 0.01–100 Hz. The experiment was implemented using the Experiment Builder software (SR Research, Ontario). At the beginning of each trial, an EEG trigger was sent by the software to the EEG data collection computer. This enabled us to timelock the eye tracking and EEG data streams offline.

Upon arrival participants were instructed about the procedure and then fitted with an EEG cap. Electrode Cz served as online reference, but the data were re-referenced offline to the average of the left and right mastoid electrodes. All electrode impedances were kept below 50 kOhm for the experiment, as is typical for high input impedance systems. The experiment was split into 2 blocks of 90 trials each, one where all the clear text items were presented and the other where all the faint text items were presented, with the order of blocks counterbalanced between participants. Trial order was randomised within blocks. The blocked design was chosen over a completely randomised design to avoid modulating pupil size between successive trials (Binda et al., 2014). Between the blocks participants were given a break. Participants read the sentences silently to themselves as their eyes were tracked, and answered yes/no comprehension questions following 27% of the trials (12 trials of the 45 total for each condition) in order to ensure attentive reading. The comprehension questions probed various parts of the sentence. The mean accuracy was 94% (standard deviation 4%), with all participants answering at least 83% correctly. Responses were indicated with a left or right trigger press on a gaming controller. The correspondence between YES and NO responses and the trigger (left or right) was counterbalanced between participants. There were 7 practice trials prior to the main experiment, with 4 of these practice sentences followed by comprehension questions. Including calibration, practice, and clean-up, the experiment took approximately 2 h.

Results

Eye movements

Eye tracking data from the 26 included subjects were preprocessed using Robodoc and Eyedry (http://blogs.umass.edu/eyelab/software/). These data were then analyzed in the R environment (R Core Team, 2019) using the tidyverse (Wickham et al., 2019) packages. Analyses focused primarily on the target word, with supplemental analyses performed for the word immediately preceding the target (see below). The target word region was defined as the characters making up the target itself as well as the leading and trailing whitespace characters. We excluded the 2.3% of trials in which the reader blinked while fixating the target on first pass reading.

The measures of interest were first fixation duration (the duration of the first fixation which fell upon the target region of the text), gaze duration (the sum of all of the durations of fixations upon the target region before exiting the region to the left or the right), go-past time (the sum of all of the durations of fixations from when the region is first fixated until it is exited to the right, including any regressive re-reading of previous regions, and any re-reading of the target region), and skipping proportion (the proportion of trials for which the reader did not fixate the target word on first pass reading, but rather made a saccade past it). If a reader skipped the target word on first pass reading, that trial was excluded from the analysis of the other three measures. The means in each condition, with by-subject standard error, are plotted in Figure 1.

Fixation times were left untransformed for our main analyses, as the question of whether the effects of the two manipulations are additive or interactive is best addressed with fixation times on their original scale; see Staub (2020) for discussion, and see Liceralde and Gordon (2018) and Schramm and Rouder (2019) for recent discussions of benefits and costs of the log transformation. We note, however, that supplementary analyses with log-transformed measures resulted in exactly the same patterns of significance for the target word in Experiment 1, and below we report any other cases in which patterns of significance differ; the full results of these models are presented in our online data repository.

First fixation duration, gaze duration, and go past time were analyzed using separate Linear Mixed Effects Models implemented using the lme4 package (Bates et al., 2014). The fixed effects were the predictability of the target word (predictable vs. unpredictable) and visual contrast (clear vs. faint). The contrasts for these effects were sum coded. The maximal models were fit first, but after failing to converge or returning singular fits, random slopes were dropped, with the slope accounting for the least variance getting dropped first, following Barr et al. (2013). This was iterated until the model converged; the resulting models are reported here. A summary of these models and all following models is provided in the Appendix. *P*-values were calculated using the lmerTest package (Kuznetsova et al.,

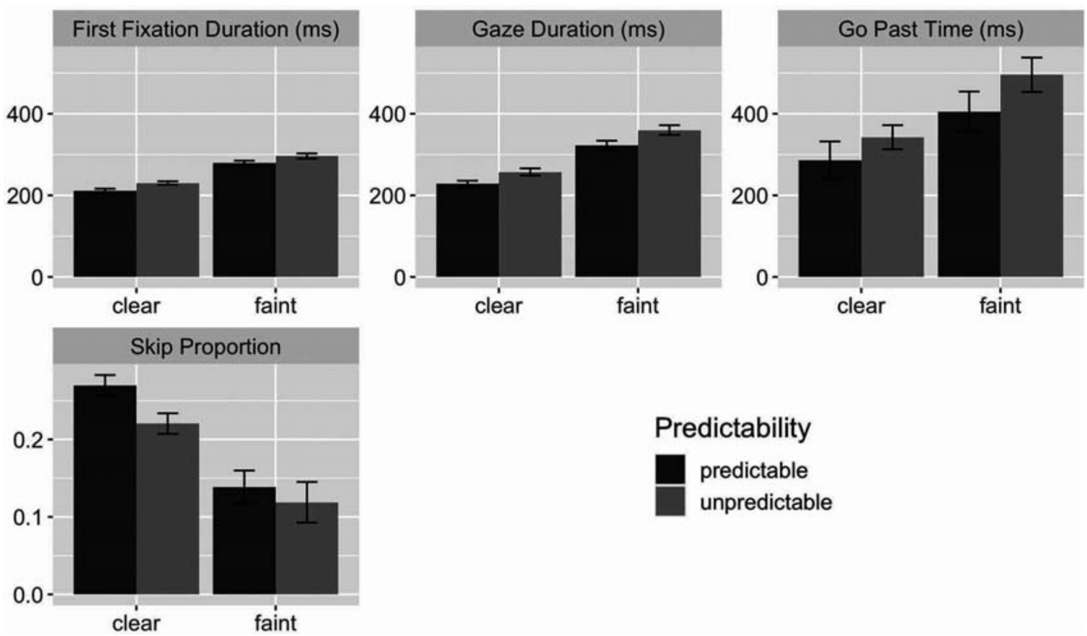

Figure 1. Condition means in Experiment 1, for each eye movement measure for the target word. Error bars represent by-subjects standard error.

Table 2. Mixed-effects model results for Experiment 1 target word eye movement data. Bolded *p*-values are statistically significant at $p < 0.05$; * indicates $p < 0.0125$, reflecting Bonferroni correction.

Measure	Effect	Estimate	SE	t- or z-value	*p*-value
First Fixation	Predictability (P)	17.350	3.160	5.491	**< 0.001** *
	Visual Contrast (VC)	66.905	6.579	10.169	**< 0.001** *
	Interaction (P x VC)	−0.485	5.646	−0.086	0.932
Gaze Duration	Predictability (P)	32.139	5.304	6.060	**< 0.001** *
	Visual Contrast (VC)	97.854	10.902	8.976	**< 0.001** *
	Interaction (P x VC)	10.656	8.762	1.216	0.224
Go Past	Predictability (P)	73.97	16.16	4.475	**< 0.001** *
	Visual Contrast (VC)	138.04	20.92	6.597	**< 0.001** *
	Interaction (P x VC)	42.36	24.04	1.762	0.078
Skip	Predictability (P)	−0.256	0.083	−3.102	**0.002** *
	Visual Contrast (VC)	−1.004	0.201	−4.990	**< 0.001** *
	Interaction (P x VC)	0.043	0.167	0.258	0.797

2017). To correct for multiple comparisons associated with multiple eye tracking measures, a Bonferroni correction was applied following Von der Malsburg and Angele (2017) to give a new significance threshold of $\alpha = 0.0125$ (0.05 / 4).

For all three reading time measures, there were statistically significant effects of both predictability and visual contrast (*p*s < 0.001) and these were in the expected direction, with predictable words receiving shorter fixations compared to unpredictable words, and clear text receiving shorter fixations compared to faint text. These two effects did not demonstrate a significant interaction in any measure.

The probability of skipping the target word was analyzed by means of a logistic mixed-effects model with similar coding. Visual contrast had a significant effect on the skipping rate ($p < 0.001$), with clear target words being skipped more frequently compared to faint target words, and predictable words were more frequently skipped as well ($p = 0.002$). Again, there was no significant interaction ($p = 0.797$) (Table 2).

Fixation-related potentials (FRPs)

Preprocessing of the EEG data was performed in MATLAB (Mathworks, Inc.) using EEGLAB (Delorme & Makeig, 2004). From the eye movement data we extracted the time that the participant initially fixated the target word, in ms from the start of the trial. We used this information to create an event list to be imported alongside the raw EEG data to add triggers corresponding to those initial fixations. Because the eye movement data were collected with a sampling rate of 1000 Hz, and EEG data were collected with a sampling rate of 500 Hz, an eye movement trigger could fall between EEG samples, in which case the time of the eye movement trigger was adjusted down by 1 ms. Trials where the participant skipped the target word were automatically excluded, and we again excluded trials in which the reader blinked while fixating the target (2.3% of trials). A .3–30 Hz bandpass filter was applied to the data. The PREP Pipeline (Bigdely-Shamlo et al., 2015) was used to identify and interpolate bad channels. The mean number of channels that were interpolated was 9 out of the total 129 (sd = 5, max = 22, min = 1). The data were then re-referenced to the mastoid electrodes. An automatic threshold-based artifact rejection algorithm removed epochs with extreme voltages (± 100µV). Because we aimed to ensure maximum comparability in analytical choices to the study of Kretzschmar et al. (2015), we did not apply further artifact correction methods, but see below for analyses assessing the role of eye movement artifact. The EEG data were then exported for analysis in R. Following all trial exclusion steps, the average number of trials per condition per subject which went into creating the FRPs after preprocessing was 30.63 (sd = 6.49, max = 43, min = 11).

