Early parafoveal semantic integration in natural reading

Yali Pan[1]*, Steven Frisson[1], Kara D. Federmeier[2], Ole Jensen[1]

[1]Centre for Human Brain Health, School of Psychology, University of Birmingham, Birmingham, UK
[2]Department of Psychology, Program in Neuroscience, and the Beckman Institute for Advanced Science and Technology, University of Illinois, Champaign, Illinois, USA

*Corresponding author. Email: Y.Pan.1@bham.ac.uk

Abstract

Humans can read and comprehend text rapidly, implying that readers might process multiple words per fixation. However, the extent to which parafoveal words are previewed and integrated into the evolving sentence context remains disputed. We investigated parafoveal processing during natural reading by recording brain activity and eye movements using MEG and an eye tracker while participants silently read one-line sentences. The sentences contained an unpredictable target word that was either congruent or incongruent with the sentence context. To measure parafoveal processing, we flickered the target words at 60 Hz and measured the resulting brain responses (i.e., *Rapid Invisible Frequency Tagging, RIFT*) during fixations on the pre-target words. Our results revealed a significantly weaker tagging response for target words that were incongruent with the sentence context compared to congruent ones, even within 100 ms of fixating the word immediately preceding the target. This reduction in the RIFT response was also found to be predictive of individual reading speed. We conclude that semantic information is not only extracted from the parafovea but can also be integrated with the sentence context before the word is fixated. This early and extensive parafoveal processing supports the rapid word processing required for natural reading. Our study suggests that theoretical frameworks of natural reading should incorporate the concept of deep parafoveal processing.

Introduction

Reading is a remarkable human skill that requires rapid processing of written words. We typically fixate each word for only 225-250 ms, but nevertheless manage to encode its visual information, extract its meaning, and integrate it into the larger context, while also doing saccade planning (Rayner, 2009). To overcome the tight temporal constraints during reading, we preview the next word in the parafovea before moving our eyes to it (Jensen et al., 2021;

Reichle and Reingold, 2013; Schotter, 2018). Substantial evidence suggests that parafoveal information can be extracted at various linguistic levels, including orthography (Drieghe et al., 2005; Inhoff, 1989; Johnson et al., 2007; White, 2008; Williams et al., 2006), phonology (Ashby et al., 2006; Ashby and Rayner, 2004; Chace et al., 2005; Miellet and Sparrow, 2004; Pollatsek et al., 1992; Rayner et al., 1995), lexicality (Kennedy and Pynte, 2005; Kliegl et al., 2006), syntax (Snell et al., 2017; Wen et al., 2019) and semantics (Rayner and Schotter, 2014; Schotter, 2013; Schotter et al., 2015; Schotter and Jia, 2016); for a comprehensive review see (Schotter et al., 2012). However, for semantics in particular, controversy remains about the extent and type of information extracted from parafoveal processing under various conditions. Moreover, it is unknown when and how the previewed semantic information can be used – i.e., integrated into the evolving sentence context – which is an integral component of the ongoing reading process.

For some time, it was claimed that parafoveal preview was limited to perceptual features of words and did not extend to semantics (Inhoff, 1982; Inhoff and Rayner, 1980; Rayner et al., 2014, 1986). However, eye tracking-based evidence for the extraction of parafoveal semantic information began to emerge from studies that used languages other than English, including Chinese (Tsai et al., 2012; Yan et al., 2012, 2009; Zhou et al., 2013) and German (Hohenstein et al., 2010; Hohenstein and Klieg, 2014), and was eventually extended into English (Rayner and Schotter, 2014; Schotter et al., 2015; Schotter and Jia, 2016; Veldre and Andrews, 2018, 2017, 2016a, 2016b). For example, Schotter and Jia (2016) showed preview benefits on early gaze measures for plausible compared to implausible words, even for plausible words that were unrelated to the target. These results demonstrate that semantic information can indeed be extracted from parafoveal words. However, due to the limitations of the boundary paradigm, which only assesses effects after target words have been fixated, it is challenging to precisely determine when and how parafoveal semantic processing takes place. Furthermore, it is generally hard to distinguish between the effects of cross-saccade integration (e.g., mismatch between the preview and the word fixated) and the effects of how differing words fit into the context itself (Veldre and Andrews, 2016a, 2016b).

Complementary evidence showing that semantic information can be extracted parafoveally, even in English, comes from electrophysiological studies. Context-based facilitation of semantic processing can be observed as reductions in the amplitude of the N400 component (Kutas and Hillyard, 1984, 1980), a negative-going event-related potential (ERP) response observed between about 300 and 500 ms after stimulus onset, which has been linked to

semantic access (DeLong et al., 2014; Federmeier, 2022; Federmeier et al., 2007; Kutas and Federmeier, 2011; Lau et al., 2008). Basic effects of contextual congruency on the N400 – smaller responses to words that do versus do not fit a sentence context (e.g., to "butter" compared to "socks" after "He spread the warm bread with …") – are also observed for parafoveally-presented words (Antúnez et al., 2022; Barber et al., 2013, 2010; López-Peréz et al., 2016; Meade et al., 2021), and, even when all words are congruent, N400 responses to words in parafoveal preview, like those to foveated words, are graded by increasing context-based predictability (Payne et al., 2019; Payne and Federmeier, 2017; Stites et al., 2017). Although many of these effects have been measured in the context of unnatural reading paradigms (e.g., the "RSVP flanker paradigm"), similar effects obtain during natural reading. Using the stimuli and procedures from Schotter and Jia (Schotter and Jia, 2016), Antúnez et al. showed that N400 responses, measured relative to the fixation before the target words (i.e., before the boundary change while the manipulated words were in parafoveal preview), were sensitive to the contextual plausibility of these previewed words (Antúnez et al., 2022). These studies suggest that semantic information is available from words before they are fixated, even if that information does not always have an impact on eye fixation patterns.

Thus, both eye tracking and electrophysiological studies have provided evidence suggesting that semantic information is extracted from words in parafoveal preview. However, most of these studies have been limited to measuring parafoveal preview from fixations to an immediately adjacent word, raising questions about exactly how far in advance semantic information might become available from parafoveal preview. Moreover, important questions remain about the extent to which parafoveally extracted semantic information can be functionally integrated into the building sentence-level representation. Although some ERP studies have found that the semantic information extracted from parafoveal preview is carried forward, affecting semantic processing when that same word is later fixated (Barber et al., 2010; Payne et al., 2019; Stites et al., 2017), other studies have not observed any downstream impact (Barber et al., 2013; Li et al., 2015). Furthermore, post-N400 ERP components, linked to more attentionally-demanding processes associated with message-building and revision, do not seem to be elicited during parafoveal preview (Li et al., 2023; Milligan et al., 2023; Payne et al., 2019; Schotter et al., 2023). Therefore, critical questions remain about the time course and mechanisms by which semantic information is extracted and used during reading.

Answering those questions requires an approach that allows a more continuous and specific assessment of sensitivity to target word semantics during parafoveal processing across

multiple fixations, and, in particular, that can speak to how attention is allocated across words during natural reading. We tackle these core issues using a new technique that combines the use of frequency tagging and the measurement of magnetoencephalography (MEG)-based signals.

Frequency tagging, also known as steady-state visually evoked potentials, involves flickering a visual stimulus at a specific frequency and then measuring the neuronal response associated with processing the stimulus (Norcia et al., 2015; Vialatte et al., 2010). It has been widely used to investigate visuospatial attention (Gulbinaite et al., 2019; Kritzman et al., 2022; Müller et al., 2003, 1998; Norcia et al., 2015; Vialatte et al., 2010) and has recently been applied to language processing (Beyersmann et al., 2021; Montani et al., 2019; Wu et al., 2023). However, the traditional frequency tagging technique flickers visual stimuli at a low frequency band, usually below 30 Hz, such that the flickering can be visible and may interfere with the ongoing task. To address this limitation, we developed the rapid invisible frequency tagging (RIFT) technique, which involves flickering visual stimuli at a frequency above 60 Hz, making it invisible and non-disruptive to the ongoing task. Responses to RIFT have been shown to increase with the allocation of attention to the stimulus bearing the visual flicker (Brickwedde et al., 2022; Drijvers et al., 2021; Duecker et al., 2021; Ferrante et al., 2023; Gutteling et al., 2022; Zhigalov et al., 2021, 2019; Zhigalov and Jensen, 2022, 2020). In our previous study, we adapted RIFT to a natural reading task and found temporally-precise evidence for parafoveal processing at the lexical level (Pan et al., 2021). The RIFT technique provides a notable advantage by generating a signal — the tagging response signal — specifically yoked to just the tagged word. This ensures a clear separation in processing the tagged word from the ongoing processing of other words, addressing a challenge faced by eye tracking and ERP/FRP approaches. Moreover, RIFT enables us to monitor the entire dynamics of attentional engagement with the tagged word, which may begin a few words before the tagged word is fixated.

In the current study, RIFT was utilized in a natural reading task to investigate parafoveal semantic integration. We recruited participants (n = 34) to silently read one-line sentences while their eye movements and brain activity were recorded simultaneously by an eye-tracker and MEG. The target word in each sentence was always unpredictable (see Behavioural pre-tests in Methods) but was semantically congruent or incongruent with the preceding sentence context (for the characteristics of words, see Table 1). The target words were tagged by flickering an underlying patch, whose luminance kept changing in a 60 Hz sinusoid

throughout the sentence presentation. The patch was perceived as grey, the same colour as the background, making it invisible. To ensure that the flicker remained invisible across saccades, we applied a Gaussian transparent mask to smooth out sharp luminance changes around the edges (Figure 1A). Parafoveal processing of the target word was indexed by the RIFT responses recorded using MEG during fixations of pre-target words.

This paradigm allows us to address three questions. First, we aimed to measure when in the course of reading people begin to direct attention to parafoveal words. Second, we sought to ascertain when semantic information obtained through parafoveal preview is integrated into the sentence context in a manner that affects reading behaviours. Modulations of pre-target RIFT responses by the contextual congruity of target words would serve as evidence that parafoveal semantic information has not only been extracted and integrated into the sentence context but that it is affecting how readers allocate attention across the text. Third, we explored whether these parafoveal semantic attention effects have any relationship to reading speed.

Table 1. Characteristics of pre-target, target, and post-target words

	Pre-target	Target	Post-target
Word frequency	124.5 (310.9)	62.2 (77.0)	3619.8 (6725.2)
Word Length	5.5 (1.1)	5.3 (1.1)	5.4 (2.0)
Position in the sentence	5.4 (1.3)	6.4 (1.3)	7.4 (1.3)

Note. Word frequency is reported as the total CELEX frequency per million (Davis, 2005). Word length is the number of letters in a given word. Position in the sentence refers to the location in the sequence of words where a given word is presented. The number of words in each sentence is 11.6 ± 1.7 (mean ± SD). All values shown here are mean with standard deviations in the parentheses. Note that the pre-target and post-target were identical for the congruent and incongruent conditions. The target word was counterbalanced over items such that it was congruent for one item and incongruent for another.

