

ACCEPTED MANUSCRIPT

This article may not exactly replicate the authoritative document published in the APA journal, *Journal of Experimental Psychology: Human Perception and Performance* (doi: 10.1037/xhp0000200). It is not the copy of record.

Is semantic preview benefit due to relatedness or plausibility?

Aaron Veldre & Sally Andrews

University of Sydney

Short title: Plausibility preview benefit in English

Contact details

Aaron Veldre

School of Psychology, University of Sydney

Sydney, NSW 2006

Australia

aaron.veldre@sydney.edu.au

+61 2 9351 8990

Abstract

There is increasing evidence that skilled readers of English benefit from processing a parafoveal preview of a semantically related word. However, in previous investigations of semantic preview benefit using the gaze-contingent boundary paradigm the semantic relatedness between the preview and target has been confounded with the plausibility of the preview word in the sentence. In the present study, preview relatedness and plausibility were independently manipulated in neutral sentences read by a large sample of skilled adult readers. Participants were assessed on measures of reading and spelling ability to identify possible sources of individual differences in preview effects. The results showed that readers benefited from a preview of a plausible word, regardless of the semantic relatedness of the preview and the target. However, there was limited evidence of a semantic relatedness benefit when the plausibility of the preview was controlled. The plausibility preview benefit was strongest for low proficiency readers, suggesting that poorer readers were more likely to program a forward saccade based on information extracted from the preview. High proficiency readers showed equivalent disruption from all non-identical previews suggesting that they were more likely to suffer interference from the orthographic mismatch between preview and target.

Keywords

Reading, eye movements, semantic preview benefit, individual differences, plausibility

Word count

7222 (excluding Abstract, References, and Footnotes)

Skilled reading involves the precise coordination of oculomotor control, attention, and lexical processing. During fixations, readers rapidly retrieve the meanings of words in foveal vision and initiate processing of yet-to-be fixated words in the parafovea. However, the nature of the information extracted from parafoveal words remains a source of debate. Studies investigating parafoveal processing in reading typically employ the gaze-contingent *boundary paradigm* (Rayner, 1975) in which a target word is masked by a preview stimulus until the reader makes a saccade across an invisible boundary that is located at the end of the pretarget word. The well-established *preview benefit*, which refers to reduced first-pass reading times on the target after an orthographically/phonologically related preview relative to an unrelated preview, suggests that readers routinely extract and use (at least) sublexical information from the parafovea in reading (see Schotter, Angele, & Rayner, 2012, for a review).

There is now increasing evidence that readers of English also benefit from previews of *semantically* related words but only under certain conditions (e.g., Rayner & Schotter, 2014; Schotter, 2013; Schotter, Lee, Reiderman, & Rayner, 2015; Veldre & Andrews, in press; but see Rayner, Balota, & Pollatsek, 1986; Rayner, Schotter, & Drieghe, 2014). The relative elusiveness of semantic preview benefit in English contrasts with languages such as Chinese (Yan, Richter, Shu, & Kliegl, 2009; see also Tsai, Kliegl, & Yan, 2012; Yan, Zhou, Shu, & Kliegl, 2012), German (Hohenstein & Kliegl, 2014), and Korean (Kim, Radach, & Vorstius, 2012). Investigating the conditions required to observe the effect in English will therefore enhance our understanding of how parafoveal processing contributes to reading.

Preview benefit effects have traditionally been interpreted as reflecting the pre-activation of information from the parafovea and integration of this information with the target when it is fixated to facilitate its identification. For example, orthographic preview

benefit is attributed to pre-activation of orthographic features that the preview shares with the target. However, the mechanisms underlying *semantic* preview benefit remain unclear.

One possibility is that a semantically related preview word directly activates the target word. This could occur by a similar mechanism to that underlying semantic priming, e.g., spreading activation in semantic memory from the preview to words that share relevant semantic features (see Neely, 1991). Consistent with this interpretation, Schotter (2013) found that a synonym (e.g., *street* as a preview for *avenue*) but not a semantic associate (*suburb*) produced a benefit to first-pass reading relative to an unrelated word (*client*). This pattern of results suggests that the preview may activate closely related synonyms of the target but not more remotely associated words. The extent of semantic pre-activation may also be facilitated by top-down expectancies derived from the sentence context. Schotter et al. (2015) found that in a moderately constraining sentence, semantic associates produced a significant preview benefit on first-pass reading.