Our analyses included 19 centro-parietal electrodes: from C3 through Cz to C4 anteriorly to P3 through Pz to P4 posteriorly, and the remaining 13 electrodes contained in the resulting rectangle.[1] Figure 2 shows the electrodes included in this ROI. These electrodes were chosen *a priori* by visually inspecting the topographic plots provided in Kretzschmar et al. (2015) and Dimigen et al. (2011), both of which utilised FRPs to examine predictability effects. We did not include electrode location in our analyses, as there was no justification to do so given the lack of spatial modulation of the N400 FRP observed in this previous work, and given that including such a factor may increase Type 1 error rates (Luck & Gaspelin, 2017).

Our primary concern was with the amplitude of the N400 FRP. Accordingly, based on the FRP morphology reported in these previous studies, we defined the N400 window as 250 ms to 500 ms post-fixation, and the baseline window as −200 ms to 0 ms, where 0 ms is the start of the fixation. We computed the average

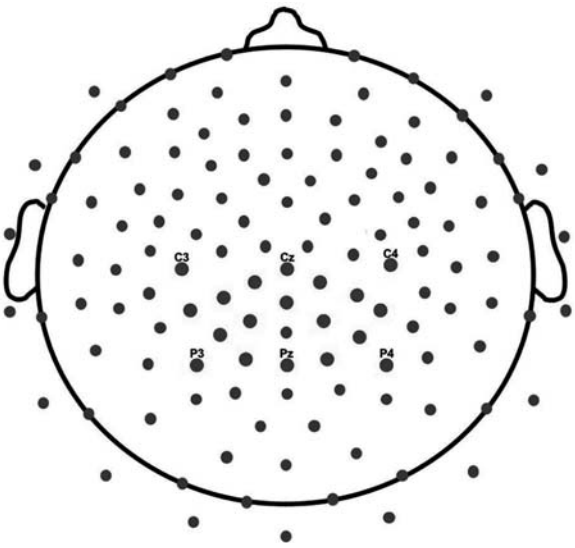

Figure 2. Layout of the 128-channel EEG cap. Highlighted electrodes were included in our predefined ROI.

voltage within these windows for each trial. The FRP associated with fixating the target word, collapsing across all 19 electrodes in our ROI, is presented in the top panel of Figure 3, while FRP for electrode CPz alone, which is centrally located in our ROI, is presented in the bottom panel of Figure 3. In addition, Figure 4 illustrates the topographical distribution of the effects of interest, plotting FRPs for electrodes both within the ROI and in more frontal, occipital, and temporal locations. From this plot it is clear that the N400 is largely restricted to centro-parietal electrodes, as expected. Moreover, while very prominent eye movement artifacts are present at frontal electrode sites, such artifacts do not appear in our ROI.

Our analysis utilised mixed effects modeling of individual trial N400 amplitude (Alday, 2019). The model included the trial specific mean baseline voltage, the predictability of the target word, visual contrast, and the interaction between predictability and contrast as predictors of the trial specific N400 amplitude. The effects of predictability and visual contrast were coded using sum contrast coding. The maximal model was fit first, but after failing to converge, random slopes were successively dropped in ascending order of the variance they accounted for until the model converged (Barr et al., 2013).

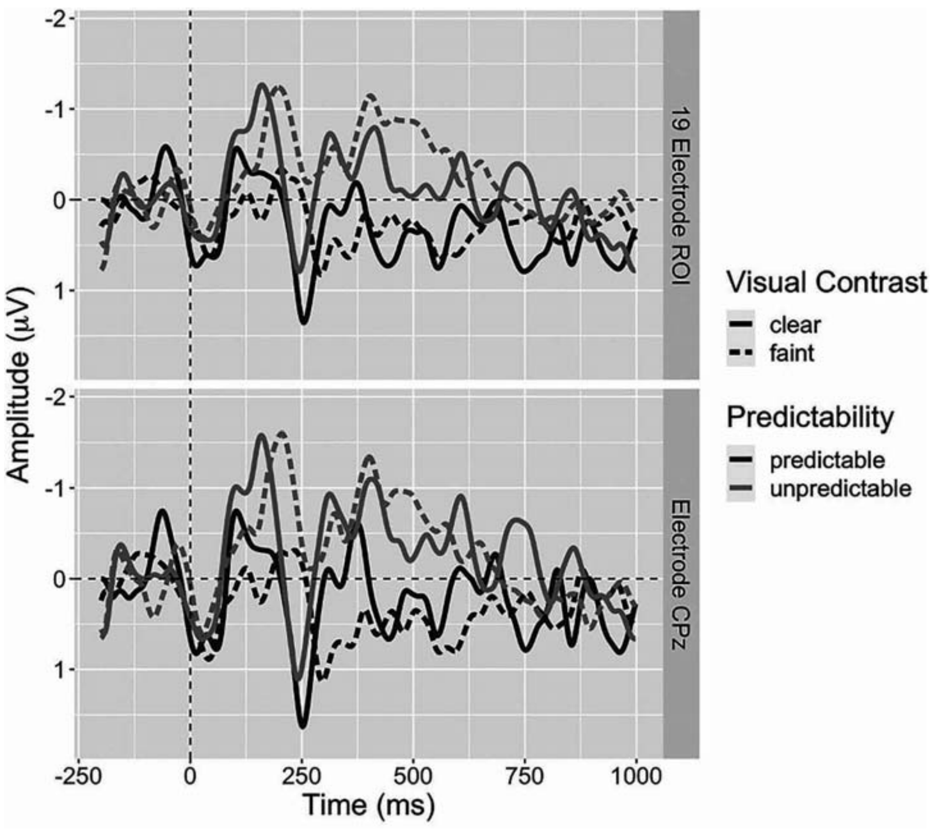

Figure 3. Grand average FRPs for all electrodes in our ROI (top) and electrode CPz (bottom) for target word fixations in Experiment 1. Voltages have been baseline corrected for visual presentation. A 10 Hz low pass filter was applied for visual presentation. Time 0 is the beginning of the first fixation on the target word.

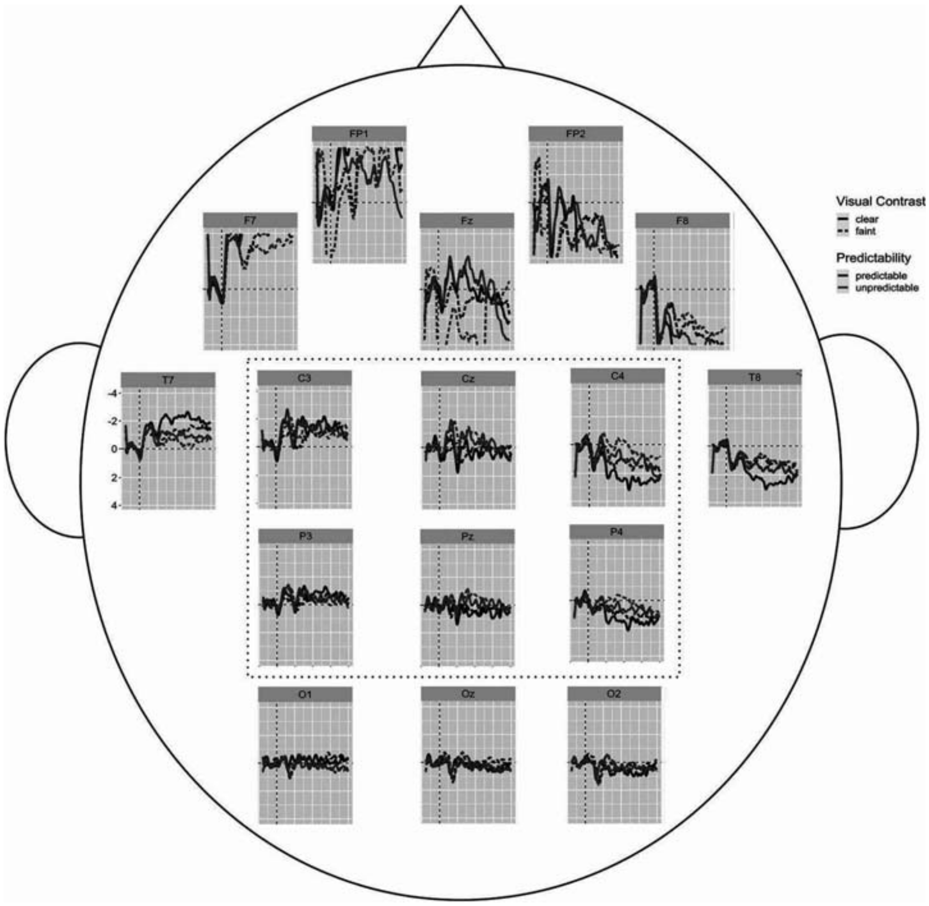

Figure 4. Grand average FRPs from −200 ms to 1000 ms for electrodes around the scalp for target word fixations in Experiment 1. Voltages have been baseline corrected for visual presentation. A 10 Hz low pass filter was applied for visual presentation. Time 0 is the beginning of the first fixation on the target word. The ROI for analysis is outlined with a dotted line in the middle of the plot.