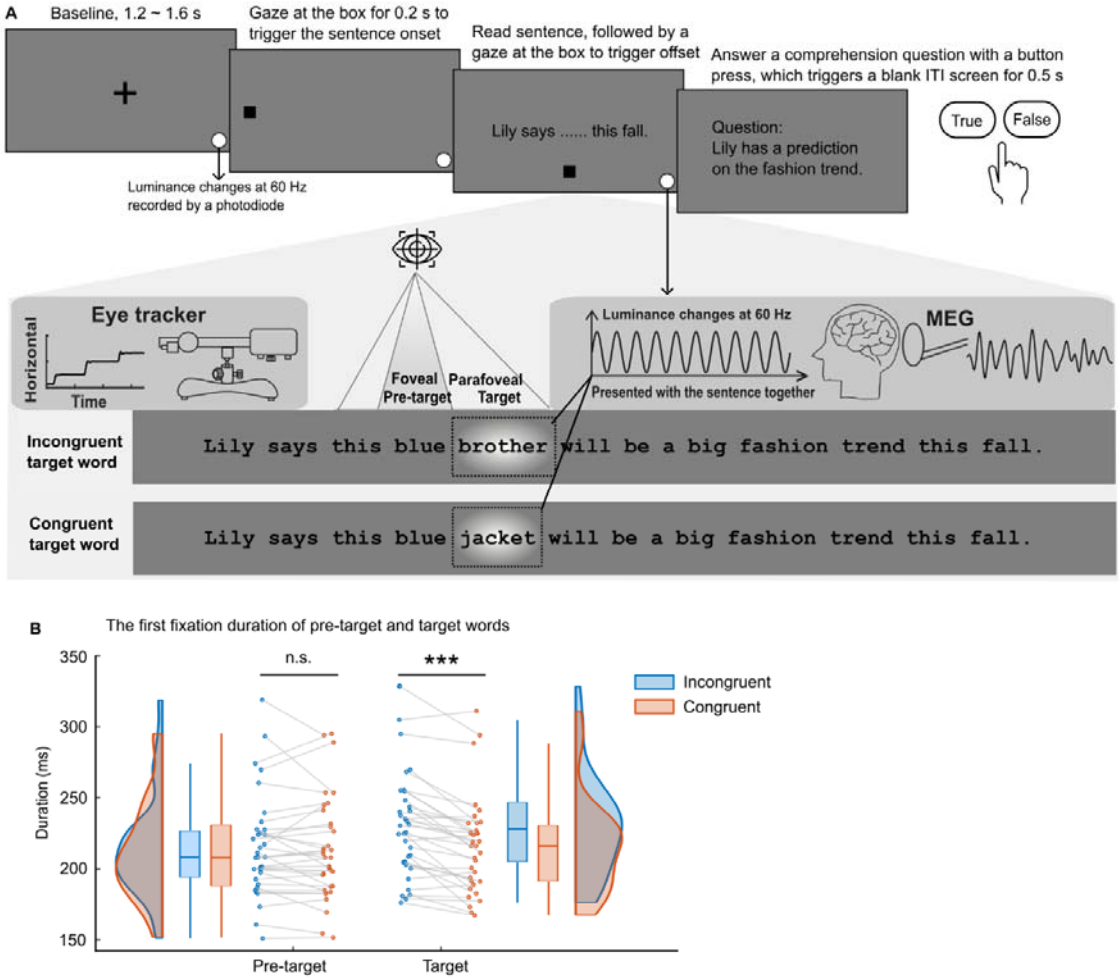


Figure 1. The paradigm and the eye movement metrics. (**A**) After the presentation of a cross-fixation at the screen centre for 1.2 −1.6 s, a gaze-contingent box appeared near the left edge of the screen. Fixating the box for 0.2 s triggered the full sentence presentation. Participants (n = 34) read 160 one-line sentences silently while brain activity and eye movements were recorded. Each sentence was embedded with one congruent or incongruent target word (see the dashed rectangle; not shown in the actual experiment). The target words could not be predicted based on the sentence context and word level properties of congruent and incongruent targets were balanced by swapping them between two sentence frames. The target words were tagged by changing the luminance of the underlying patch (with a Gaussian mask) in a 60 Hz sinusoid throughout the sentence presentation (depicted as a bright blob, not shown in the actual experiment). Additionally, we included a small disk at the bottom right of the screen that displayed the tagging signal and was recorded by a photodiode throughout each trial. After reading, gazing at the bottom box for 0.2 s triggered the sentence offset. Twelve percent of the sentences were followed by a simple yes-or-no comprehension question. (**B**) The first fixation durations on the pre-target and target words when the target words were incongruent (in blue) or congruent (in orange) with the sentence context. Each dot indicates one participant. ***$p < .001$; n.s., not statistically significant; ITI, inter-trial interval.

Results

No evidence for semantic parafoveal processing in the eye movement data

Like prior work measuring eye fixations during English reading (Inhoff, 1982; Inhoff and Rayner, 1980; Rayner et al., 2014, 1986), we found no evidence for parafoveal semantic processing in the eye movement data (Figure 1B, left). A paired t-test comparing first fixation durations on the pre-target word showed no effect of contextual (in)congruity ($t_{(33)} = .84$, $p = .407$, $d = .14$, two-sided). However, first fixation durations on the target word were significantly longer when they were incongruent (versus congruent) with the context ($t_{(33)} = 5.99$, $p = 9.83\times10^{-7}$, $d = 1.03$, two-sided pairwise *t*-test; Figure 1B, right). In addition, we found that the contextual congruity of target words affected later eye movement measures (i.e., total gaze duration and the likelihood of refixation after the first pass reading), with additional processing evident when the target words were incongruent with the context compared with when they were congruent (Supplementary Figure 1).

Parafoveal processing measured by Rapid Invisible Frequency Tagging (RIFT)

First, we performed a selection procedure to identify MEG sensors that responded to RIFT. We measured neural responses to the flickering target words by calculating the coherence between the MEG sensors and the tagging signal measured by a photodiode. A MEG sensor was considered a good tagging response sensor if it showed significantly stronger 60 Hz coherence during the pre-target intervals (with flicker) compared to the baseline intervals before the sentence presentation (without flicker). Both pre-target and baseline intervals were 1-second epochs. We then applied a cluster-based permutation test and identified sensor clusters that showed a robust tagging response ($p_{\text{cluster}} < .01$; Figure 2A). Tagging response sensors were found in 29 out of 34 participants, and all subsequent analyses were based on these tagging response sensors (7.9 ± 4.5 sensors per participant, M ± SD). The sources of these responses were localized to the left visual association cortex (Brodmann area 18; Figure 2B) using Dynamic Imaging Coherent Sources (DICS) (Gross et al., 2001).

Next, we characterized the temporal dynamics of attentional allocation to the flickering target word by calculating the 60 Hz coherence during fixations on several words surrounding the target word (Figure 2C). The resulting RIFT response curve revealed that significant attention was allocated to the target word as far as three words prior, spanning 15.3 ± 2.7 letters (M ± SD), including the spaces between words. This range is consistent with previous estimations of the perceptual span of 12−15 letters during English reading (McConkie and Rayner, 1975;

Rayner, 2009, 1975; Underwood and McConkie, 1985), as reported in the eye movement literature. Moreover, the RIFT response curve was left skewed, indicating a higher allocation of attentional resources to the flickering target words before fixating on them. The normal size and left skewness of the perceptual span in our study suggests that RIFT did not influence attention distribution during natural reading. Notably, the strongest RIFT responses were observed during fixations on the pre-target word (i.e., word position N-1, Figure 2C), highlighting the suitability of RIFT for measuring neuronal activity associated with parafoveal processing during natural reading.

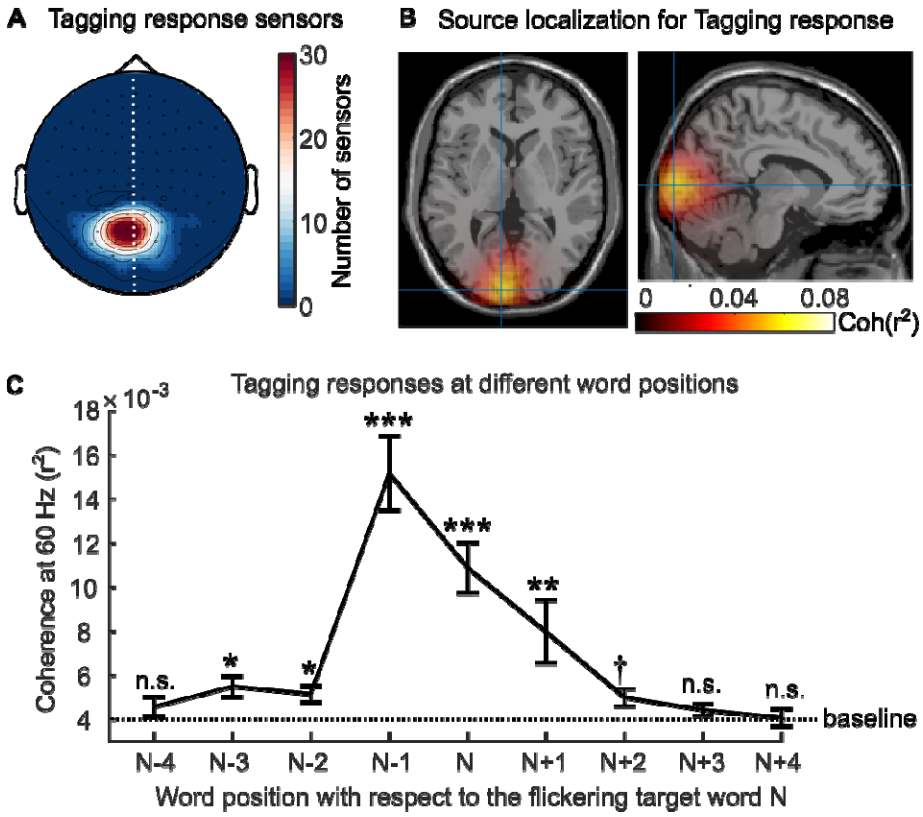


Figure 2. Neural responses to the rapid invisible frequency tagging (RIFT). (**A**) Topography of the RIFT response sensors over all participants (7.9 ± 4.5 sensors per participant, M ± SD). These sensors showed significantly stronger coherence to the tagging signal during the pre-target interval (with target words flickering in the parafovea) compared with the baseline interval (no flicker). Further analyses only included participants who had a RIFT response (n = 29). (**B**) The source of the RIFT response sensors was localized to the left visual association cortex (MNI coordinates [-9 -97 3] mm, Brodmann area 18). (**C**) The averaged 60 Hz coherence over the RIFT response sensors when participants fixated on words at different positions, where N indicates the target word and N-1 indicates the pre-target word. Error bars indicate the SE over participants (n = 29). The shaded area indicates the RIFT responses when previewing the flickering target words. We compared the RIFT response at each word position with the baseline (the dashed line). ***$p < .001$, **$p < .01$, *$p < .05$, † $p = .051$; n.s., not statistically significant.

Neural evidence for semantic parafoveal integration

Importantly, evidence for parafoveal semantic integration was found using RIFT (Figure 3). The pre-target coherence was weaker when the sentence contained a contextually incongruent word, compared to when it was congruent (Figure 3A). We conducted a pairwise t-test and found a significant effect on the averaged pre-target coherence at 60 Hz ($t_{(28)}$ = -2.561, p = .016, d = .476, two-sided pairwise; Figure 3B). To avoid any contamination of the parafoveal measure with activity from target fixation, pre-target coherence was averaged over the minimum pre-target fixation duration across both conditions for each participant (97.4 ± 14.1 ms, M ± SD, denoted as a dashed rectangle). Next, we conducted a jackknife-based latency estimation and found that the congruency effect on the 60 Hz pre-target coherence had a significantly later onset when previewing an incongruent (116.0 ± 1.9 ms, M ± SD) compared to a congruent target word (91.4 ± 2.1 ms, M ± SD, denoted as a dashed rectangle; $t_{(28)}$ = -2.172, p = .039, two-sided; Figure 3C). Therefore, both the magnitude and onset latency of the pre-target coherence were modulated by the contextual congruency of the target word, providing neural evidence that semantic information is integrated into the context during parafoveal processing, detectable within 100 ms after readers fixate the pre-target word.

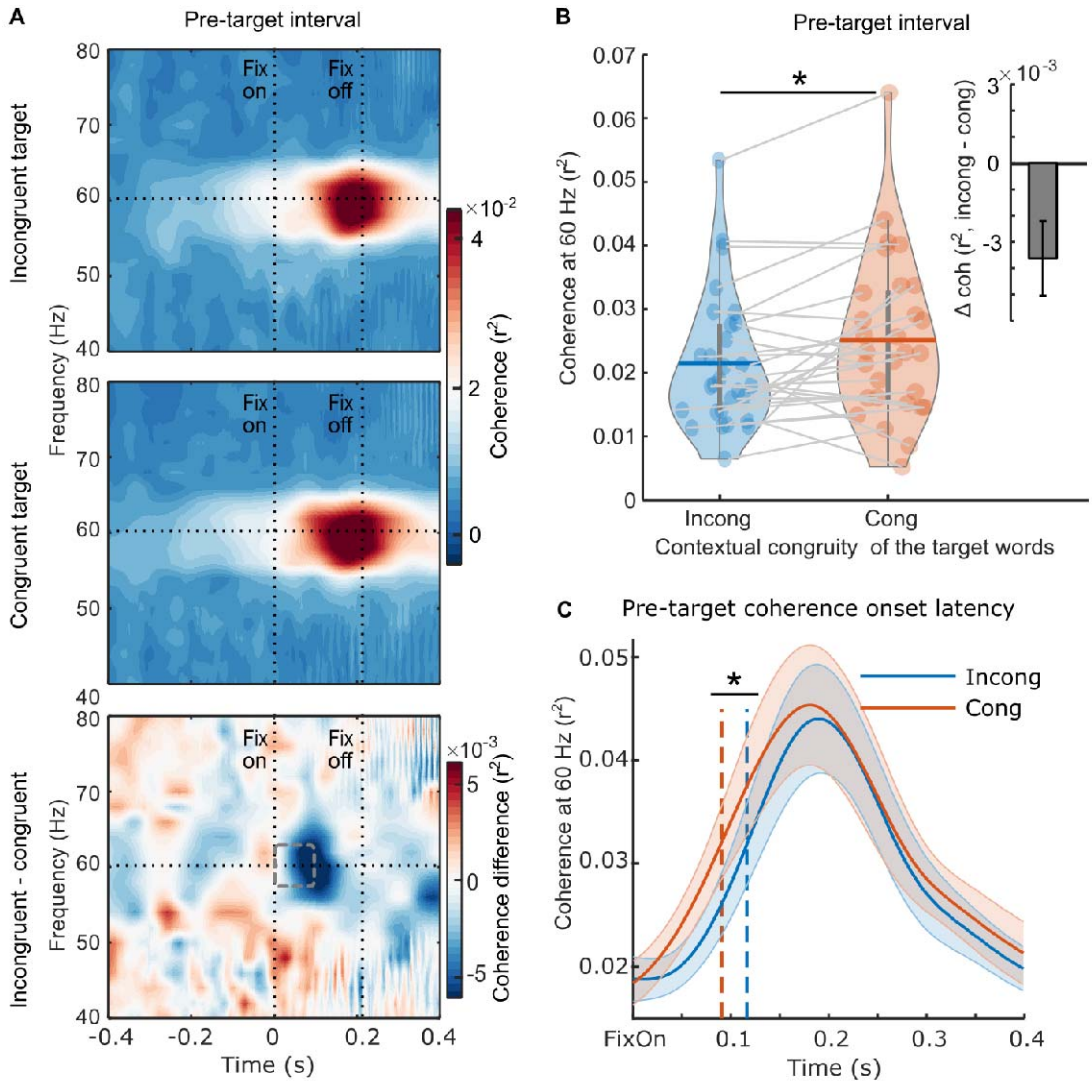


Figure 3. Neural evidence for parafoveal semantic integration. (**A**) The pre-target coherence spectrum averaged over the RIFT response sensors at the group level (n = 29) when the subsequent target words were incongruent with the sentence context (top panel), congruent with the sentence context (middle panel), and the difference between the two conditions (bottom panel). The horizontal line indicates the tagging frequency at 60 Hz. The two vertical lines indicate the first fixation onset of the pre-target words and the average fixation offset. (**B**) The averaged 60 Hz coherence during the minimum pre-target intervals for each participant (97.4 ± 14.1 ms, M ± SD; denoted as a dashed rectangle) with respect to the incongruent and congruent target words. Each dot indicates one participant, the horizontal lines inside of the violins indicate the mean values. The upright inserted figure shows the pre-target coherence difference over participants with the error bar as SE. (**C**) The onset latency of the pre-target coherence at the group level (n = 29). The onset latency refers to the time when the coherence curve reached its half maximum, denoted by the dotted lines. Zero time-point indicates the first fixation onset of the pre-target words. The shaded area shows SE around the mean value. *p < .05.