An alternative explanation of semantic preview benefit is that, rather than arising from spreading activation between the preview and target, it may actually reflect the degree of fit between the preview and the sentence context: facilitation for previews that are semantically plausible/acceptable in the sentence and/or a cost associated with the incompatibility of an implausible word. Current evidence does not rule out this explanation because, in all studies that have provided positive evidence for semantic preview benefit in English (Rayner & Schotter, 2014; Schotter, 2013; Schotter et al., 2015; Veldre & Andrews, in press), the semantically related word preview has differed from the unrelated baseline in both its relatedness to the target *and* in its acceptability in the sentence context. Specifically, in the studies cited above, the semantically related word was always an acceptable continuation of the pre-target sentence, while the unrelated word was usually an implausible word in the sentence.¹ Thus, rather than arising from pre-activation of the target, the

facilitation from a semantically related preview may reflect the extent to which the preview is a plausible continuation of the sentence.

Effects of preview plausibility have previously been observed in Chinese. Yang, Wang, Tong, and Rayner (2012) found that there was no semantic preview benefit when both the related and unrelated previews were implausible in the sentence context. However, there was a significant *plausibility* preview benefit because a preview of a word that was a plausible continuation of the pre-target sentence facilitated first-pass reading relative to an implausible preview. There was also an additional semantic benefit from a plausible related word relative to a plausible unrelated word, but this effect was restricted to single fixation duration. Taken together these results suggest that semantic preview benefit in Chinese depends, in part, on preview plausibility. Yang, Li, Wang, Slattery, and Rayner (2014) extended these findings by showing that an initially plausible preview (that was incongruous with the post-target sentence context) facilitated first-pass reading of the target. The authors argued that the plausibility effect was not due to integration of the plausible preview with the sentence context because the plausible preview did not produce more regressions back to the target than an implausible preview. The data therefore suggested that a plausible preview word directly facilitated processing of the *target* word.

It is unclear whether preview plausibility also plays a role in English. Assessing the plausibility of a word is usually assumed to occur at a postlexical stage of processing, reflecting the integration of semantic information with the context. Postlexical processing of parafoveal words may occur relatively infrequently in English because the orthographic depth of the writing system arguably makes parafoveal (lexical) processing more demanding (Laubrock, & Hohenstein, 2012). Thus there may be limited opportunities to observe effects of the plausibility of preview words in English. To explore this issue, Schotter (2013) conducted follow-up analyses to determine whether the synonym preview benefit observed in

her study was due to unrelated items that were semantically and/or syntactically anomalous. These analyses revealed that effects of preview anomaly were restricted to the late measures of go-past duration and regressions out of the target and did not affect early first-pass measures. Schotter claimed that the late effects of preview anomaly implicated postlexical integration processes, rather than word identification, and suggested that preview plausibility was not a major driver of semantic preview benefit in English.

However, effects of plausibility have been observed early in the time-course of reading in English. Several studies have shown that highly implausible words cause inflated fixation durations on first-pass reading (Rayner, Warren, Juhasz, & Liversedge, 2004; Warren & McConnell, 2007) and these effects have been reported from as early as the duration of the first fixation on the word (Matsuki et al., 2011; Staub, Rayner, Pollatsek, Hyönä, & Majewski, 2007). Critically, these early effects do not violate the assumption that plausibility only influences postlexical processing. Since lexical processing of an upcoming word is often initiated when it is in the parafovea, postlexical integration difficulty can be observed on the earliest fixation measures (Abbott & Staub, 2015).

A framework for explaining early semantic/plausibility effects is provided by Schotter, Reichle, and Rayner's (2014; see also Schotter et al., 2015) recent conceptualization of semantic preview effects within the E-Z Reader model of eye movements in reading (Reichle, Pollatsek, Fisher, & Rayner, 1998). Even though this serial attention model of eye movement control assumes that words are processed one at a time, it provides a detailed, computationally specified account of how parafoveal preview effects can arise early in the eye movement record. E-Z Reader's account of parafoveal processing hinges on the core assumption that the completion of an initial, cursory stage of lexical processing (L_1) initiates the programming of a saccade to the next word. During the approximately 150 ms before that saccade executes, processing of the fixated word continues to complete the second stage

(L₂) of lexical processing, at which point attention shifts to the next word even if the saccade to that word has not yet occurred. The saccadic program triggered by the completion of L₁ is initially labile (M₁), but can be re-programmed if L₁ on the next word completes before it reaches a non-labile state (M₂).

Within this model, preview effects are contingent on the amount of processing time available from when attention shifts to the preview word after completion of lexical access (L₂) for the pre-boundary word until the point at which new visual information becomes available after the execution of a saccade to the target word. It was initially assumed that the parafoveal processing arising from the attentional shift rarely allowed enough time to extract semantic information (see e.g., Hohenstein & Kliegl, 2014). However, Schotter et al.'s (2014) recent simulations showed that semantic preview effects can occur in cases where L₁ processing of the preview word in the parafovea completes too late to cancel the planned saccade to the target word, but early enough to trigger programming of a new saccade away from the target location before the target is actually presented. Specifically, if L₁ processing of the preview word completes at any point during non-labile saccade programming (M₂), during the saccade itself, or during the preattentive visual processing stage (V) of the target, a new saccade is programmed away from the target word based on the information extracted from the *preview* word, rather than the target. If this saccade is executed, it will produce an apparent preview benefit that has nothing to do with the preview-target relationship, because the saccade was programmed before the target word was presented.