A summary of the statistical results is presented in Table 3. We observed a significant effect of baseline amplitude on N400 amplitude. Replicating prior co-registration research, we observed a significant effect of predictability on the N400 amplitude. Contrary to all of the eye movement measures, visual contrast was not a significant predictor of N400 amplitude. As this null effect is of theoretical interest in the present context, we performed a supplemental Bayes Factor analysis to assess the strength of the evidence for the null, using the brms package (Bürkner, 2017). The model utilised the full random effects structure, was run for 10,000 iterations with 2,000 iterations of burn-in, and utilised the following priors: the main effect of predictability had a Normal prior distribution with a mean of −2 and a standard deviation of 1; all of the other effects (the main effect of contrast, the main effect of the baseline voltage and the interaction between predictability and contrast) used Normal priors with means of 0 and standard deviations of 1. This analysis found that there was moderate evidence for a null effect of visual contrast on the N400 (BF = 3.22; estimate: −0.18; 95% CI: −0.58–0.21). Finally, as was the case with the eye tracking measures, predictability and visual contrast did not interact.

Table 3. Results of mixed-effects model of target word FRPs in Experiment 1. Bolded *p*-values are statistically significant at $p < 0.05$.

Effect	Estimate	SE	t-value	*p*-value
Baseline	0.319	0.015	21.897	**< 0.001**
Predictability (P)	−0.719	0.198	−3.641	**< 0.001**
Visual Contrast (VC)	−0.192	0.169	−1.137	0.255
Interaction (P x VC)	−0.345	0.336	−1.028	0.304

We additionally fit a linear mixed effects model that used the centered first fixation duration on a trial as a predictor for that trial's elicited N400 FRP amplitude; this was the same model as the one summarised in Table 3, but with the additional predictor of the first fixation duration. This model revealed no significant effect of the first fixation duration on the amplitude of the N400 ($p = 0.71$). Yet another model was fit using the centered gaze duration on the target as a predictor for the amplitude of the N400. This too revealed no significant effect of gaze duration on the N400 ($p = 0.14$).

These results indicate that the FRP effect we find in the N400 time window is unlikely to be artifact of the fixation duration differences between conditions, as in the present experiment fixation duration did not predict N400 amplitude (cf. Dimigen & Ehinger, 2021).

Pre-target effects

To investigate potential effects of the experimental manipulations emerging prior to direct fixation on the target, we conducted an additional set of eye movement and FRP analyses. For the eye movement analysis, we investigated the same four measures as for the target word, but for the pre-target word. This word tended to be short, and therefore was frequently skipped; recall that this word was identical for the high- and low-predictability sentences with a given target word. For the FRP analysis, we time-locked the EEG to the start of the fixation immediately prior to the first fixation on the target word, whether this fixation fell on the pre-target word or an earlier word (as in Kretzschmar et al., 2015). We did so for two main reasons. First, there is evidence that effects of parafoveal processing are driven by properties of the saccade goal only (McDonald, 2006; Kliegl et al., 2013), so time-locking to the fixation immediately prior to the first fixation on the target word should better capture such processing. Second, the pre-target word was skipped almost half of the time, so time-locking the pre-target FRP to the first fixation on the pre-target word would have resulted in very substantial data loss. Only trials that did contain a first-pass target fixation were included; like for the target word FRP analysis, trials on which the target was skipped were not included.

The eye movement data are presented in Figure 5. The results of the statistical analyses[2], using the same procedures as for the target word, are summarised in Table 4. The analyses of the pre-target word revealed significant effects of visual contrast on first fixation duration, gaze duration, and go past time ($ps < 0.007$), as well as a significant effect of target word predictability on gaze duration ($p = 0.007$). We discuss this latter finding below.

The grand average FRPs for the pre-target fixation are presented in Figure 6, and the model estimates are summarised in Table 5. The statistical model did not reveal an influence of the target's predictability on the pre-target N400. The only significant predictor of the pre-target N400 amplitude was the pre-target baseline period amplitude. Figure 6 does suggest a predictability-based effect in a later time window, with predictable targets resulting in less negative FRPs. However, this is likely to be a reflection of the target word N400 that we have already identified, offset by the approximately 200–250 ms of the pre-target fixation.

Discussion

In this co-registration experiment, predictability of the target word influenced both fixation duration and the amplitude of the N400 FRP, time-locked to the onset of the first fixation on the target. On the other hand,

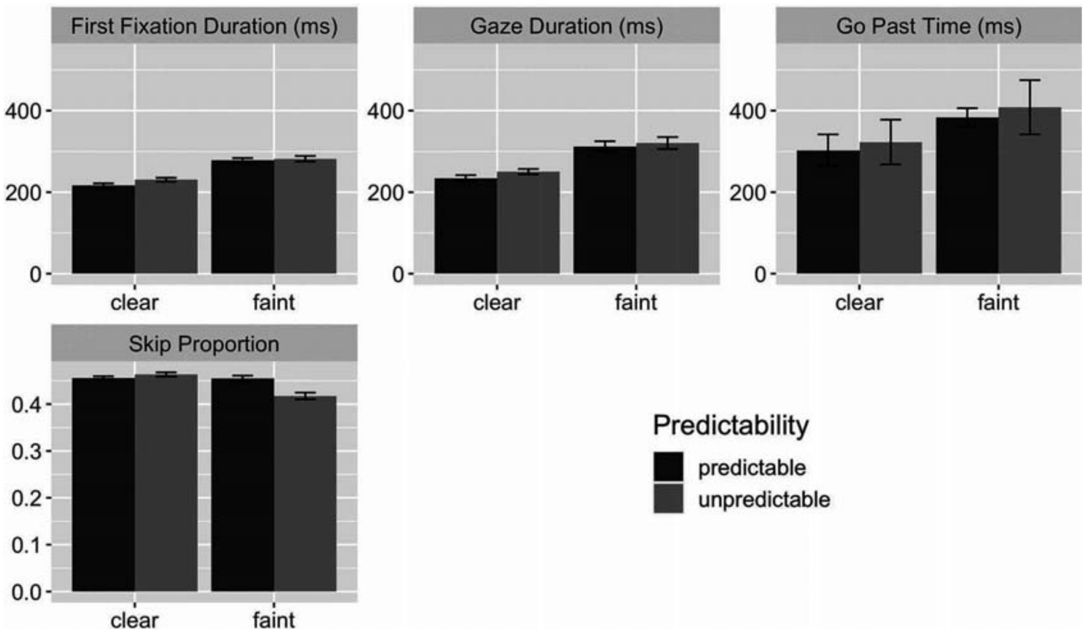

Figure 5. Condition means in Experiment 1, for each eye movement measure for the pre-target word. Error bars represent by-subjects standard error.

Table 4. Mixed-effects model results for Experiment 1 pre-target word eye movement data. Bolded *p*-values are statistically significant at $p < 0.05$; * indicates $p < 0.0125$, reflecting Bonferroni correction.

Measure	Effect	Estimate	SE	t-value	*p*-value
First Fixation	Predictability (P)	9.411	4.033	2.334	**0.022**
	Visual Contrast (VC)	56.836	5.358	10.608	**< 0.001** *
	Interaction (P x VC)	−11.180	6.552	−1.706	0.088
Gaze Duration	Predictability (P)	13.493	5.038	2.678	**0.007 ***
	Visual Contrast (VC)	75.445	5.053	14.931	**< 0.001** *
	Interaction (P x VC)	−10.438	10.182	−1.025	0.305
Go Past	Predictability (P)	27.135	14.332	1.893	0.058
	Visual Contrast (VC)	80.893	27.572	2.934	**0.007 ***
	Interaction (P x VC)	1.576	28.832	0.055	0.956
Skip	Predictability (P)	−0.073	0.068	−1.083	0.279
	Visual Contrast (VC)	−0.120	0.130	−0.924	0.355
	Interaction (P x VC)	−0.170	0.134	−1.273	0.203

while visual contrast affected fixation duration on the target word, it had no detectable effect on the amplitude of the N400. The eye movement effects are expected based on previous results (e.g. Staub, 2020). The FRP effects confirm predictions laid out in the Introduction, which were based both on specific previous results (e.g. Holcomb, 1993) and on theoretical considerations regarding the interpretation of the N400 in light of its timing. As in Holcomb (1993), we see a dissociation between the effect of stimulus quality on a behavioral measure of lexical processing and its effect on the N400. Like lexical decision latency, eye fixation duration is lengthened when text is faint, but in both cases the N400 amplitude is not affected. It is especially notable that the effect of faint text on first fixation and gaze duration, in the present experiment, was actually more than five times as large as the effect of predictability, in absolute terms (e.g. model estimates of 75 ms vs. 13 ms, for gaze duration), but even in this context the contrast manipulation did not have a measurable effect on the N400.