We conducted a similar analysis of the coherence measured when participants fixated the target word and found no significant modulations related to the contextual congruity of that target word, in either the magnitude ($t_{(28)}$ = .499, p = .622, d = .093, two-sided pairwise) or onset latency ($t_{(28)}$ = -.280, p = .782); Figure 4) of the RIFT response. Thus, the parafoveal semantic integration effect identified during the pre-target intervals cannot be attributed to signal contamination from fixations on the target word induced by the temporal smoothing of filters.

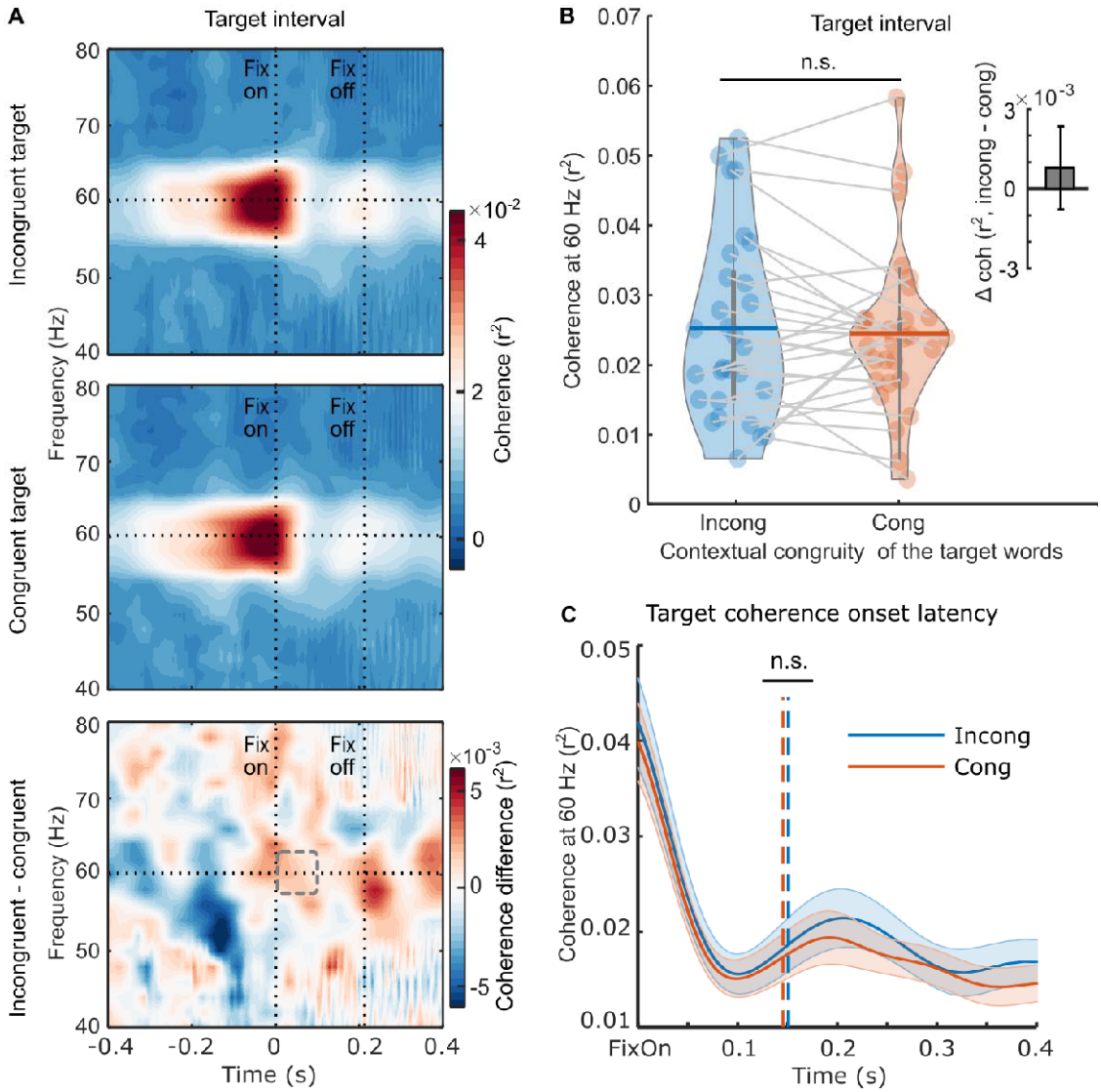


Figure 4. Neural responses to the rapid invisible frequency tagging (RIFT) during the target interval. (**A**) The target coherence spectrum averaged over the RIFT response sensors

at the group level (n = 29) when the target words were incongruent with the sentence context (top panel), or congruent with the sentence context (middle panel); the bottom panel shows the difference between the two conditions. The horizontal line indicates the tagging frequency at 60 Hz. The two vertical lines indicate the first fixation onset of the target words and the averaged fixation offset. (**B**) We averaged the 60 Hz coherence within the minimum target fixation duration over participants (97.6 ± 15.7 ms, M ± SD; denoted as a dashed rectangle). Each dot indicates one participant, the horizontal lines inside of the violins indicate the mean values. The upright inserted figure shows the target coherence difference over participants with the error bar as SE. (**C**) A jackknife-based method was used to calculate the onset latency of the average coherence at the group level. The onset latency refers to the time when the coherence curve reached its half maximum, denoted by the dotted lines. n.s., not statistically significant.

Parafoveal semantic integration is related to individual reading speed

The RIFT effects of congruency observed during parafoveal preview of the targets showed that readers tend to allocate less attention to upcoming text when an upcoming word is semantically incongruent compared to when all words are congruent. If readers differ in the extent to which their attention is "repelled" by incongruent words, then we might expect that the magnitude of the RIFT effect would be related to reading speed. Therefore, we conducted a correlation analysis to investigate this relationship (Figure 5). Individual reading speed was quantified as the number of words read per second from the congruent sentences in the study. We found a positive correlation between the pre-target coherence difference (incongruent - congruent) and individual reading speed ($r_{(27)} = .503$, $p = .006$; Spearman's correlation). This suggests that readers who show greater shifts in attentional allocation in response to semantic incongruity read more slowly on average.

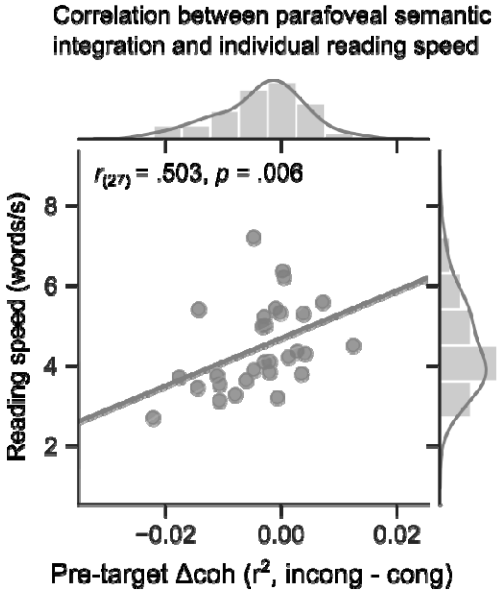


Figure 5. Individual reading speed is correlated with the magnitude of the RIFT congruency effect. Reading speed was measured as the number of words read per second in the congruent sentences. The RIFT effect was measured as the coherence difference during the pre-target fixations for sentences containing incongruent and congruent target words. Each dot indicates one participant (n = 29, Spearman correlation). The shaded area represents the 95% CI.

Discussion

In the current natural reading study, we utilized RIFT to probe for evidence that readers are sensitive to the effect of contextual congruity of an upcoming target word during parafoveal processing. We found no significant modulation of fixation durations of pre-target words based on the contextual congruity of the target word (Figure 1B). However, we observed a significant difference in the amount of covert attention allocated to the target when previewing congruent and incongruent target words (Figure 3). Specifically, we found lower RIFT responses for parafoveal words that were incongruent compared to congruent with the sentence context. Because the target words were always of low predictability, their semantic congruence could only be appreciated if they had been integrated (to some extent) with the unfolding context. Thus, the RIFT patterns provide compelling neural evidence that semantic information can not only be extracted but also integrated during parafoveal processing.

More specifically, we observed that pre-target coherence was weaker in magnitude (Figure 3B) and had a later onset latency (Figure 3C) in response to a contextually incongruent target word compared to a congruent one. Two possible explanations for these findings can be considered. First, the decreased RIFT responses may be due to changes in the pattern of allocation of attention across the text during reading. When reading in English, attention continuously shifts from left to right. If the semantic information previewed in the parafovea cannot be easily integrated into the context, this pattern may be interrupted, leading to delayed and/or reduced allocation of attention to the parafoveal word, possibly because readers shift more attention to the currently fixated word or to previous words to ensure that they have decoded and understood what they have read thus far. On this view, the RIFT finding may reflect a covert “regression” of attention, similar to overt eye-movement regressions that sometimes occur when readers encounter semantically incongruous words (Antúnez et al., 2022; Braze et al., 2002; Ni et al., 1998; Rayner et al., 2004) (also see Supplementary Figure 1A). Alternatively, the reduction in RIFT responses could arise if readers shift attentional resources away from the text altogether. Previous work has

demonstrated that tagging responses decrease as attention shifts from an external task (e.g., counting visual targets) to an internal task (e.g., counting heartbeats) (Kritzman et al., 2022). Similarly, in a reading scenario, visually perceiving the flickering word constitutes an external task, while the internal task involves the semantic integration of previewed information into the context. If more attentional resources are internally directed when faced with the challenge of integrating a contextually incongruent word, fewer attentional resources would remain for processing the flickering word. This may be the kind of shift reflected in the reduction in RIFT responses. On either account, the reduced forward allocation of attention diminishes parafoveal processing, and, in turn, may tend to slow reading speed, as supported by our correlation results (Figure 5).

Our results also provide information about the time course of semantic integration, as we found evidence that readers appreciated the incongruity – and thus must have begun to integrate the semantics of the parafoveal words with their ongoing message-level representation – by as early as within 100 ms after fixating on the pre-target word. The timing of this parafoveal semantic effect appears remarkably early, considering that typical semantic access for a single word occurs no earlier than around 200 ms, as demonstrated in the visual word recognition literature (Carreiras et al., 2014). For instance, in a Go/NoGo paradigm, the earliest distinguishable brain activity related to category-related semantic information of a word occurs at 160 ms (Amsel et al., 2013; Hauk et al., 2012). Therefore, the RIFT results presented here suggest that natural reading involves parallel processing that spans multiple words. The level of (covert) attention allocated to the target word, as indexed by the significant difference in RIFT responses compared to the baseline interval, was observed even three words in advance (see Figure 2C). This initial increase in RIFT coincided with the target entering the perceptual span (McConkie and Rayner, 1975; Rayner, 1975; Underwood and McConkie, 1985), likely aligning with the initial extraction of lower-level perceptual information about the target. The emerging sensitivity of the RIFT signal to target plausibility, detected around 100 ms after the fixation on the pre-target word, suggests that readers at that time had accumulated sufficient semantic information about the target words and integrated that information with the evolving sentence context. Therefore, it is plausible that the initial semantic processing of the target word commenced even before the pre-target fixation and was distributed across multiple words. This parallel processing of multiple words facilitates rapid and fluent reading.

Our findings have significant implications for theories of reading. The occurrence and early onset of semantic integration in parafoveal vision suggests that words are processed in an exceptionally parallel manner, posing a challenge for existing serial processing models (Reichle et al., 2009, 2006, 2003, 1998). At the same time, it is important to note that the fact that semantic integration begins in parafoveal vision does not mean that it is necessarily completed before a word is fixated. The fact that we observed semantic congruency effects on the fixation durations of the target words (Figure 1B) suggests that additional processing is required to fully integrate the semantics with overt attention in foveal vision. This also aligns with previous studies that found some ERP responses to semantic violations, including the LPC (Late Positive Component), are elicited only during foveal processing, but not during parafoveal processing (Li et al., 2023; Milligan et al., 2023; Payne et al., 2019; Schotter et al., 2023).

Thus, RIFT measures complement eye tracking (and other) measures, providing unique information revealing multiple mechanisms at work during natural reading. The results of the present study are aligned with the SWIFT model of eye movement control in natural reading (Engbert et al., 2005), wherein the activation field linked to a given word is hypothesized to be both temporally and spatially distributed. Indeed, we found that the initial increase in covert attention to the target word occurred as early as three words before, as measured by RIFT responses (Figure 2C). These covert processes enable the detection of semantic incongruity (Figure 3B and Figure 3C). However, it may occur at the non-labile stage of saccade programming, preventing its manifestation in fixation measures of the currently fixated pre-target word (Figure 1B). Therefore, the RIFT technique's capacity to yoke patterns to a specific word offers a unique opportunity to track the activation field of word processing during natural reading. Additional processes, which do impact overt eye movement patterns, are then brought to bear when the target words are fixated, resulting in increased fixation durations for incongruous words. At that same point (i.e., the target word), however, the RIFT responses showed a null effect of congruency (Figure 4); it may be that the RIFT technique is better suited to capturing parafoveal compared to foveal attentional processes, in part because there are more motion-sensitive rod cells in the parafoveal than foveal area. Finally, even after readers move away from fixating the word, attention to the target can persist or be reinstated, as evidenced by patterns of regressions (Supplementary Figure 1A). Therefore, during natural reading, attention is distributed across multiple words. The highly flexible and distributed allocation of attention allows readers to be parallel

processors and thereby read fluently and effectively (Engbert et al., 2005, 2002; Snell et al., 2018; Snell and Grainger, 2019). Our natural reading paradigm, where all words are available on the screen and saccadic eye movements are allowed, makes it possible to capture the extensive parallel processing. Moreover, saccades have been found to coordinate our visual and oculomotor systems, further supporting the parallel processing of multiple words during natural reading (Pan et al., 2023).

Two noteworthy limitations exist in the current study. Firstly, the construction of pretarget–target word pairs consistently follows an adjective–noun phrase structure, potentially leading to semantic violations arising from immediate local incongruence rather than a broader incongruence derived from the entire sentential context. While the context preceding target words was deliberately minimized to ensure a pure effect of bottom-up parafoveal processing rather than the confounding impact of top-down prediction, it is essential to recognize that information from both local and global contexts can exert distinct effects on word processing during natural reading (Wong et al., 2022). Future investigations should incorporate more information-rich contexts to explore the extent to which the parafoveal semantic integration effect observed in this study can be generalized. Secondly, the correlation analysis between the pre-target RIFT effect and individual reading speed (Figure 5) does not establish a causal relationship between parafoveal semantic integration and reading performance. Given that the comprehension questions in the current study were designed primarily to maintain readers' attention and the behavioural performance reached a ceiling level, employing more intricate comprehension questions in future studies would be ideal to accurately measure reading comprehension and reveal the impact of semantic parafoveal processing on it.

In summary, our findings show that parafoveal processing is not limited to simply extracting word information, such as lexical features, as demonstrated in our previous study (Pan et al., 2021). Instead, the previewed parafoveal information from a given word can begin to be integrated into the unfolding sentence representation well before that word is fixated. Moreover, the impact of that parafoveal integration further interacts with reading comprehension by shaping the time course and distribution of attentional allocation – i.e., by causing readers to move attention away from upcoming words that are semantically incongruous. These results support the idea that words are processed in parallel and suggest that early and deep parafoveal processing may be important for fluent reading.

Materials and Methods

Participants

We recruited 36 native English speakers (24 females, 22.5 ± 2.8 years old, mean ± SD) with normal or corrected-to-normal vision. All participants are right-handed and without any history of neurological problems or a language disorder diagnosis. Two participants were excluded from analysis due to poor eye tracking or falling asleep during the recordings, which left 34 participants (23 females). The study was approved by the University of Birmingham Ethics Committee. The informed consent form was signed by all participants after the nature and possible consequences of the studies were explained. Participants received £15 per hour or course credits as compensation for their participation.

Stimuli

In total participants read 277 sentences, of which 117 sentences were fillers from a published paper (White, 2008). The filler sentences were all plausible and were included to make sure the incongruent sentences were less than one third of the sentence set. We constructed the remaining 160 sentences with 80 pairs of target words. In all sentences the context was low constraint; i.e., none of the target words could be predicted by the prior context (see Behavioural pre-tests below for details). The target word in each sentence was either incongruent or congruent with the sentence. To focus on semantic integration and avoid any confounds of word-level properties, we embedded each pair of target words in two different sentence frames. By swapping the target words within a pair of sentences, we created four sentences: two congruent ones and two incongruent ones. These were then counterbalanced over participants. In this way, we counterbalanced across lexical characteristics of the target words and characteristics of the sentence frames within each pair. Each participants read one version of the sentence set (A or B). For example, for the target pair *brother/jacket*, one participant read them in the congruent condition in the sentence set version A; while another participant read them in the incongruent condition in version B (see below, targets are in italic type for illustration, but in normal type in the real experiment).

A. Last night, my lazy *brother* came to the party one minute before it was over.

Lily says this blue *jacket* will be a big fashion trend this fall.

B. Last night, my lazy *jacket* came to the party one minute before it was over.

Lily says this blue *brother* will be a big fashion trend this fall

For all sentences, the pre-target words were adjectives, and the target words were nouns (for detailed characteristics of the words please see Table 1). The word length of pre-target words was from 4 to 8 letters, and for target words was from 4 to 7 letters. The sentences were no longer than 15 words or 85 letters. The target words were embedded somewhere in the middle of each sentence and were never the first three or the last three words in a sentence. Please see Appendix in the Supplementary material for the full list of the sentence sets that were used in the current study.

Behavioural pre-tests

We recruited native English speakers for two behavioural pre-tests of the sentence sets. These participants did not participate in the MEG session.

Predictability of target words

We carried out a cloze test to estimate the predictability of the target words and the contextual constraint of the sentences. Participants read sentence fragments consisting of the experimental materials up to but not including the target words. Then participants were asked to write down the first word that came to mind that could continue the sentence (no need to complete the whole sentence). Example:

1. Last night, my lazy _____
2. Lily says this blue _____

The predictability of a word was estimated as the percentage of participants who wrote down exactly this word in the cloze test. A target word with less than 10% predictability was deemed to be not predicted by the sentence context. In addition, sentences for which no word was predicted with 50% or greater probability were low constraint. Twenty participants (6 males, 24.2 ± 2.0 years old, mean ± SD) took part in the first round of pre-test. Eight sentences were replaced with new sentences because the target words were too predictable and/or the sentence was too constraining. We then conducted a second round of the predictability test with 21 new participants (7 males, 25.0 ± 6.0 years old). None of the target words in this final set were predictable (2.3% ± 4.8%, mean ± SD), and all the sentence contexts were low constraint (25.2% ± 11.8%).

Plausibility of sentences

Two groups of participants were instructed to rate how plausible (or acceptable) each sentence was in the sentence set version A or B separately. Plausibility was rated on a 7-point scale with plausibility increasing from point 1 to 7. Sentences in the experiment were designed to be either highly implausible (the incongruent condition) or highly plausible (the congruent condition). To occupy the full range of the scale, we constructed 70 filler sentences with middle plausibility (e.g., sentence 1 below). In this example, sentences 2 and 3 were the incongruent and congruent sentences from the experiment.

	Implausible 1	2	3	4	5	6	Plausible 7
1. Kate said that she saw lots of stars twinkling in the sky at noon.	1	2	3	4	5	6	7
2. Lily says this blue brother will be a big fashion trend this fall.	1	2	3	4	5	6	7
3. Little Jimmy picked up a box and put some coins inside of it.	1	2	3	4	5	6	7
……	1	2	3	4	5	6	7

For version A we recruited 27 participants (4 males, 22.8 ± 6.1 years old, mean ± SD): The plausibility rating for the incongruent sentences was 2.08 ± 0.79 (mean ± SD); while for the congruent sentences was 6.18 ± 0.56. For sentence set version B we recruited 22 participants (4 males, 21.1 ± 2.3 years old, one invalid dataset due to incomplete responses): The plausibility rating was 1.81 ± 0.41 (mean ± SD) for the sentences in the incongruent condition and 6.15 ± 0.47 for the sentences in the congruent condition. These results showed that in both versions of the sentences set, incongruent sentences were viewed as highly implausible and congruent sentences as highly plausible.

Experimental procedure

Participants were seated 145 cm away from the projection screen in a dimly lit magnetically shielded room. The MEG gantry was set at 60 degrees upright and covered the participant's whole head. We programmed in Psychophysics Toolbox -3(Kleiner et al., 2007) to present the one-line sentences on a middle-grey screen (RGB [128 128 128]). All words were displayed in black (RGB [0 0 0]) with an equal-spaced Courier New font. The font size was 20 and the font type was bold so that each letter and space occupied 0.316 visual degrees. The visual angle of the whole sentence was no longer than 27 visual degrees in the horizontal

direction. The sentence set was divided into 5 blocks, each of which took about 7 minutes. There was a break of at least 1 minute between blocks and participants pressed a button to continue the experiment at any time afterwards. Participants were instructed to read each sentence silently at their own pace and to keep their heads and body as stable as possible during the MEG session. Eye movements were acquired during the whole session. In total, the experiment took no longer than 55 minutes. While the current study was conducted using MEG, these procedures might also work with EEG. If so, this would make our approach accessible to more laboratories as EEG is less expensive. However, there are currently no studies directly comparing the RIFT response in EEG versus MEG. Therefore, it would be of great interest to investigate if the current findings can be replicated using EEG.

Within a trial, there was first a fixation cross presented at the centre of a middle-grey screen for 1.2 – 1.6 s. This was followed by a black square with a radius of 1 degree of visual angle. This square was placed at the vertical centre, 2 degrees of visual angle away from the left edge of the screen. Participants had to gaze at this black ‘starting square’ for at least 0.2 s to trigger the onset of the sentence presentation. Afterwards, the sentence would start from the location of the square (Figure 1A). The sentence was presented with an ‘ending square’ 5 degrees of visual angle below the screen centre. The ‘ending square’ was the same size as the ‘starting square’ but in grey colour (RGB [64 64 64]). A gaze at this ‘ending square’ for at least 0.1s would end the presentation of the sentence. Then the trial ended with a blank middle-grey screen that lasted for 0.5s. Randomly, 12% of the trials were followed by a statement about the content of the sentence that was just presented, and participants needed to answer “True or “False” by pressing a button. For example, the statement for sentence 2 was “Lily has a prediction about the fashion trend in this fall”, and the correct answer was “True”. The statement for sentence 3 was “Little Jimmy didn’t have a box”, and the correct answer was “False”. All participants read the sentences carefully as shown by the high accuracy of answering (96.3% ± 4.7%, mean ± SD).

Rapid invisible frequency tagging (RIFT)

Projection of the sentence stimuli

We projected the sentences from the stimulus computer screen in the experimenter room to the projection screen inside of the MEG room using a PROPixx DLP LED projector (VPixx Technologies Inc., Canada). The refresh rate of the PROPixx projector was up to 1440 Hz, while the refresh rate of the stimulus screen was only 120 Hz (1920 × 1200 pixels resolution).

We displayed the sentence repeatedly in four quadrants of the stimulus computer screen. In each quadrant, the words were coded in three colour channels as RGB. The projector then interpreted these 12 colour channels (3 channels × 4 quadrants) as 12 individual grayscale frames, which were projected onto the projection screen in rapid succession. Therefore, the projection screen refreshed at 12 times the rate of the stimulus computer screen.

Flickering of the target word

We added a square patch underneath the target word to frequency tag the target word. The side length of the square patch was the width of the target word plus the spaces on both sides (2 to 3° visual angle). We flickered the patch by changing its luminance from black to white at a 60 Hz sinusoid (Figure 1A). To reduce the visibility of the patch edges across saccades, we applied a Gaussian smoothed transparent mask on top of the square patch. The mask was created by a two-dimensional Gaussian function (Equation 1):

$$\text{mask} = \exp\left(-\frac{x^2}{2\sigma^2} - \frac{y^2}{2\sigma^2}\right)$$

where, x and y are the mesh grid coordinates for the flickering patch, and σ is the x and y spread of the mask with σ = 0.02 degrees.