It is worth noting a few discrepancies between the present eyetracking results and the results of some previous studies. First, while the present results regarding the effects of visual contrast on eye movements are almost entirely consistent with Staub (2020), one exception is that the effect of visual contrast on skipping was significant in the present experiment, but did not quite reach significance in Staub (2020). However, in other studies such as Warrington et al. (2018), this effect has reached significance. Thus, this result is not particularly surprising. A more surprising result was the effect of target word predictability on gaze duration on the pre-target word. Such "parafoveal-on-foveal" effects of predictability have usually not been present in the previous literature (Brothers et al., 2017). One possibility is that because these pre-target words were very short, some of the fixations on these words reflected attempted skips that undershot their intended landing site on the target word (e.g. Drieghe et al., 2008). Another possibility

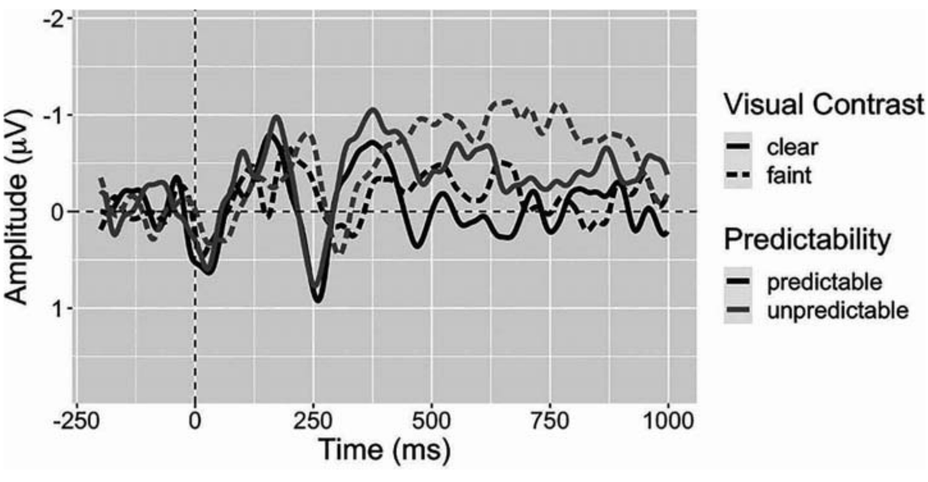

Figure 6. Grand average FRPs for all electrodes in our ROI for pre-target fixations in Experiment 1. Voltages have been baseline corrected for visual presentation. A 10 Hz low pass filter was applied for visual presentation. Time 0 is the beginning of the pre-target fixation.

is that these pre-target words were themselves somewhat predictable in the predictable contexts; the predictability of these words was never normed. But most likely, this result is simply a Type I error. In the valid preview conditions of Staub and Goddard (2019, Experiment 2), which used the same materials, there was no hint of such an effect; first fixation durations and gaze durations were actually numerically longer on the pre-target word when it was followed by a predictable target. Moreover, this effect was not significant in our supplementary analysis of log gaze duration (see footnote 2).

In Experiment 2, we attempt to confirm a predicted dissociation between eye movements and FRPs in the opposite direction from the dissociation that emerged in Experiment 1. In Experiment 1, we saw that visual contrast affects fixation durations, but not N400 amplitude. In Experiment 2, we test whether the predictability of the target word influences N400 amplitude in the absence of parafoveal preview of this word, though under these circumstances it does not influence fixation durations. On the assumption that the predictability effect on the N400 indexes an obligatory, but relatively late, stage of processing, this effect should emerge regardless of how early encoding is influenced by denying parafoveal preview.

Experiment 2

Methods

Participants

Forty undergraduate students at the University of Massachusetts, Amherst participated in exchange for course credit or monetary compensation; none participated in Experiment 1. All participants were right-handed, native speakers of American English and had no reported history of neurological or language impairments. The same exclusion criteria were used as in Experiment 1, leading to 4 participants being dropped from the analysis due to excessive bridging and 3 due to out of bounds voltage. The data from 33 participants (28 female; mean age: 19.52 years old; standard deviation: 1.54 years) were used for the full analysis.

Table 5. Results of mixed-effects model of pre-target FRPs in Experiment 1. Bolded *p*-values are statistically significant at $p < 0.05$.

Effect	Estimate	SE	t-value	*p*-value
Baseline	0.391	0.014	26.693	**< 0.001**
Predictability (P)	−0.305	0.170	−1.797	0.072
Visual Contrast (VC)	0.073	0.171	0.427	0.669
Interaction (P x VC)	0.099	0.340	0.292	0.770

Materials

The same sentences were used in Experiment 2 as in Experiment 1. However, rather than crossing the predictability of a target word and the visual contrast of the stimulus, in Experiment 2, predictability was crossed with preview validity. These materials are exemplified in Table 6. The invalid previews were the word previews from Staub and Goddard (2019, Experiment 2), which were contextually implausible English words which were unrelated in meaning to the targets they masked. Previews and targets were matched in length and the position of ascending and descending letters.

After the experiment concluded, participants' awareness of display changes was assessed by asking if they noticed anything strange about the display. Of the 33 included participants, 13 failed to notice any anomaly. The 20 who did report noticing some anomaly were then asked to report on what percentage of trials they noticed this anomaly. The mean percentage was 24.15% (sd = 17.07); only 5 participants reported noticing an anomaly on 40% of trials or more. In fact, 50% of trials featured a display change.

Comprehension questions again followed 27% of the sentences. The mean accuracy was 94% (standard deviation: 0.04). All participants answered at least 83% of questions correctly.

Procedure

The same procedure was followed as in Experiment 1, with the exception that conditions were not blocked; trials from the valid preview and invalid preview conditions were randomly intermixed.

Results

Eye movements

We analyzed the same four eye movement measures for the target word that were analyzed in Experiment 1. We

Table 6. Example stimuli for Experiment 2 with target word *rose* (not underlined in actual experiment). The word to the left of the "|"was the preview.

	Valid Preview	Invalid Preview
Predictable Target	On Valentine's Day the woman received a single red {rose|rose} from her secret admirer.	On Valentine's Day the woman received a single red {cane|rose} from her secret admirer.
Unpredictable Target	The traffic cop finally admitted that the red {rose|rose} that fell out of the car wasn't meant for him.	The traffic cop finally admitted that the red {cane|rose} that fell out of the car wasn't meant for him.

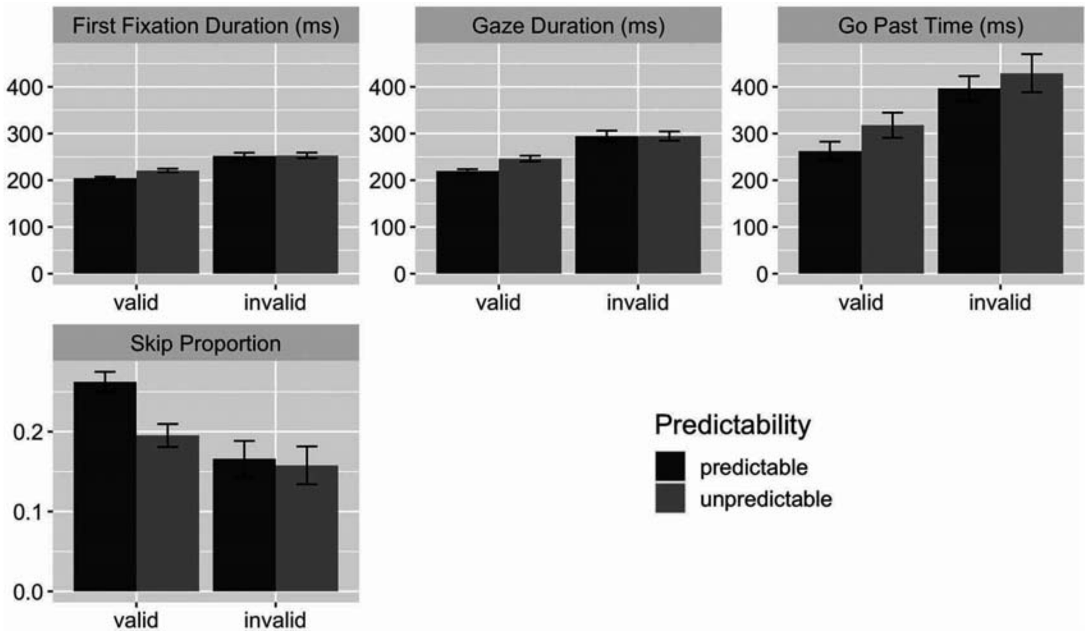

Figure 7. Condition means in Experiment 2, for each eye movement measure for the target word. Error bars represent by-subjects standard error.

again removed trials in which the participant blinked during first pass reading of the target; 3.6% of trials were removed on this basis. In addition, 5.9% of trials were removed because the display change was not executed within 10 ms of the start of fixation on the target word.