On average, the patch was perceived as middle-grey, the same colour as the background screen, which made it invisible to participants. The target word was still black, the same colour as the other words on the screen. To record the tagging signal, we attached a custom-made photodiode (Aalto NeuroImaging Centre, Finland) to the disk at the bottom right corner of the screen. The luminance of the disk varied the same as that of the flickering patch underneath the target word. The photodiode was plugged into the MEG system as an external channel.

Data acquisition

MEG

Brain data were acquired with a 306-sensor TRIUX Elekta Neuromag system, which consisted of 204 orthogonal planar gradiometers and 102 magnetometers (Elekta, Finland). After participants signed the consent form, we attached four head-position indicator coils (HPI coils) to their heads: two on the left and right mastoid bone, and two on the forehead with at least 3 cm distance in between. Afterwards, we used a Polhemus Fastrack

electromagnetic digitizer system (Polhemus Inc, USA) to digitize the locations for three bony fiducial points: the nasion, left and right preauricular points. Then we digitized the four HPI coils. Furthermore, at least 200 extra points were acquired, which were distributed evenly and covered the whole scalp. These points were used later in the source analysis when spatially co-register the MEG head model with individual structural MRI images. The sampling rate of the MEG system was 1,000 Hz. Data were band-pass filtered prior to sampling from 0.1 to 330 Hz to reduce aliasing effects.

Eye movements

We used an EyeLink 1000 Plus eye-tracker (long-range mount, SR Research Ltd, Canada) to track eye movements throughout the whole MEG session. The eye tracker was placed on a wooden table in front of the projection screen. The centre of the eye tracker was at the middle line of the projection screen, and the top of the eye tracker reached the bottom edge of the screen. The distance between the eye-tracker camera and the centre of the participant's eyes was 90 cm. We recorded the horizontal and vertical positions as well as the pupil size from the left eye, at a sampling rate of 1,000 Hz. Each session began with a nine-point calibration and validation test. The test was accepted if the eye-tracking error was below 1 visual degree both horizontally and vertically. During the session, we performed a one-point drift checking test every three trials and after the break between blocks. If the drift checking failed or the sentence presentation was unable to be triggered through gazing, a nine-point calibration and validation test was conducted again.

MRI

After MEG data acquisition, participants were asked to come to the laboratory another day to have an MRI image acquired. We acquired the T1-weighted structural MRI image using a 3-Tesla Siemens PRISMA scanner (TR = 2000 ms, TE = 2.01 ms, TI = 880 ms, flip angle = 8 degrees, FOV = 256×256×208 mm, 1 mm isotropic voxel). For 11 participants who dropped out of the MRI acquisition, the MNI template brain (Montreal, Quebec, Canada) was used instead in the source analysis later.

Eye movement data analysis

We extracted the fixation onset events from the EyeLink output file. The EyeLink parsed fixation events based on the online detection of saccade onset using the following parameters: the motion threshold as 0.1 degrees, the velocity threshold as 30 degrees/sec, and the

acceleration threshold as 8000 degrees/sec2. These conservative settings were suggested by the EyeLink user manual for reading studies, as they can prevent false saccade reports and reduce the number of micro-saccades and lengthen fixation durations.

Only the fixation that first landed on a given word was selected. The first fixation durations were averaged within the incongruent and congruent conditions for pre-target and target words. Pairwise, two-sided *t*-test were conducted on the first fixation durations of pre-target and target words separately (conducted in R (Team, 2013)). In addition to the early eye movement measure of the first fixation duration, we also conducted *t*-tests for two later eye movement measures. The likelihood of refixation was measured as the proportion of trials on which there was at least one saccade that regressed back to that word. The total gaze duration was the sum of all fixations on a given word, including those fixations during regression or re-reading.

MEG data analyses

The data analyses were performed in MATLAB R2020a (Mathworks Inc, USA) by using the FieldTrip (Oostenveld et al., 2011) toolbox (version 20200220), following the FLUX MEG analysis pipeline(Ferrante et al., 2022), and custom-made scripts.

Pre-processing

We first band-pass filtered the MEG data from 0.5 to 100 Hz using phase preserving two-pass Butterworth filters. Subsequently, detrending was applied individually to each channel of the continuous raw data to factor out the linear trend. Malfunctioning sensors were removed based on inspecting the data quality during online recording (0 to 2 sensors per participant). Afterwards, the data were decomposed into independent components using an independent component analysis (ICA) (Ikeda and Toyama, 2000). The number of components was the same as the number of good MEG sensors in the dataset (306 or less). We only removed bad components that related to eye blinks, eye movements, and heartbeat by visually inspecting the components (3.4 ± 0.7 components per participant, M ± SD, range from 2 to 5 components).

MEG segments were extracted from -0.5 to 0.5 s intervals aligned with the first fixation onset of the pre-target and target words respectively (see Eye movement data analysis, above, for information on how fixation onsets were defined). Segments with fixation durations shorter than 0.08 s or longer than 1 s were discarded. We also extracted 1 s long baseline segments,

which were aligned with the cross-fixation onset before the sentence presentation. We manually inspected all segments to further identify and remove segments that were contaminated by muscle or movement artefacts.

Coherence calculation

We calculated the coherence between the MEG sensors and the photodiode (i.e., the tagging signal) to quantify the tagging responses. The amplitude of photodiode channel was normalized across each segment. To estimate the coherence spectrum in the frequency domain over time, we filtered the segments using hamming tapered Butterworth bandpass filters (4th order, phase preserving, two-pass). The frequency of interest was from 40 to 80 Hz in a step of 2 Hz. For each centre frequency point, the spectral smoothing was ± 5 Hz. For example, the filter frequency range for 60 Hz was from 55 to 65 Hz. We performed a Hilbert transform to obtain the analytic signals for each centre frequency point, which then were used to estimate the coherence (Equation 2):

$$\mathrm{coh}(t) = \frac{\left|\frac{1}{n} \sum_{j=1}^{n} m_{x_j}(t) m_{y_j}(t) e^{i\emptyset_{xy_j}(t)}\right|^2}{\sum_{j=1}^{n} m_{x_j}(t)^2 \, m_{y_j}(t)^2}$$

where n is the number of trials. For the time point t in the trial j, $m_x(t)$ and $m_y(t)$ are the time-varying magnitude of the analytic signals from a MEG sensor (x) and the photodiode (y) respectively, $\emptyset_{xy}(t)$ is the phase difference as a function of time (for detailed description, please see (Cohen, 2014).

Selection for the RIFT response sensors

MEG sensors that showed significantly stronger coherence at 60 Hz during the pre-target segments than the baseline segments were selected as the RIFT response sensors. We used a non-parametric Monte-Carlo method (Maris et al., 2007) to estimate the statistical significance. The pre-target segments were constructed by pooling the target contextual congruity conditions together. Several previous RIFT studies from our lab observed robust tagging responses from the visual cortex for flicker above 50 Hz (Drijvers et al., 2021; Duecker et al., 2021; Zhigalov et al., 2019; Zhigalov and Jensen, 2020). Thus, this sensor selection procedure was confined to the MEG sensors in the visual cortex (52 planar sensors). Here, the pre-target segments and baseline segments were treated as two conditions. For each combination of MEG sensor and photodiode channel, coherence at 60 Hz was estimated over

trials for the pre-target and baseline conditions separately. Then, we calculated the z-statistic value for the coherence difference between pre-target and baseline using the following equation (for details please see (Maris et al., 2007)) (Equation 3):

$$\mathrm{Z} = \frac{(\tanh^{-1}(|\mathrm{coh}_1|) - \mathrm{bias}_1) - (\tanh^{-1}(|\mathrm{coh}_2|) - \mathrm{bias}_2)}{\sqrt{\mathrm{bias}_1 + \mathrm{bias}_2}},$$

$$\mathrm{bias}_1 = \frac{1}{2n_1 - 2},\ \mathrm{bias}_2 = \frac{1}{2n_2 - 2}$$

where coh_1 and coh_2 denote the coherence value for pre-target and baseline segments, bias_1 and bias_2 is the term used to correct for the bias from trial numbers of the pre-target (n_1) and baseline condition (n_2). All trials from the pre-target and baseline conditions were used.

After obtaining the z statistic value for the empirical coherence difference, we ran a permutation procedure to estimate the statistical significance of this comparison. We randomly shuffled the trial labels between pre-target and baseline conditions 5,000 times. During each permutation, coherence was computed for both conditions (with shuffled labels), then entered Equation 3 to obtain a z score for the coherence difference. After all randomizations were performed, the resulting z-values established the null distribution. Since a tagging response sensor was supposed to have stronger coherence during the pre-target segments compared with the baseline segments, the statistical test was right sided. If the z-value of the empirical coherence difference was larger than 99% of z-values in the null distribution, this sensor was selected as the RIFT response sensor (right-sided, p = .01). For each participant, the coherence values were averaged over all sensors with significant tagging response to obtain an averaged coherence for further analyses. Please note that the tagging response sensors may vary in number across participants (7.9 ± 4.5 sensors per participant, M ± SD). Additionally, they may have a different but overlapping spatial layout, primarily over the visual cortex. For the topography of all tagging response sensors, please refer to Figure 2A.

Coherence response curve

We first extracted MEG segments for the words N-4, N-3, N-2, N+1, N+2, and N+3 following the same procedure described in the pre-processing when extracted MEG segments for the pre-target (N-1) and target words (N). All segments were 1 s long, aligned with the first fixation onset to the word. Then, we calculated the coherence at 60 Hz during these

segments for participants who have RIFT response sensors (n = 29). Next for each participant, the 60 Hz coherence was first averaged over the RIFT response sensors, then averaged within a time window of [0 0.2] s (the averaged fixation duration for words). We got an averaged 60 Hz coherence for the word at each position. We also got the 60 Hz coherence for the baseline interval averaged over [0 0.2] s, aligned with the cross-fixation onset. Then a pairwise *t* test was performed between the baseline coherence and the coherence at each word position.

Coherence comparison between conditions

The coherence comparison analyses were only conducted for the participants who had sensors with a reliable tagging response (n = 29). To avoid any bias from trial numbers, an equal number of trials under the different contextual congruity conditions was entered the coherence analysis per participant. We randomly discarded the redundant trials from the condition that had more trials for both the pre-target and target segments.

To compare the pre-target coherence amplitude between conditions, the coherence values at 60 Hz were averaged across the minimum fixation duration of all pre-target words. The time window for averaging was defined for each participant so that the coherence signal from the target fixation was not involved. Similarly, we averaged the 60 Hz coherence for the target segments over the minimum target fixation duration. Then, a two-sided pairwise Student's *t*-test was performed to estimate the statistical significance of the coherence difference as shown in Figure 3B and Figure 4B.

To assess the coherence onset latency difference between conditions, we used a leave-one-out Jackknife-based method (Miller et al., 1998). We extracted the 60 Hz coherence during the 1 s long pre-target segments for each participant. Then, during each iteration of participants, we randomly chose and left out one participant. For the remaining participants, coherences were calculated for the incongruent and congruent target conditions. Then, the coherence was averaged over the remaining participants to estimate the onset latency for both conditions. Here, the onset latency was defined as the time point when the averaged coherence value reached its half-maximum ($\mathrm{coh}_{\mathrm{min}} + (\mathrm{coh}_{\mathrm{max}} - \mathrm{coh}_{\mathrm{min}})/2$). We computed the onset latency difference by subtracting the onset latency for the incongruent target condition from the congruent condition. After all iterations, onset latency differences from all these subsamples were pooled together to estimate a standard error (S_D) using the following equation (Equation 4):

$$S_D = \sqrt{\frac{n-1}{n} \cdot \sum_{i=1}^{n} \left(D_{-i} - \bar{J}\right)^2}$$

where $\bar{J}$ is the average onset latency difference over all the subsamples, D_{-i} is the coherence difference obtained from the subsample when participant i was left out, n is the number of participants. We also computed the onset latency difference from the overall sample set (without leaving any participant out) and divided it by the S_D to obtain its *t*-value. A standard t table (pairwise, two-tailed) provided the statistical significance for the coherence onset latency difference between the incongruent and congruent target conditions. This procedure was conducted for both the pre-target and target segments as shown in Figure 3C and Figure 4C.

Source analysis for RIFT

We used a beamforming-based approach, Dynamic Imaging Coherent Sources (DICS) (Gross et al., 2001), to estimate the neural sources that generated the responses to RIFT. The DICS technique was applied to the pre-target segments (0 to 0.5 s aligned with fixation onset to the pre-target word) regardless of the target contextual congruity conditions, with a focus of 60 Hz in the frequency domain. In this source analysis, only participants with robust tagging responses were included (n = 29).

First, we constructed a semi-realistic head model, where spherical harmonic functions were used to fit the brain surface (Nolte, 2003). We aligned the individual structural MRI image with the head shape that was digitized during the MEG session. This was done by spatially co-registering the three fiducial anatomical markers (nasion, left and right ear canal) and extra points that covered the whole scalp. For participants whose MRI image was unavailable, the MNI template brain was used instead. The aligned MRI image was segmented into a grid, which was used to prepare the single-shell head model.

Next, we constructed the individual source model by inverse-warping a 5 mm spaced regular grid in the MNI template space to each participant's segmented MRI image. We got the regular grid from the Fieldtrip template folder, which was constructed before doing the source analysis. In this way, the beamformer spatial filters were constructed on the regular grid that mapped to the MNI template space. Even though after this warping procedure grid points in the individual native space were no longer evenly spaced, the homologous grid

points across participants were located at the same location in the normalized template space. Thus, the reconstructed sources can be directly averaged across participants on the group level.

Next, the Cross-Spectral Density (CSD) matrix was calculated at 60 Hz for both the pre-target and baseline segments. The CSD matrix was constructed for all possible combinations between the MEG sensors and the photodiode channel. No regularisation was performed to the CSD matrices (lambda = 0).

Finally, a common spatial filter was computed based on the individual single-shell head model, source model, and CSD matrices. This spatial filter was applied to both the pre-target and baseline CSD matrices for calculating the 60 Hz coherence. This was done by normalizing the magnitude of the summed CSD between the MEG sensor and the photodiode channel by their respective power. After the grand average over participants, the relative change for pre-target coherence was estimated as the following formula, $(coh_{pretarget} - coh_{baseline})/coh_{baseline}$.

Acknowledgments

We thank Jonathan L. Winter for providing help with the MEG recordings. The computations described in this paper were performed using the University of Birmingham's BlueBEAR HPC service, which provides a High Performance Computing service to the University's research community. See http://www.birmingham.ac.uk/bear for more details.

Funding

This study was supported by the following grants to O.J.: the James S. McDonnell Foundation Understanding Human Cognition Collaborative Award (grant number 220020448), Wellcome Trust Investigator Award in Science (grant number 207550), and the BBSRC grant (BB/R018723/1) as well as the Royal Society Wolfson Research Merit Award. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Y.P., S.F., K.D.F., and O.J. devised and designed the experiments, Y.P. made the sentences with assistance from S.F., Y.P. programmed and conducted the experiments, Y.P. carried out

the analyses with assistance from O.J. and S.F., Y.P., O.J., K.D.F., and S.F. wrote the paper together.

Competing interests

Authors declare that they have no competing interests.

Data availability

We have deposited the following data in the current study on figshare (https://figshare.com/projects/Semantic/149801): the epoch data after pre-processing, the raw EyeLink files, the Psychotoolbox data, and the head models after the co-registration of T1 images with the MEG data. Any additional information will be available from the authors upon reasonable request.

Code availability

The experiment presentation scripts (Psychtoolbox), statistics scripts (R), scripts and data to generate all figures (Matlab) are available on GitHub (https://github.com/yalipan666/Semantic).

References

Amsel BD, Urbach TP, Kutas M. 2013. Alive and grasping: Stable and rapid semantic access to an object category but not object graspability. *Neuroimage* **77**:1–13. doi:10.1016/J.NEUROIMAGE.2013.03.058

Antúnez M, Milligan S, Andrés Hernández-Cabrera J, Barber HA, Schotter ER. 2022. Semantic parafoveal processing in natural reading: Insight from fixation-related potentials & eye movements. *Psychophysiology* **59**:e13986. doi:10.1111/PSYP.13986

Ashby J, Rayner K. 2004. Representing syllable information during silent reading: Evidence from eye movements. *Lang Cogn Process* **19**:391–426. doi:10.1080/01690960344000233

Ashby J, Treiman R, Kessler B, Rayner K. 2006. Vowel processing during silent reading: Evidence from eye movements. *J Exp Psychol Learn Mem Cogn* **32**:416–424. doi:10.1037/0278-7393.32.2.416

Barber HA, Doñamayor N, Kutas M, Münte T. 2010. Parafoveal N400 effect during sentence reading. *Neurosci Lett* **479**:152–156. doi:10.1016/j.neulet.2010.05.053

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

Beyersmann E, Montani V, Ziegler JC, Grainger J, Stoianov IP. 2021. The dynamics of

reading complex words: evidence from steady-state visual evoked potentials. *Sci Reports 2021 111* **11**:1–14. doi:10.1038/s41598-021-95292-0

Braze D, Shankweiler D, Ni W, Palumbo LC. 2002. Readers' eye movements distinguish anomalies of form and content. *J Psycholinguist Res* **31**:25–44. doi:10.1023/A:1014324220455/METRICS

Brickwedde M, Bezsudnova Y, Kowalczyk A, Jensen O, Zhigalov A. 2022. Application of rapid invisible frequency tagging for brain computer interfaces. *J Neurosci Methods* **382**:109726. doi:10.1016/J.JNEUMETH.2022.109726

Carreiras M, Armstrong BC, Perea M, Frost R. 2014. The what, when, where, and how of visual word recognition. *Trends Cogn Sci* **18**:90–98. doi:10.1016/j.tics.2013.11.005

Chace KH, Rayner K, Well AD. 2005. Eye movements and phonological parafoveal preview: Effects of reading skill. *Can J Exp Psychol* **59**:209–217. doi:10.1037/h0087476

Cohen MX. 2014. Analyzing neural time series data: Theory and practice. MIT Press.

Davis CJ. 2005. N-watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behav Res Methods* **37**:65–70. doi:10.3758/BF03206399

DeLong KA, Quante L, Kutas M. 2014. Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia* **61**:150–162. doi:10.1016/J.NEUROPSYCHOLOGIA.2014.06.016

Drieghe D, Rayner K, Pollatsek A. 2005. Eye movements and word skipping during reading revisited. *J Exp Psychol Hum Percept Perform* **31**:954–969. doi:10.1037/0096-1523.31.5.954

Drijvers L, Jensen O, Spaak E. 2021. Rapid invisible frequency tagging reveals nonlinear integration of auditory and visual information. *Hum Brain Mapp* **42**:1138–1152. doi:10.1002/HBM.25282

Duecker K, Gutteling TP, Herrmann CS, Jensen O. 2021. No Evidence for Entrainment: Endogenous Gamma Oscillations and Rhythmic Flicker Responses Coexist in Visual Cortex. *J Neurosci* **41**:6684–6698. doi:10.1523/JNEUROSCI.3134-20.2021

Engbert R, Longtin A, Kliegl R. 2002. A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Res* **42**:621–636. doi:10.1016/S0042-6989(01)00301-7

Engbert R, Nuthmann A, Richter EM, Kliegl R. 2005. Swift: A dynamical model of saccade generation during reading. *Psychol Rev* **112**:777–813. doi:10.1037/0033-295X.112.4.777

Federmeier KD. 2022. Connecting and considering: Electrophysiology provides insights into comprehension. *Psychophysiology* **59**:e13940.

Federmeier KD, Wlotko EW, De Ochoa-Dewald E, Kutas M. 2007. Multiple effects of sentential constraint on word processing. *Brain Res* **1146**:75–84. doi:10.1016/j.brainres.2006.06.101

Ferrante O, Liu L, Minarik T, Gorska U, Ghafari T, Luo H, Jensen O. 2022. FLUX: A pipeline for MEG analysis. *Neuroimage* **253**:119047. doi:10.1016/J.NEUROIMAGE.2022.119047

Ferrante O, Zhigalov A, Hickey C, Jensen O. 2023. Statistical Learning of Distractor Suppression Down-regulates Pre-Stimulus Neural Excitability in Early Visual Cortex. *J*

Neurosci JN-RM-1703-22. doi:10.1523/JNEUROSCI.1703-22.2022

Gross J, Kujala J, Hämäläinen M, Timmermann L, Schnitzler A, Salmelin R. 2001. Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proc Natl Acad Sci U S A* **98**:694–699. doi:10.1073/pnas.98.2.694

Gulbinaite R, Roozendaal DHM, VanRullen R. 2019. Attention differentially modulates the amplitude of resonance frequencies in the visual cortex. *Neuroimage* **203**:1–40. doi:10.1016/j.neuroimage.2019.116146

Gutteling TP, Sillekens L, Lavie N, Jensen O. 2022. Alpha oscillations reflect suppression of distractors with increased perceptual load. *Prog Neurobiol* **214**:102285. doi:10.1016/J.PNEUROBIO.2022.102285

Hauk O, Coutout C, Holden A, Chen Y. 2012. The time-course of single-word reading: Evidence from fast behavioral and brain responses. *Neuroimage* **60**:1462. doi:10.1016/J.NEUROIMAGE.2012.01.061

Hohenstein S, Klieg R. 2014. Semantic preview benefit during reading. *J Exp Psychol Learn Mem Cogn* **40**:166–190. doi:10.1037/a0033670

Hohenstein S, Laubrock J, Kliegl R. 2010. Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *J Exp Psychol Learn Mem Cogn* **36**:1150–1170. doi:10.1037/a0020233

Ikeda S, Toyama K. 2000. Independent component analysis for noisy data - MEG data analysis. *Neural Networks* **13**:1063–1074. doi:10.1016/S0893-6080(00)00071-X

Inhoff AW. 1989. Parafoveal processing of words and saccade computation during eye fixations in reading. *J Exp Psychol Hum Percept Perform* **15**:544–555.

Inhoff AW. 1982. Parafoveal word perception: A further case against semantic preprocessing. *J Exp Psychol Hum Percept Perform* **8**:137–145.

Inhoff AW, Rayner K. 1980. Parafoveal word perception: A case against semantic preprocessing. *Percept Psychophys* **27**:457–464. doi:10.3758/BF03204463

Jensen O, Pan Y, Frisson S, Wang L. 2021. An oscillatory pipelining mechanism supporting previewing during visual exploration and reading. *Trends Cogn Sci*.

Johnson RL, Perea M, Rayner K. 2007. Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *J Exp Psychol Hum Percept Perform* **33**:209–229. doi:10.1037/0096-1523.33.1.209

Kennedy A, Pynte J. 2005. Parafoveal-on-foveal effects in normal reading. *Vision Res* **45**:153–168. doi:10.1016/j.visres.2004.07.037

Kleiner M, Brainard D, Pelli D. 2007. What's new in Psychtoolbox-3? **14**.

Kliegl R, Nuthmann A, Engbert R. 2006. Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *J Exp Psychol Gen* **135**:12–35. doi:10.1037/0096-3445.135.1.12

Kritzman L, Eidelman-Rothman M, Keil A, Freche D, Sheppes G, Levit-Binnun N. 2022. Steady-state visual evoked potentials differentiate between internally and externally directed attention. *Neuroimage* **254**:119133.

Kutas M, Federmeier KD. 2011. Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annu Rev Psychol* **62**:621–647. doi:10.1146/annurev.psych.093008.131123

Kutas M, Hillyard SA. 1984. Brain potentials during reading reflect word expectancy and semantic association. *Nature* **307**:161–163. doi:10.1038/307161a0

Kutas M, Hillyard SA. 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science (80-)* **207**:203–205. doi:10.1126/science.7350657

Lau EF, Phillips C, Poeppel D. 2008. A cortical network for semantics: (De)constructing the N400. *Nat Rev Neurosci* **9**:920–933. doi:10.1038/nrn2532

Li C, Midgley KJ, Holcomb PJ. 2023. ERPs reveal how semantic and syntactic processing unfold across parafoveal and foveal vision during sentence comprehension. *Lang Cogn Neurosci* **38**:88–104. doi:10.1080/23273798.2022.2091150

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

López-Peréz PJ, Dampuré J, Hernández-Cabrera JA, Barber HA. 2016. Semantic parafoveal-on-foveal effects and preview benefits in reading: Evidence from Fixation Related Potentials. *Brain Lang* **162**:29–34. doi:10.1016/j.bandl.2016.07.009

Maris E, Schoffelen JM, Fries P. 2007. Nonparametric statistical testing of coherence differences. *J Neurosci Methods* **163**:161–175. doi:10.1016/j.jneumeth.2007.02.011

McConkie GW, Rayner K. 1975. The span of the effective stimulus during a fixation in reading. *Percept Psychophys* **17**:578–586. doi:10.3758/BF03203972

Meade G, Declerck M, Holcomb PJ, Grainger J. 2021. Parallel semantic processing in the flankers task: Evidence from the N400. *Brain Lang* **219**:104965. doi:10.1016/J.BANDL.2021.104965

Miellet S, Sparrow L. 2004. Phonological codes are assembled before word fixation: Evidence from boundary paradigm in sentence reading. *Brain Lang* **90**:299–310. doi:10.1016/S0093-934X(03)00442-5

Miller J, Patterson T, Ulrich R. 1998. Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology* **35**:99–115. doi:10.1017/S0048577298000857

Milligan S, Nestor B, Antúnez M, Schotter ER. 2023. Out of Sight, Out of Mind: Foveal Processing is Necessary for Semantic Integration of Words into Sentence Context. *J Exp Psychol Hum Percept Perform*.