Condition means are presented in Figure 7, and the statistical analyses are summarised in Table 7. Predictability and preview validity both had significant effects on first fixation duration and gaze duration. In addition, these factors interacted in both measures, with a sizable predictability effect appearing in the valid preview condition, but essentially absent in the invalid preview condition. The main effects remained significant in go-past time, though their interaction was no longer significant ($p = 0.17$). The analysis of skipping additionally revealed two significant main effects; after correction for multiple comparisons, the interaction did not reach significance ($p = 0.03$).[3]

Fixation-related potentials (FRPs)

The same preprocessing procedures as in Experiment 1 were followed. The average number of channels that were interpolated during preprocessing was 10 out of 129 (sd = 4, max = 22, min = 3). The average number of trials per condition per participant after exclusions was 29.15 (sd = 5.36; max = 40, min = 13). The resulting FRP of the entire 19-electrode ROI is presented in the top panel of Figure 8 and the FRP associated with electrode CPz alone is presented in the bottom panel. A topographic plot is presented in Figure 9. The results from the linear mixed effects model of the data are presented in Table 8.

Table 7. Mixed-effects model results for Experiment 2 target word eye movement data. Bolded *p*-values are statistically significant at $p < 0.05$; * indicates $p < 0.0125$, reflecting Bonferroni correction.

Measure	Effect	Estimate	SE	t-value	*p*-value
First Fixation	Predictability (P)	9.174	2.342	3.917	**< 0.001 ***
	Preview Validity (PV)	−39.622	5.689	−6.965	**< 0.001 ***
	Interaction (P x PV)	15.578	4.686	3.324	**< 0.001 ***
Gaze Duration	Predictability (P)	14.715	3.555	4.140	**< 0.001 ***
	Preview Validity (PV)	−61.603	9.244	−6.664	**< 0.001 ***
	Interaction (P x PV)	28.611	7.113	4.023	**< 0.001 ***
Go Past	Predictability (P)	46.96	11.41	4.114	**< 0.001 ***
	Preview Validity (PV)	−122.87	17.81	−6.899	**< 0.001 ***
	Interaction (P x PV)	24.93	18.24	1.367	0.172
Skip	Predictability (P)	−0.260	0.084	−3.105	**0.002 ***
	Preview Validity (PV)	0.561	0.121	4.651	**< 0.001 ***
	Interaction (P x PV)	−0.318	0.149	−2.140	**0.032**

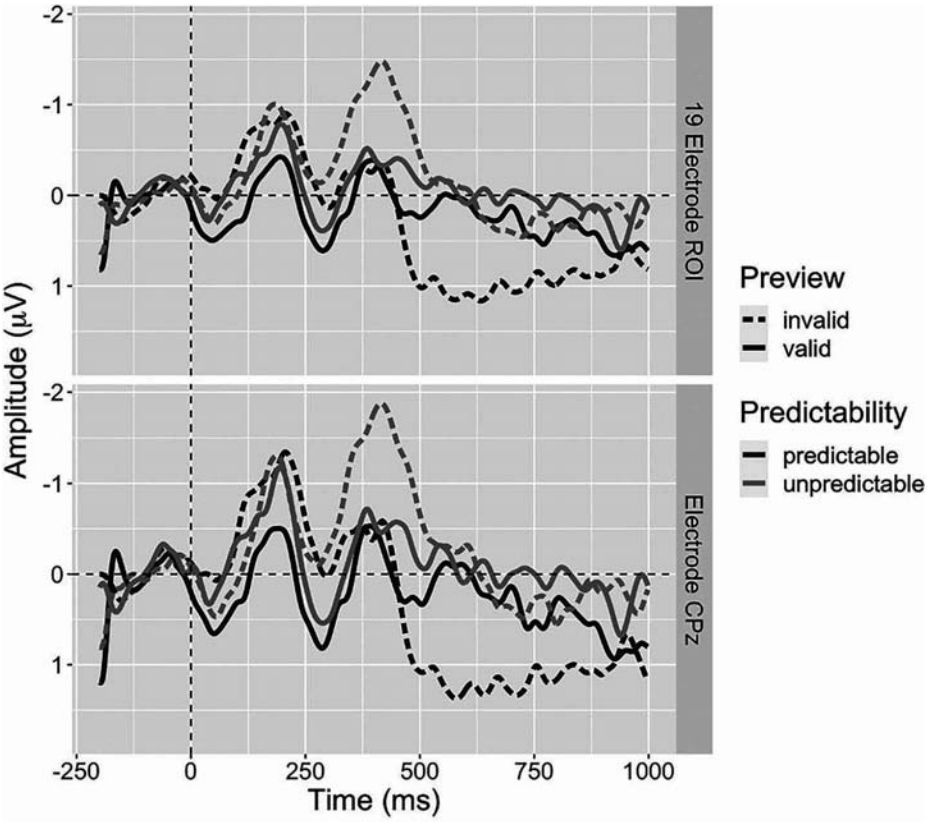

Figure 8. Grand average FRPs for all electrodes in our ROI (top) and electrode CPz (bottom) for target word fixations in Experiment 2. Voltages have been baseline corrected for visual presentation. A 10 Hz low pass filter was applied for visual presentation. Time 0 is the beginning of the first fixation on the target word.

There were significant main effects of predictability, with unpredictable targets eliciting more negative N400 amplitudes, and of preview validity, with invalid previews eliciting more negative N400 amplitudes. In addition, there was a significant interaction between these two factors ($p = 0.04$). Notably, this interaction was in the opposite direction of the early eye tracking measures, where invalid preview eliminated the predictability effect; as indicated in Figure 8, predictability strongly affected the N400 in the invalid preview conditions, while having little effect in the valid preview conditions.

As can be clearly seen in Figure 8, the difference between the predictable and unpredictable invalid preview conditions persisted well beyond the traditional N400 window. From roughly 450 ms to 800 ms post-fixation, the invalid/predictable condition appears to be more positive than the invalid/unpredictable condition. As we did not plan analyses of this later time window, we do not present significance tests. The fact that the difference between these conditions persists largely unchanged suggests that it could be regarded as reflecting a single effect extending from the N400 window out to 800 ms post-fixation, rather than separate N400 and late positivity components. Future research should directly examine the possibility that invalid preview of a highly predictable word results in a late positivity.

It is notable that the baseline amplitude was not a significant predictor of the N400 amplitude in this experiment ($p = 0.726$). However, Alday (2019) demonstrated that for late components such as the N400, a correlation between the baseline value and the component's amplitude should not be assumed. Additionally, as in Experiment 1, we fit models that used the centered first fixation duration or gaze duration on a trial as a predictor for the N400 FRP amplitude. As in Experiment 1, these models revealed no significant effect of the first fixation duration on the amplitude of the N400 ($p = 0.99$), and no significant effect of the gaze duration on the amplitude of the N400 ($p = 0.82$). Again, statistical modelling shows that the FRP effects we find in the N400 time window are unlikely to be due to eye movement differences between conditions.

Pre-target effects

As for Experiment 1, we explore potential effects of the experimental manipulations prior to direct fixation on the

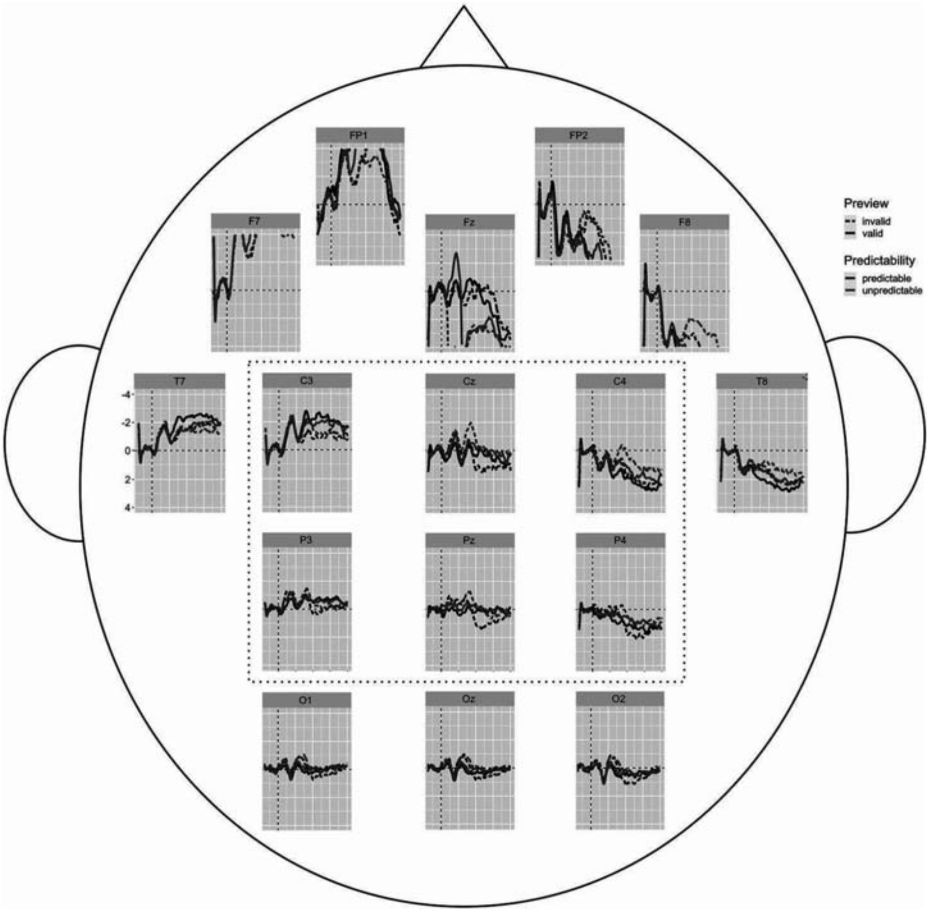

Figure 9. Grand average FRPs from −200 ms to 1000 ms for electrodes around the scalp for target word fixations in Experiment 2. Voltages have been baseline corrected for visual presentation. A 10 Hz low pass filter was applied for visual presentation. Time 0 is the beginning of the first fixation on the target word. The ROI for analysis is outlined with a dotted line in the middle of the plot.

target, by examining eye movement measures for the pre-target word, and FRPs time-locked to the pre-target fixation. Again, we note that because of the differences in the pre-target eye movement and FRP analyses (with the former focusing on the pre-target word, and the latter related to the pre-target fixation, whether this fixation fell on the preceding word or an even earlier one), direct comparison of the results should proceed with caution.