Montani V, Chanoine V, Stoianov IP, Grainger J, Ziegler JC. 2019. Steady state visual evoked potentials in reading aloud: Effects of lexicality, frequency and orthographic familiarity. *Brain Lang* **192**:1–14. doi:10.1016/j.bandl.2019.01.004

Müller MM, Malinowski P, Gruber T, Hillyard SA. 2003. Sustained division of the attentional spotlight. *Nature* **424**:309–312. doi:10.1038/nature01812

Müller MM, Picton TW, Valdes-Sosa P, Riera J, Teder-Sälejärvi WA, Hillyard SA. 1998. Effects of spatial selective attention on the steady-state visual evoked potential in the 20–28 Hz range. *Cogn Brain Res* **6**:249–261. doi:10.1016/S0926-6410(97)00036-0

Ni W, Fodor JD, Crain S, Shankweiler D. 1998. Anomaly Detection: Eye Movement Patterns. *J Psycholinguist Res* **27**:515–539. doi:10.1023/A:1024996828734/METRICS

Nolte G. 2003. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoenchephalography forward calculation in realistic volume conductors. *Phys Med Biol* **48**:3637–3652. doi:10.1088/0031-9155/48/22/002

Norcia AM, Gregory Appelbaum L, Ales JM, Cottereau BR, Rossion B. 2015. The steady-state visual evoked potential in vision research: A review. *J Vis* **15**:4–4. doi:10.1167/15.6.4

Oostenveld R, Fries P, Maris E, Schoffelen J-M. 2011. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput Intell Neurosci* **2011**.

Pan Y, Frisson S, Jensen O. 2021. Neural evidence for lexical parafoveal processing. *Nat Commun 2021 121* **12**:1–9. doi:10.1038/s41467-021-25571-x

Pan Y, Popov T, Frisson S, Jensen O. 2023. Saccades are locked to the phase of alpha oscillations during natural reading. *PLOS Biol* **21**:e3001968.

Payne BR, Federmeier KD. 2017. Event-related brain potentials reveal age-related changes in parafoveal-foveal integration during sentence processing. *Neuropsychologia* **106**:358–370. doi:10.1016/j.neuropsychologia.2017.10.002

Payne BR, Stites MC, Federmeier KD. 2019. Event-related brain potentials reveal how multiple aspects of semantic processing unfold across parafoveal and foveal vision during sentence reading. *Psychophysiology* **56**:1–15. doi:10.1111/psyp.13432

Pollatsek A, Lesch M, Morris RK, Rayner K. 1992. Phonological Codes Are Used in Integrating Information Across Saccades in Word Identification and Reading. *J Exp Psychol Hum Percept Perform* **18**:148–162. doi:10.1037/0096-1523.18.1.148

Rayner K. 2009. Eye movements and attention in reading, scene perception, and visual search. *Q J Exp Psychol* **62**:1457–1506. doi:10.1080/17470210902816461

Rayner K. 1975. The perceptual span and peripheral cues in reading. *Cogn Psychol* **7**:65–81.

Rayner K, Balota DA, Pollatsek A. 1986. Against parafoveal semantic preprocessing during eye fixations in reading. *Artic Can J Psychol Rev Can Psychol* **40**:473–483. doi:10.1037/h0080111

Rayner K, Schotter ER. 2014. Semantic preview benefit in reading english: The effect of initial letter capitalization. *J Exp Psychol Hum Percept Perform* **40**:1617–1628. doi:10.1037/a0036763

Rayner K, Schotter ER, Drieghe D. 2014. Lack of semantic parafoveal preview benefit in reading revisited. *Psychon Bull Rev* **21**:1067–1072. doi:10.3758/s13423-014-0582-9

Rayner K, Sereno SC, Lesch MF, Pollatsek A. 1995. Phonological Codes Are Automatically Activated During Reading: Evidence From an Eye Movement Priming Paradigm. *Psychol Sci* **6**:26–32. doi:10.1111/J.1467-9280.1995.TB00300.X

Rayner K, Warren T, Juhasz BJ, Liversedge SP. 2004. The Effect of Plausibility on Eye Movements in Reading. doi:10.1037/0278-7393.30.6.1290

Reichle ED, Pollatsek A, Fisher DL, Rayner K. 1998. Toward a model of eye movement control in reading. *Psychol Rev* **105**:125–157. doi:10.1016/b978-0-444-70113-8.50043-6

Reichle ED, Pollatsek A, Rayner K. 2006. E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cogn Syst Res* **7**:4–22. doi:10.1016/j.cogsys.2005.07.002

Reichle ED, Rayner K, Pollatsek A. 2003. The E-Z reader model of eye-movement control in reading: Comparisons to other models. *Behav Brain Sci* **26**:445–476. doi:10.1017/S0140525X03000104

Reichle ED, Reingold EM. 2013. Neurophysiological constraints on the eye-mind link. *Front Hum Neurosci* **7**:361. doi:10.3389/fnhum.2013.00361

Reichle ED, Warren T, McConnell K. 2009. Using E-Z reader to model the effects of higher level language processing on eye movements during reading. *Psychon Bull Rev*. doi:10.3758/PBR.16.1.1

Schotter ER. 2018. Reading Ahead by Hedging Our Bets on Seeing the Future: Eye Tracking and Electrophysiology Evidence for Parafoveal Lexical Processing and Saccadic Control by Partial Word Recognition. *Psychol Learn Motiv - Adv Res Theory* **68**:263–298. doi:10.1016/BS.PLM.2018.08.011

Schotter ER. 2013. Synonyms provide semantic preview benefit in English. *J Mem Lang* **69**:619–633. doi:10.1016/j.jml.2013.09.002

Schotter ER, Angele B, Rayner K. 2012. Parafoveal processing in reading. *Attention, Perception, Psychophys* **74**:5–35. doi:10.3758/s13414-011-0219-2

Schotter ER, Jia A. 2016. Semantic and plausibility preview benefit effects in English: Evidence from eye movements. *J Exp Psychol Learn Mem Cogn* **42**:1839–1866. doi:10.1037/xlm0000281

Schotter ER, Lee M, Reiderman M, Rayner K. 2015. The effect of contextual constraint on parafoveal processing in reading. *J Mem Lang* **83**:118–139. doi:10.1016/j.jml.2015.04.005

Schotter ER, Milligan S, Estevez VM. 2023. Event-related potentials show that parafoveal vision is insufficient for semantic integration. *Psychophysiology* e14246. doi:10.1111/PSYP.14246

Snell J, Grainger J. 2019. Readers are parallel processors. *Trends Cogn Sci* **23**:537–546. doi:10.1016/j.tics.2019.04.006

Snell J, Meeter M, Grainger J. 2017. Evidence for simultaneous syntactic processing of multiple words during reading. *PLoS One* **12**:1–17. doi:10.1371/journal.pone.0173720

Snell J, van Leipsig S, Grainger J, Meeter M. 2018. OB1-reader: A model of word recognition and eye movements in text reading. *Psychol Rev* **125**:969–984. doi:10.1037/rev0000119

Stites MC, Payne BR, Federmeier KD. 2017. Getting ahead of yourself: Parafoveal word expectancy modulates the N400 during sentence reading. *Cogn Affect Behav Neurosci* **17**:475–490. doi:10.3758/s13415-016-0492-6

Team RC. 2013. R: A language and environment for statistical computing **201**.

Tsai JL, Kliegl R, Yan M. 2012. Parafoveal semantic information extraction in traditional Chinese reading. *Acta Psychol (Amst)* **141**:17–23. doi:10.1016/J.ACTPSY.2012.06.004

Underwood NR, McConkie GW. 1985. Perceptual Span for Letter Distinctions during Reading. *Read Res Q* **20**:153. doi:10.2307/747752

Veldre A, Andrews S. 2018. Parafoveal preview effects depend on both preview plausibility and target predictability. *Q J Exp Psychol* **71**:64–74. doi:10.1080/17470218.2016.1247894

Veldre A, Andrews S. 2017. Parafoveal preview benefit in sentence reading: Independent effects of plausibility and orthographic relatedness. *Psychon Bull Rev* **24**:519–528. doi:10.3758/s13423-016-1120-8

Veldre A, Andrews S. 2016a. Is Semantic Preview Benefit Due to Relatedness or Plausibility? *J Exp Psychol Hum Percept Perform* **42**:939–952. doi:10.1037/xhp0000200

Veldre A, Andrews S. 2016b. Semantic preview benefit in English: Individual differences in the extraction and use of parafoveal semantic information. *J Exp Psychol Learn Mem Cogn* **42**:837–854. doi:10.1037/xlm0000212

Vialatte FB, Maurice M, Dauwels J, Cichocki A. 2010. Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives. *Prog Neurobiol* **90**:418–438. doi:10.1016/J.PNEUROBIO.2009.11.005

Wen Y, Snell J, Grainger J. 2019. Parallel, cascaded, interactive processing of words during sentence reading. *Cognition* **189**:221–226. doi:10.1016/j.cognition.2019.04.013

White SJ. 2008. Eye movement control during reading: Effects of word frequency and orthographic familiarity. *J Exp Psychol Hum Percept Perform* **34**:205–223. doi:10.1037/0096-1523.34.1.205

Williams CC, Perea M, Pollatsek A, Rayner K. 2006. Previewing the neighborhood: The role of orthographic neighbors as parafoveal previews in reading. *J Exp Psychol Hum Percept Perform* **32**:1072–1082.

Wong R, Veldre A, Andrews S. 2022. Are There Independent Effects of Constraint and Predictability on Eye Movements During Reading? *J Exp Psychol Learn Mem Cogn*. doi:10.1037/XLM0001206

Wu M, Bosker HR, Riecke L, Nl W. 2023. Sentential contextual facilitation of auditory word processing builds up during sentence tracking. *J Cogn Neurosci*.

Yan M, Richter EM, Shu H, Kliegl R. 2009. Readers of Chinese extract semantic information from parafoveal words. *Psychon Bull Rev* **16**:561–566. doi:10.3758/PBR.16.3.561

Yan M, Zhou W, Shu H, Kliegl R. 2012. Lexical and sublexical semantic preview benefits in Chinese reading. *J Exp Psychol Learn Mem Cogn* **38**:1069–1075. doi:10.1037/A0026935

Zhigalov A, Duecker K, Jensen O. 2021. The visual cortex produces gamma band echo in response to broadband visual flicker. *PLOS Comput Biol* **17**:e1009046. doi:10.1371/JOURNAL.PCBI.1009046

Zhigalov A, Herring JD, Herpers J, Bergmann TO, Jensen O. 2019. Probing cortical excitability using rapid frequency tagging. *Neuroimage* **195**:59–66. doi:10.1016/j.neuroimage.2019.03.056

Zhigalov A, Jensen O. 2022. Travelling waves observed in MEG data can be explained by two discrete sources. *bioRxiv*.

Zhigalov A, Jensen O. 2020. Alpha oscillations do not implement gain control in early visual cortex but rather gating in parieto-occipital regions. *Hum Brain Mapp* **41**:5176–5186. doi:10.1002/hbm.25183

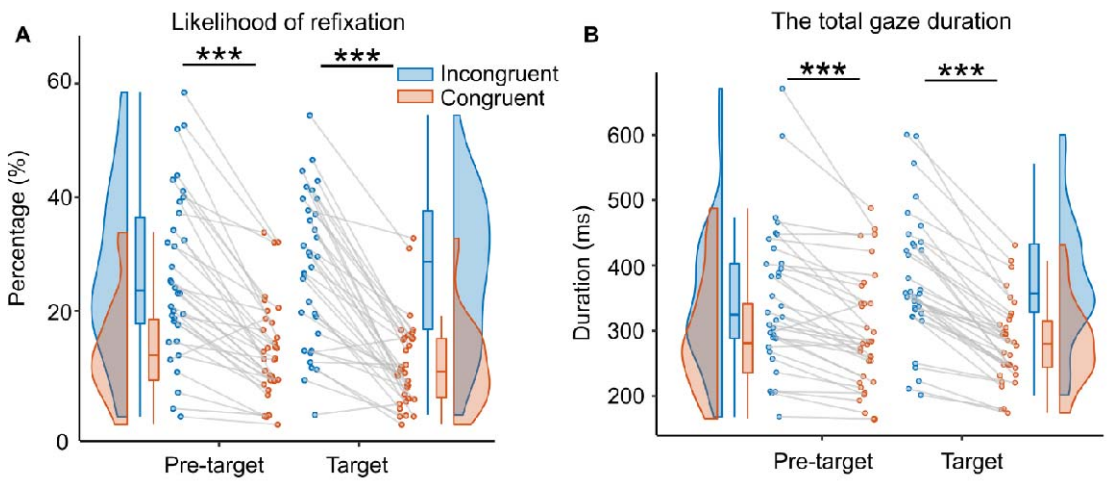
Zhou W, Kliegl R, Yan M. 2013. A validation of parafoveal semantic information extraction in reading Chinese. *J Res Read* **36**:S51–S63. doi:10.1111/J.1467-9817.2013.01556.X

Supplementary Materials for

Early parafoveal semantic integration in natural reading

Yali Pan*, Steven Frisson, Kara D. Federmeier, Ole Jensen
*Correspondence: Y.Pan.1@bham.ac.uk

This PDF file includes:
Supplementary Figure 1
Appendix



Supplementary Figure 1. The likelihood of refixation and total gaze duration of eye movement data. (**A**) The likelihood of refixation into a word was defined as the proportion of trials that have at least one regression from a later part of the sentence back to that word. We found that when the target words were contextually incongruent with the sentence compared with congruent, there was a significantly higher probability of regression into pre-target words ($t_{(33)} = 7.83$, $p = 5.04 \times 10^{-9}$, $d = 1.34$, two-sided pairwise *t*-test) and target words ($t_{(33)} = 9.13$, $p = 1.49 \times 10^{-10}$, $d = 1.57$, two-sided pairwise *t*-test). Each dot indicates one participant. (**B**) The total gaze duration was defined as the sum of all fixations on a given word, including those fixations during re-reading. Significantly longer total gaze durations were found for pre-target words ($t_{(33)} = 5.78$, $p = 1.86 \times 10^{-6}$, $d = .99$, two-sided pairwise *t*-test) and target words ($t_{(33)} = 10.55$, $p = 4.20 \times 10^{-12}$, $d = 1.81$, two-sided pairwise *t*-test) when the target words were incongruent with the context compared with congruent. ***$p < .001$; n.s., not statistically significant.

Appendix

Experimental sentence set

Here we share all 160 sentences embedded with congruent target words. For sentence set version A, we swapped the target words within each pair for sentences 1 to 80 and made them incongruent, while sentences 81 to 160 were kept congruent. For sentence set version B, target words in sentences 1 to 80 were kept the same but target words in sentences 81 to 160 were swapped within each pair to make them incongruent. The sequence of the sentences was shuffled to make sure that no more than 3 sentences in a row were in the same condition. For illustration, the target words are shown in italic type here, but they were in normal type in the experiment. For the 117 filler sentences, please see the Appendix in (White, 2008).

1. Last night, my lazy *brother/jacket* came to the party one minute before it was over.
2. Lily says this blue *jacket/brother* will be a big fashion trend this fall.

3. This area has been populated by many *hikers/coins* over the last year.
4. Little Jimmy picked up a box and put some *coins/hikers* inside of it.

5. Joey became an avid *student/ring* during his adolescence.
6. He could only afford a cheap *ring/student* without a diamond for his fiancée.

7. This morning the noisy *kids/ideas* played happily in the backyard.
8. My parents had no firm *ideas/kids* about what I should become.

9. The unfortunate pupil lost his beloved *pony/crisis* just before his birthday.
10. Experts say that the severe *crisis/pony* will cause oil prices to triple.

11. The construction of this ancient *castle/worker* cost a lot of money.
12. After the meeting, the anxious *worker/castle* sighed in the hallway.

13. Peter's love for this sporting *match/collar* inspired all his friends.
14. We could see from her torn *collar/match* that she had been in a fight.

15. With the help of his clever *friend/burger* Jack, he made the first pot of gold.
16. I always like to order a filling *burger/friend* from the local pub.

17. She looked at the tired *fireman/scan* with a satisfied smile.
18. He felt relieved after completing the complex *scan/fireman* within an hour.

19. Scientists found a steep *boulder/cousin* sitting in the middle of the canyon.
20. Last week his friendly *cousin/boulder* passed out for no apparent reason.

21. They asked the selfish *maid/roof* where her huge sums of money came from.
22. It took Tom a month to mend the broken *roof/maid* all by himself.

23. Under stress, the crafty *boss/rifles* promised customers a full refund.
24. The cowboys hung the stolen *rifles/boss* high up on the wall.

25. She submitted the crucial *file/queen* that can prove her innocence.
26. According to history books, the proud *queen/file* never accepted any criticism.

27. Ana complained that the tall *herbs/sport* behind the house had dried up.
28. He failed in his chosen *sport/herbs* with hopes of success fading with each effort.

29. She said that the corrupt *company/bush* offered high salaries to young graduates.
30. In the last few years, the thick *bush/company* died back dramatically.

31. It turned out that the last- minute *trip/tree* lasted for six hours.
32. Linda found that the slender *tree/trip* dead from a pest infestation.

33. Suddenly, the warm *coffee/flower* stained his brand new shirt.
34. Plenty of rain will make the vivid *flower/coffee* blossom well.

35. Jack became a humble *chef/vehicle* specializing in French cuisine.
36. The young man's shiny *vehicle/chef* vanished slowly out of sight.

37. To the north, the steep *hills/colonel* stretched for many miles.
38. Before sleeping, the nervous *colonel/hills* smoked a cigarette.

39. Decades ago, that algae- covered *pond/player* was enough to irrigate the crops.
40. He saw the smart *player/pond* throw the ball, causing chaos among the opposition.

41. In recent days, the cruel *murder/cream* has scared citizens from going out.
42. Mary told me that the light *cream/murder* was low in fat but hard to whip.

43. News said that the painful *disease/ball* would continue to affect many children.
44. The boy found his lost *ball/disease* under the tree and stopped crying at once.

45. Politicians hated the brief *report/drone* criticizing the government's incompetence.
46. My favourite gift is the shiny *drone/report* from my dad last year.

47. Every year, the sandy *shore/officer* attracts thousands of tourists.
48. After taking a deep breath, the junior *officer/shore* entered the room.

49. Last week, the caring *family/plaza* rescued a stray dog and kept it as a pet.
50. During the air raid, the spacious public *plaza/family* happened to be ruined.

51. With his sharp criticism, the young *actor/storm* annoyed his agent as usual.
52. Laura was told that the sudden *storm/actor* delayed the bus for two days.

53. Lily said that the vacant *cottage/picture* belonged to her grandparents.
54. In the small house, a comic *picture/cottage* adorned the reception room.

55. Facing the lion, the brave *hunter/engine* showed no fear.
56. Out of repair, the rattling *engine/hunter* was about to be scrapped.

57. Jack had to admit that this planned *visit/aunt* turned out to be embarrassing.
58. Tom admired the way his devoted *aunt/visit* always volunteers on weekends.

59. Rob felt that the brief *letter/clerk* from his wife expressed a hint of sadness.

60. Alone at home, the tired *clerk/letter* cooked a beef patty.

61. After the surgery, Rob's poor *health/dusk* left him barely able to get out of bed.
62. Sam's train arrived before *dusk/health* and we were able to give him a ride home.

63. She gave the dog a quick *bath/joke* after they came back from the outside.
64. Michael made a mean *joke/bath* about Boris Johnson's hair.

65. They didn't realize the harsh *impact/crown* that their products could have.
66. In the museum, we saw the golden *crown/impact* that belonged to the first king.

67. Bill is a superb *partner/night* because he is easy to get along with.
68. The explorer made his way through the gloomy *night/partner* with a small torch.

69. The TV show was an obvious *flop/lady* after the actress joined the cast.
70. On rainy days, the careful *lady/flop* reminded herself to go slowly.

71. Jane complained that her white *kitten/problem* hadn't come home for two days.
72. I guess no one can solve the hard *problem/kitten* without outside help.

73. The new event was such a huge *failure/suspect* that people kept talking about it.
74. Before committing the crime, the anxious *suspect/failure* drank a lot of alcohol.

75. Toby kept his money in a small *shed/deer* because he lived on a farm.
76. They noticed the young *deer/shed* eating acorns in the forest.

77. Amy wanted some more of the sliced *pear/canal* for afternoon snack.
78. Laura went down to the narrow *canal/pear* to watch the boats.

79. I wondered if the noisy *club/pain* would be a good place for the bachelorette party.
80. Tara always has an acute *pain/club* in her tooth after eating ice cream.

81. Before the war, the brave *general/potato* assembled an army.
82. She began to slice up a large *potato/general* for the dinner.

83. The child had a large *face/mist* with big, expressive eyes.
84. Last night, there was a dense *mist/face* when they left the cinema.

85. Marla enjoyed seeing the chubby *cats/court* playing with each other.
86. Jim entered the giant *court/cats* to try out for the basketball team.

87. They visited the antique *chapel/fans* before booking their wedding.
88. After the defeat, the crazy *fans/chapel* kept cursing and crying.

89. She approached the rusty *gate/toast* before realizing it was locked.
90. Many people like to eat crispy *toast/gate* with their morning coffee at breakfast.

91. They stepped into the messy *garage/hawk* that had high wooden shelves.
92. We watched the large hungry *hawk/garage* swoop down to get the poor chicken.

93. Alexandra used a short *hammer/museum* when she created the stone statue.
94. Ruth visited the public *museum/hammer* that she had read about all these years.

95. We'd better buy some tasty *chips/speech* before we watch the big game.
96. Historians believe the rousing *speech/speech* heralded the start of the revolution.

97. Under the tree, there is a little *hare/moon* running happily.
98. In the darkness, only the misty *moon/hare* lit up the street.

99. The prince inherited the supreme *power/sheep* from the late king.
100. Look over there, a fluffy *sheep/power* seems to be lost.

101. The man was a young *teacher/opinion* who always worked late into the night.
102. As for this scandal, Jo has a clear *opinion/teacher* but she won't say it.

103. They had no idea that the blue *liquid/justice* shrinks all woollen clothes.
104. The report was sent to the honest *justice/liquid* three days before the trial.

105. Every night, this tired *captain/bottle* drank wine before going to sleep.
106. The shopkeeper said the metal *bottle/captain* would sell well this year.

107. For the locals, the salt *lake/animal* triggered a political issue.
108. Near the small brook, a hungry *animal/lake* hunts quietly for hours.

109. Every night, this deep *secret/surgeon* makes the pianist toss and turn.
110. In the lab, a young *surgeon/secret* examined the victim's body.

111. Sadly, the lonely *poet/flour* died before he could finish his last poem.
112. Due to the moist weather, the wheat *flour/poet* became mouldy quickly.

113. Sue's colleagues say that her warm *heart/screen* makes everyone like her.
114. On the wall, the small green *screen/heart* shows the room temperature precisely.

115. To his surprise, the yummy *dish/nanny* was not expensive.
116. Eventually, the greedy *nanny/dish* disclosed all the details about this affair.

117. The sight of the cotton *factory/patient* was something to behold.
118. It was obvious that the weak *patient/factory* was getting weaker day by day.

119. In the past month alone, the gentle *scholar/pots* published five papers.
120. The filthy and rusty *pots/scholar* made the food taste terrible.

121. Just after dawn, an armed *ship/shirt* approached the pretty lagoon slowly.
122. At last, she found the wool *skirt/ship* hanging in the wardrobe.

123. Villagers said that the newly built *school/crowd* was well equipped.
124. In the downtown market, the agitated *crowd/school* began the parade.

125. Suzy really likes eating *sugar/music* because she wasn't allowed to eat it as a kid.
126. Ali said he really enjoyed modern *music/sugar* when he was at college.

127. Tina wants a spacious *yard/chief* because she likes to lie on the grass and read.
128. In an open field, the violent *chief/yard* executed prisoners with a gun.

129. After working overtime for a month, the wronged *manager/grass* wanted to jump ship.
130. In the Stone Age, the spiny *grass/manager* prevailed over the land.

131. The sparrow was being chased by some fluffy *hens/cups* under the hot sun.
132. When the ball was scored, they tapped their *cups/hens* to show their joy.

133. They recorded the details of the stolen *cars/legs* carefully on a spreadsheet.
134. The poor boy stood in the snow with bruised *legs/cars* and cried sadly.

135. Little Roy likes to play with the plastic *bricks/lawyer* at the Lego store.
136. It was said that the honest *lawyer/bricks* convened the committee meeting.

137. I learned about the muddy *trail/jury* through a friend on the last hike.
138. Just now, the calm *jury/trail* delivered a guilty verdict in this notorious case.

139. The holiday was neglected by this busy *parent/story* but her son was used to it.
140. This widely spread *story/parent* reflected the distortion of human nature.

141. We are meeting at the newly built *airport/editor* tonight for our trip to Europe.
142. Every day before leaving work, the tall *editor/airport* cleans her desk.

143. Nobody knew when the excited *puppy/area* urinated on the floor.
144. Everyone knows that entire *area/puppy* has restricted access.

145. The man's cunning *excuse/truck* relieved him of the fine.
146. Roy repaired the broken *truck/excuse* over the weekend.

147. Ana was glad that the gentle *nurse/meeting* said her little boy was out of danger.
148. In the company, the annual *meeting/nurse* marks the end of a year’s hard work.

149. He carefully placed the sharp *sword/desk* down after the fight.
150. She found an empty *desk/sword* where she could put her computer.

151. They danced a slow *tango/note* together after dinner.
152. Steph noticed a torn *note/tango* and looked for the other half.

153. Mindy’s dog has a strange *smell/tape* and likes to bark a lot.
154. Patty likes to cut some pink *tape/smell* to decorate her notebooks.

155. We could hear the angry *priest/card* shouting at the little girl.
156. David was happy to receive a nice *card/priest* from his daughter at Christmas.

157. She always meets the same happy *couple/hole* when she walks in the park.
158. The stray dog lives in a hidden *hole/couple* that protects it from the cold weather.

159. Jack failed to submit his concise *paper/baker* before the deadline.

160. I have heard that the young *baker/paper* makes the best baguettes in town.

References

White SJ. 2008. Eye movement control during reading: Effects of word frequency and orthographic familiarity. J Exp Psychol Hum Percept Perform 34:205–223. doi:10.1037/0096-1523.34.1.205