The eye tracking measures are presented in Figure 10, and the statistical results are presented in Table 9.[4] The preview manipulation had a significant effect on each of the reading time measures on the pre-target word, with longer reading times when this word was followed by an invalid preview. Target word predictability did not significantly influence any of the eye tracking measures, nor was there a significant interaction between the two manipulations. We again observe very high rates of skipping of this pre-target word, and the skipping rate was not modulated by the experimental manipulations.

The FRP associated with the pre-target fixation is plotted in Figure 11. A summary of the linear mixed effects model is presented in Table 10. There was a significant effect of preview validity on the N400 amplitude, with invalid previews yielding more negative amplitudes. While we did not find a significant effect of target word predictability on the pre-target word N400, we did find a significant interaction between the two factors. The form of this interaction is clearly visible in Figure 11: Predictability affected the N400 amplitude time-locked to the pre-target fixation in the valid preview conditions, but not in the invalid preview conditions, where the word to the right of fixation was never actually the predictable target. Notably, the interaction associated with the N400 time-

Table 8. Results of mixed-effects model of target word FRPs in Experiment 2. Bolded *p*-values are statistically significant at $p < 0.05$.

Effect	Estimate	SE	t-value	*p*-value
Baseline	−0.005	0.016	−0.350	0.726
Predictability (P)	−0.440	0.103	−4.256	**< 0.001**
Preview Validity (PV)	−0.374	0.135	−2.769	**0.009**
Interaction (P x PV)	−0.423	0.207	−2.045	**0.041**

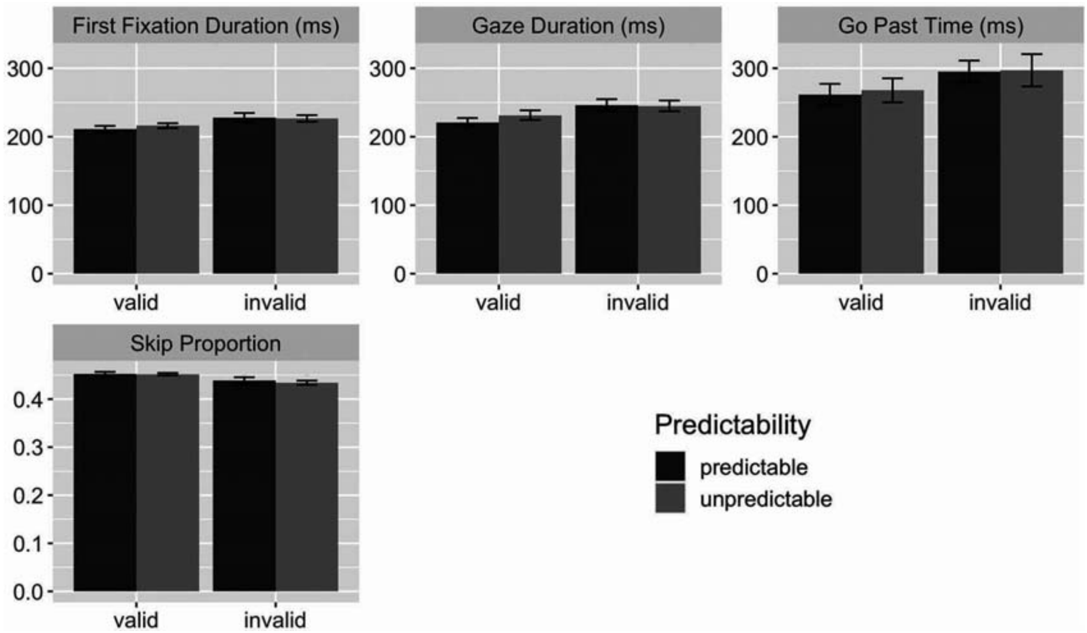

Figure 10. Condition means in Experiment 2, for each eye movement measure for the pre-target word. Error bars represent by-subjects standard error.

locked to the pre-target fixation is the opposite pattern of the interaction associated with the N400 elicited by the target word fixation, where the predictability effect was more pronounced in the invalid preview conditions.

Discussion

We begin by summarising the eye movement data. The results for the target word very closely replicate the findings of Staub and Goddard (2019). Most notably, for the early reading time measures of first fixation duration and gaze duration, the predictability effect is largely confined to the valid preview conditions, resulting in a statistical interaction between predictability and preview validity. Also like Staub and Goddard, we find additive effects of the two manipulations on go-past time; it appears that predictability influences this late measure, which reflects regressive eye movements as well as first pass fixations, even in the absence of valid preview. The effects of the two manipulations on skipping are also extremely similar to the effects observed by Staub and Goddard, who found, like in the present study, a numerical pattern whereby predictability affected skipping only with valid preview; but also like in the present study, this interaction did not quite reach significance.

On the pre-target word, the only effect to reach significance was an effect of preview validity, with longer reading times when the word that was visible to the right of fixation was the invalid preview for the target, which was a word that did not fit semantically into the sentence. This effect is arguably surprising, given the conclusions of Brothers et al. (2017) regarding the non-existence of lexical parafoveal-on-foveal effects. Staub

Table 9. Mixed-effects model results for Experiment 2 pre-target word eye movement data. Bolded *p*-values are statistically significant at $p < 0.05$; * indicates $p < 0.0125$, reflecting Bonferroni correction.

Measure	Effect	Estimate	SE	t-value	*p*-value
First Fixation	Predictability (P)	1.838	2.959	0.621	0.536
	Preview Validity (PV)	13.547	2.596	5.218	**< 0.001** *
	Interaction (P x PV)	−7.353	5.188	−1.417	0.157
Gaze Duration	Predictability (P)	4.051	4.523	0.896	0.376
	Preview Validity (PV)	17.803	6.095	2.921	**0.006** *
	Interaction (P x PV)	−15.025	7.997	−1.879	0.064
Go Past	Predictability (P)	4.508	6.897	0.654	0.513
	Preview Validity (PV)	30.111	6.898	4.365	**< 0.001** *
Interaction (P x PV)	−8.602	13.809	−0.623	0.533	
Skip	Predictability (P)	−0.019	0.072	−0.270	0.786
	Preview Validity (PV)	−0.083	0.065	−1.284	0.199
	Interaction (P x PV)	−0.073	0.123	−0.590	0.555

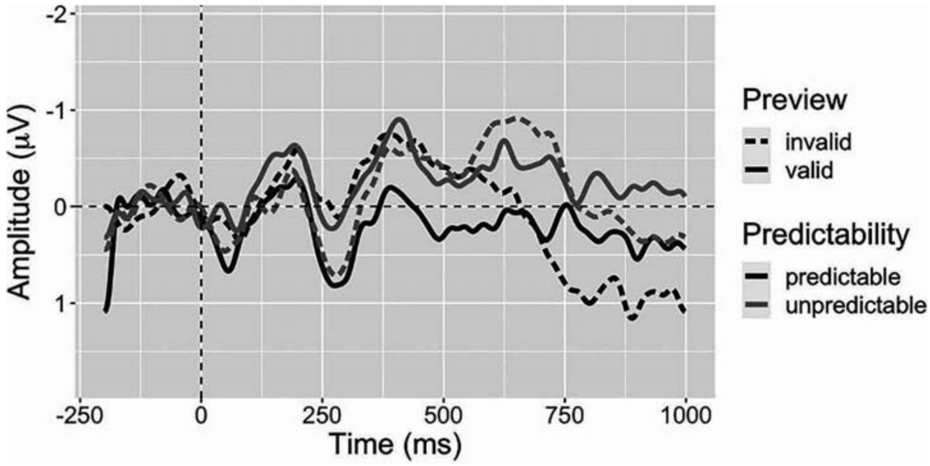

Figure 11. Grand average FRPs for all electrodes in our ROI for pre-target fixations in Experiment 2. Voltages have been baseline corrected for visual presentation. A 10 Hz low pass filter was applied for visual presentation. Time 0 is the beginning of the pre-target fixation.

and Goddard (2019) did obtain a somewhat similar effect with these materials, though in their experiment it was driven almost entirely by inflated reading times in an additional set of conditions in which the invalid preview was a random letter string. At present, we are unsure how seriously to take the present effect, just as we are unsure how seriously to take the apparent predictability-based parafoveal-on-foveal effect in Experiment 1. Providing more indication of the ephemeral nature of these effects, there was little hint of a predictability-based parafoveal-on-foveal effect in the valid preview conditions of the present experiment.

Turning to the FRP results, we found a pattern that partially confirmed our predictions. We expected the predictability effect on the N400 for the target word to be unaffected by preview validity; we did not expect invalid preview to eliminate the predictability effect on the N400, as it does for eye fixation duration measures. But there was an interaction, due to an especially *large* effect with invalid preview. In contrast to Experiment 1, the effect of predictability on the N400 with valid preview was very small; indeed, it is barely visible in Figure 8. On the other hand, the effect with invalid preview was substantial, and indeed it appears larger in magnitude than in either the clear or faint text conditions of Experiment 1.

Table 10. Results of mixed-effects model of pre-target FRPs in Experiment 2. Bolded *p*-values are statistically significant at $p < 0.05$.

Effect	Estimate	SE	t-value	*p*-value
Baseline	−0.010	0.015	−0.678	0.498
Predictability (P)	0.094	0.101	0.930	0.352
Preview Validity (PV)	−0.341	0.101	−3.391	**< 0.001**
Interaction (P x PV)	0.542	0.201	2.692	**0.007**

Turning attention to the N400 associated with the pre-target fixation makes more sense of the pattern in the valid preview conditions. Time-locking to the pre-target fixation, the valid preview conditions showed a substantial effect of predictability on the N400 (with valid preview of a predictable target resulting in a more positive N400), while the invalid preview conditions showed no effect. Thus, both the valid and invalid preview conditions demonstrate the N400's sensitivity to predictability, but this is earlier with valid preview, showing up primarily in the pre-target N400. The sensitivity of the N400 to the predictability of a word that is visible in the parafovea is consistent with results from the flanker-RSVP paradigm (e.g. Stites et al., 2017; Barber et al., 2011). Given the high constraint contexts in our materials, the present result is also consistent with the FRP findings of a parafoveal N400 effect in more normal reading circumstances, under conditions of near-perfect predictability (Kretzschmar et al., 2009). Still, there remains uncertainty about why the timing of the N400 effect for the valid preview conditions in Experiment 2 is different from the timing in Experiment 1, where both clear and faint text were displayed with valid preview. In Experiment 1, the predictability of the target had a clear effect on the N400 for the target word, but no effect for the pre-target fixation; in the valid preview conditions of Experiment 2, the predictability of the target had only a very modest effect on the N400 for the target word (Figure 8) but a more pronounced effect for the pre-target fixation (Figure 11).

There is, however, a clear interpretation of the large target word N400 effect in the invalid preview conditions of Experiment 2. The N400 in these conditions appears to show a morphology that is similar to the classic N400 obtained with sequential presentation of words at target fixation, with a relatively discrete onset and rapid rise, compared to the less temporally distinct negativity that is visible in the other conditions in the present experiments. In all other conditions, the actual onset of processing of the target word may be assumed to vary from trial to trial; on each trial, the target word has already received some degree of parafoveal processing prior to the start of fixation on the target – time 0 for computing the FRP – but this amount varies, resulting in an FRP that is in effect an average of EEG waveforms with different starting points. In the invalid preview conditions of Experiment 2, on the other hand, time 0 for computing the FRP genuinely marks the beginning of perceptual processing of the target word on every trial, resulting in a well-defined N400 that mimics the N400 observed with word-by-word presentation.

Finally, the target FRPs also exhibited an unexpected finding, a late positivity in response to predictable targets following invalid preview. As noted above, we did not investigate this effect in more detail for reasons of statistical conservativeness. However, we note that the time course of the positivity effect seems to lend further support to the observation that positivity effects depend on foveally perceived input. More specifically, using a flanker-RSVP design Payne et al. (2019) found a late positivity for incongruent target words that was confined to foveal presentation, whereas the N400 effect to incongruent targets already emerged parafoveally.

In sum, Experiment 2 demonstrated an even more striking dissociation between the predictability effect on eye movements and on FRPs. With invalid preview, the predictability effect on early fixation measures was eliminated, but the predictability effect on the target word N400 not only survived – it was particularly pronounced.

General discussion

In two co-registration experiments, we replicated existing results from the eye movement literature, while uncovering new findings regarding the effects of visual contrast and parafoveal preview validity on the N400 component of FRPs. With regard to eye movements, we found in Experiment 1 that target word predictability and visual contrast have additive effects on fixation duration measures, and we found in Experiment 2 that predictability and preview validity have interactive effects, with an invalid parafoveal preview essentially eliminating the predictability effect on early fixation duration measures; these results replicate Staub (2020) and Staub and Goddard (2019), respectively.

The novel contribution of the present study comes from the simultaneously collected EEG data. First, we found that visual contrast had no detectable effect on the amplitude of the N400, nor did the effect of visual contrast interact with the effect of predictability on N400 amplitude. While we cannot, of course, entirely rule out the possibility that there is some effect of visual contrast on the N400, we found a Bayes Factor of >3 in favor of the null. On the other hand, we found that denying parafoveal preview of a target word did interact with predictability in influencing the N400 component of the FRP to the target word, but in the opposite direction from the effect on eye fixation durations: The effect of predictability on the N400 was made larger, not smaller, when the reader lacked valid preview of this word. The opposite effect was observed for the FRP for the pre-target fixation: A predictability effect on the N400 was in evidence only in the valid preview condition, i.e. when the target was visible to the right of fixation.

The eye movement findings reinforce the conclusion that eye fixation durations in reading are sensitive to both the difficulty of visual encoding of the stimulus and to a range of linguistic factors such as a word's predictability. The additivity of the effects of contrast and predictability (as well as frequency; Staub, 2020) suggests that visual contrast influences an early encoding stage that is largely separate from the lexical processing stages that are influenced by predictability. On the other hand, the interactive effects of parafoveal preview validity and predictability suggests that the latter effect may depend on initial perceptual evidence being at least somewhat equivocal; it appears that when a word is first encountered in foveal vision, as is the case when preview is invalid, the perceptual evidence is too clear for predictability to have much effect.

We acknowledge that there is a tension between these interpretations; if the predictability effect appears only when initial perceptual evidence is equivocal, as when a word is in low-acuity parafoveal vision, then it seems reasonable to expect that the predictability effect should be made larger by making perceptual evidence even more equivocal, i.e. by reducing visual contrast. However, this does not appear to be the case. We do not have a straightforward resolution of this issue, at present, and regard it as an important topic for future research. One possibility is that there is effectively a maximum size to the predictability effect, such that it is simply not made larger by faint text; but this is just a redescription of the data, rather than a theoretical explanation. Another possibility is that the lack of predictability effect on eye movements with invalid preview is caused by an actual change in the predictability of the target when an invalid preview is presented; early encoding of the invalid preview may lead readers to abandon or reduce their confidence in their expectations (Parker et al., 2017).

The lack of visual contrast effect on the amplitude of the N400 is consistent with several previous studies failing to find an effect of factors such as stimulus degradation (Holcomb, 1993) and font characteristics (e.g. Chauncey et al., 2008; Vergara-Martínez et al., 2015) on N400 amplitude, and extends this null result of visual manipulations to a more natural reading paradigm. Thus, while presenting text with low contrast or in a difficult font (e.g. Rayner et al., 2006; Staub, 2020) increases the time the eyes remain on a word before moving on, these manipulations do not seem to influence the amplitude of the N400. This dissociation is explained if the duration of the eyes' fixation on a word reflects how long it takes for the processing system to complete some stage of lexical processing, and this latency reflects the latencies of both perceptual and linguistic substages (e.g. Reichle et al., 2003), while the amplitude of the N400 reflects the operation of a conceptual/semantic stage of processing that is substantially downstream of any perceptual encoding stages (e.g. Federmeier, 2022; Holcomb, 1993).

The FRP results from Experiment 2 were in one sense inconsistent with our initial expectations; we had predicted that while the effect of predictability on early eye movement measures would be eliminated with invalid preview, the effect of predictability on the target FRP would be unchanged. But in fact there was a predictability x preview interaction in FRPs, though in the opposite direction from the eye movement interaction. In hindsight, however, this pattern is quite sensible. As we discussed above, the invalid preview conditions are functionally like the sequential word presentation conditions in which the classic N400 is observed, where the starting point of lexical processing is under precise experimental control. It is unsurprising that we would observe a relatively large N400 on invalid preview trials, due to the precise time-alignment of target word processing on these trials. What is most notable, from a theoretical perspective, is that Experiment 2 even more clearly demonstrates a dissociation between eye movement and EEG measures: With invalid preview, predictability has little if any effect on first fixation duration or gaze duration on the target word, while having a pronounced effect on the amplitude of the N400.

We account for this pattern by supposing that a word's predictability influences at least two distinct processing stages, and that the overt influence of this variable on eye movements and EEG actually reflect influences on the first and second of these stages, respectively. First, predictability influences a relatively early stage of lexical access; in the context of a model such as the DRC (Coltheart et al., 2001), this may be conceptualised as an influence on a word's activation level in the orthographic lexicon, with the orthographic entry for a predictable word receiving some pre-activation, such that less activation from the input is required in order for the entry to reach a threshold level. It is by means of this influence that predictability typically influences first fixation duration, and it is this influence that is neutralised when parafoveal preview is denied, i.e. when the first encounter with the word is in foveal vision, where the perceptual input is very clear. When the input is very clear, the pre-activation of a predictable entry in the orthographic lexicon may make little difference to the timeline of word recognition. Second, predictability influences what Federmeier (2022) has called "semantic access", "a linking between the current input and long-term stores of experience and knowledge" (p. 4). It is this process that the N400 reflects, and N400 amplitude is reduced when a word is predictable because "some or all of the information that would normally be evoked by the stimulus has already recently been activated" (p. 4). We note that this description also evokes Holcomb's (1993) observation, quoted above, that the "N400 might index activation of meaning in a postlexical semantic memory system" (p. 60). In order for language comprehension to proceed successfully, such a process of access to semantic memory is entirely obligatory, and the work that is required in order to complete this process does not depend on how easily a word's orthographic lexical entry is initially accessed.

To reinforce the interpretation that fixation durations and N400 amplitudes are influenced by distinct processes, we also highlight the fact that in both Experiments 1 and 2, trial-by-trial variation in first fixation duration and gaze duration on the target were not significant predictors of the elicited N400 amplitude. We note that in a similar analysis, Dimigen et al. (2011) also found that first fixation duration on a trial was not a significant predictor of N400 amplitude. Dimigen et al. did find gaze duration to be a significant predictor of N400 amplitude, and the explanation of this discrepancy from the present results will require further investigation.

In sum, we suggest that the fact that predictability influences both eye movements and the N400 has been somewhat misleading, insofar as this common influence has led to the supposition that eye fixation durations and N400 amplitudes reflect common processing operations. We propose that these dependent measures reflect distinct processing operations, and the effect of predictability on both may be due to an early influence of predictability on orthographic access, which shows up in an effect on eye movements, and a later influence of predictability on semantic access, in the sense of Federmeier (2022), which shows up in N400 amplitude.

Notes

1. The full set of electrodes was: E31, E36, E37, E42, E52, E53, E54, E55, E61, E62, E78, E79, E80, E86, E87, E92, E93, E104 and E129.
2. Results from the models using log-transformed reading times for the pre-target word were different in two respects: In the log first fixation model, the interaction effect was significant; in the log gaze duration model the effect of target word predictability was not significant.
3. The interaction between preview validity and predictability was significant with log go-past time as the dependent measure. All other patterns remained the same with log transformed reading time measures.
4. Models using log reading times resulted in the same statistical patterns.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was funded by a grant from the NSF (BCS 1732008).

Data availability statement

Eye movement and EEG data for both experiments, as well as analysis scripts, are available at: https://osf.io/6nr9t/?view_only=61eb1851bba4454f822563e0f5c7bb92.

References

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

Alday, P. M. (2019). How much baseline correction do we need in ERP research? Extended GLM model can replace baseline correction while lifting its limits. *Psychophysiology*, *56*(12), e13451. https://doi.org/10.1111/psyp.13451.

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

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. (2014). Fitting linear mixed-effects models using lme4. arXiv:1406.5823.

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(3), 389–401. https://doi.org/10.1037/0096-1523.3.3.389

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K. M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*, 16. https://doi.org/10.3389/fninf.2015.00016

Binda, P., Pereverzeva, M., & Murray, S. O. (2014). Pupil size reflects the focus of feature-based attention. *Journal of Neurophysiology*, *112*(12), 3046–3052. https://doi.org/10.1152/jn.00502.2014

Brothers, T., Hoversten, L. J., & Traxler, M. J. (2017). Looking back on reading ahead: No evidence for lexical parafoveal-on-foveal effects. *Journal of Memory and Language*, *96*, 9–22. https://doi.org/10.1016/j.jml.2017.04.001

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Chauncey, K., Holcomb, P. J., & Grainger, J. (2008). Effects of stimulus font and size on masked repetition priming: An event-related potentials (ERP) investigation. *Language and Cognitive Processes*, *23*(1), 183–200. https://doi.org/10.1080/01690960701579839

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Dambacher, M., Dimigen, O., Braun, M., Wille, K., Jacobs, A. M., & Kliegl, R. (2012). Stimulus onset asynchrony and the timeline of word recognition: Event-related potentials during sentence reading. *Neuropsychologia*, *50*(8), 1852–1870. https://doi.org/10.1016/j.neuropsychologia.2012.04.011

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., & 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., 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

Drieghe, D., Rayner, K., & Pollatsek, A. (2008). Mislocated fixations can account for parafoveal-on-foveal effects in eye movements during reading. *Quarterly Journal of Experimental Psychology*, *61*(8), 1239–1249. https://doi.org/10.1080/17470210701467953

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*(4), 777–813. https://doi.org/10.1037/0033-295X.112.4.777

Federmeier, K. D. (2022). Connecting and considering: Electrophysiology provides insights into comprehension. *Psychophysiology*, *59*(1), e13940. https://doi.org/10.1111/psyp.13940.

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, *30*(1), 47–61. https://doi.org/10.1111/j.1469-8986.1993.tb03204.x

Kliegl, R., Hohenstein, S., Yan, M., & McDonald, S. A. (2013). How preview space/time translates into preview cost/benefit for fixation durations during reading. *The Quarterly Journal of Experimental Psychology*, *66*(3), 581–600. https://doi.org/10.1080/17470218.2012.658073

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 co-registration 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*(1), 1–26. https://doi.org/10.18637/jss.v082.i13

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics:(de) constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920–933. https://doi.org/10.1038/nrn2532

Liceralde, V. T., & Gordon, P. C. (2018, October 10). Consequences of power transforms as a statistical solution in linear mixed-effects models of chronometric data. https://doi.org/10.31234/osf.io/f73mh

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT Press.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, *54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Luke, S. G. (2018). Influences on and consequences of parafoveal preview in reading. *Attention, Perception, & Psychophysics*, *80*(7), 1675–1682. https://doi.org/10.3758/s13414-018-1581-0

McDonald, S. A. (2006). Parafoveal preview benefit in reading is only obtained from the saccade goal. *Vision Research*, *46*(26), 4416–4424. https://doi.org/10.1016/j.visres.2006.08.027

Nieuwland, M. S. (2019). Do 'early'brain responses reveal word form prediction during language comprehension? A critical review. *Neuroscience & Biobehavioral Reviews*, *96*, 367–400. https://doi.org/10.1016/j.neubiorev.2018.11.019

Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D. I., Ferguson, H. J., Fu, X., Heyselaar, E., Huettig, F., Matthew Husband, E., Ito, A., Kazanina, N., Kogan, V., Kohút, Z., Kulakova, E., Mézière, D., Politzer-Ahles, S., Rousselet, G., … Von Grebmer Zu Wolfsthurn, S. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1791), 20180522. https://doi.org/10.1098/rstb.2018.0522

Parker, A. J., Kirkby, J. A., & Slattery, T. J. (2017). Predictability effects during reading in the absence of parafoveal preview. *Journal of Cognitive Psychology*, *29*(8), 902–911. https://doi.org/10.1080/20445911.2017.1340303

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), e13432. https://doi.org/10.1111/psyp.13432.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *16*(1), 65–81. https://doi.org/10.1016/0010-0285(75)90005-5

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z reader model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 720–732. https://doi.org/10.1037/0096-1523.30.4.720

Rayner, K., & Clifton Jr, C. (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. (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., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, *21*(3), 448–465. https://doi.org/10.1037/0882-7974.21.3.448

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The EZ reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, *26*(4), 445–476. https://doi.org/10.1017/S0140525X03000104

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the EZ reader model. *Psychological Science*, *17*(9), 742–746. https://doi.org/10.1111/j.1467-9280.2006.01775.x

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

Schramm, P., & Rouder, J. (2019, March 5). Are Reaction Time Transformations Really Beneficial?. https://doi.org/10.31234/osf.io/9ksa6

Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128*(3), 302–319. https://doi.org/10.1016/j.cognition.2013.02.013

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

Staub, A. (2020). Do effects of visual contrast and font difficulty on readers' eye movements interact with effects of word frequency or predictability? *Journal of Experimental Psychology: Human Perception and Performance*, *46*(11), 1235–1251. https://doi.org/10.1037/xhp0000853

Staub, A., & Goddard, K. (2019). The role of preview validity in predictability and frequency effects on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(1), 110–127. https://doi.org/10.1037/xlm0000561

Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1280–1293. https://doi.org/10.1037/a0016896

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

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, *18*(4), 380–393. https://doi.org/10.3758/BF03197127

Vergara-Martínez, M., Gómez, P., Jiménez, M., & Perea, M. (2015). Lexical enhancement during prime–target integration: ERP evidence from matched-case identity priming. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(2), 492–504. https://doi.org/10.3758/s13415-014-0330-7

Von der Malsburg, T., & Angele, B. (2017). False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of Memory and Language*, *94*, 119–133. https://doi.org/10.1016/j.jml.2016.10.003

Warrington, K. L., McGowan, V. A., Paterson, K. B., & White, S. J. (2018). Effects of aging, word frequency, and text stimulus quality on reading across the adult lifespan: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(11), 1714–1729. https://doi.org/10.1037/xlm0000543.

White, S. J., & Staub, A. (2012). The distribution of fixation durations during reading: Effects of stimulus quality. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(*3*), 603–617. http://doi.org/10.1037/a0025338

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., … & Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. doi:10.21105/joss.01686