This account implies, however, that semantic preview effects may be influenced by the contextual fit of the preview in the sentence. E-Z Reader 10 (Reichle, Warren, & McConnell, 2009) assumes that lexical retrieval of each word is followed by *integration* processes that assess whether “comprehension is proceeding without difficulty, making it unnecessary to interrupt the default forward movement of the eyes” (Reichle, Pollatsek, &

Rayner, 2012, p. 159). Integration failures lead to the cancellation of any planned saccades. These early, incremental integration processes provide a mechanism to explain plausibility preview effects. Plausible previews may be successfully integrated without delaying the planned saccade resulting in a preview benefit on first-pass reading. In such cases, the preview word is misidentified as the target, which may be detected during late integration processes leading to late regressions back to the target word. However, if integration failures for implausible previews are detected early enough, they will terminate the planned saccade leading to prolonged gaze duration or regressions back to the source of processing difficulty (Reichle et al., 2013). Plausibility preview effects on first-pass measures might, therefore, reflect a combination of benefits for plausible previews and costs for implausible previews.

The Present Study

The present study was designed to examine whether semantic preview benefit in English is due to semantic activation of the target word or the plausibility of the preview word in the sentence. In order to directly test this, we independently manipulated preview relatedness and plausibility, two factors that have been confounded in previous investigations of semantic preview benefit in English. This design allowed for a plausible related word to be compared to a plausible unrelated word to assess whether readers showed an additional benefit from shared semantic features between the preview and target when both words fit the sentence context.

We also examined whether preview plausibility effects are modulated by reading proficiency. Our recent investigations of individual differences among skilled readers have revealed that higher proficiency readers engage in deeper and more extensive parafoveal processing than poorer readers (Veldre & Andrews, 2014, 2015a, 2015b). These studies all assessed proficiency using measures of spelling ability in addition to passage comprehension tests to capture the precise orthographic knowledge that is central to Perfetti's (1992, 2007)

lexical quality hypothesis of reading skill. Of most direct relevance is recent evidence that reading and spelling ability independently modulate semantic preview benefit. Veldre and Andrews (in press) found that reading and spelling ability yielded opposite effects of semantically related previews on target gaze duration. Specifically, higher reading ability was associated with a larger semantic preview benefit suggesting that better readers were more likely to extract semantic information from the parafovea. In contrast, while higher spelling ability was associated with higher overall skipping rates, implying deeper parafoveal processing, better spellers showed less semantic preview benefit than poorer spellers. This pattern of results suggested that, for trials on which the target was fixated, better spellers were more likely to suffer interference from the competition between the orthography of the preview and target which counteracted the benefit of shared semantic features between the preview and target (Veldre & Andrews, in press). The present study investigated whether proficiency effects reflect differences in sensitivity to preview-target relatedness or to contextual plausibility of the preview, or whether both factors play a role.

METHOD

Participants

The final sample comprised 97 students from the University of Sydney (62 female; mean age 18.9 years) who received course credit for participation.² All had normal or corrected-to-normal vision and reported English as the first language they learned to read and write.

Materials and Design

Participants read 80 sentences (mean length: 11.8; range: 8-16 words) in which the preview of a critical target word (range 4-7 letters) was manipulated (see Figure 1 for an example and Appendix for complete list of materials). An invisible boundary was located at the end of the pre-target word (mean length: 5.3, range: 4-9 letters). Prior to crossing the

boundary the preview of the target was replaced with one of five preview words: identical; a plausible continuation of the sentence and semantically related to the target; implausible and semantically related; plausible and unrelated; or implausible and unrelated. The targets and plausible related words were high in semantic feature overlap and have previously been found to produce a semantic preview benefit relative to an unrelated word in an experiment that did not manipulate preview plausibility (Veldre & Andrews, in press). The plausible unrelated preview and the two implausible previews were always incompatible with the post-target text. All sentences appeared in all preview conditions across five counterbalanced lists. The stimulus characteristics and norming data are presented in Table 1.

--- INSERT FIGURE 1 ABOUT HERE ---

Stimulus norming. A separate group of 20 participants provided cloze norming data. They were given each sentence frame up to the pre-target word and asked to write down the word that was most likely to come next. The results of the cloze task confirmed that the preview words were low in predictability ($< 2\%$, on average).

Another group of 20 participants provided plausibility ratings on a 7-point scale for the sentence frames up to and including the preview word. The sentence fragments ending in the identical, plausible related, and plausible unrelated previews were rated as highly acceptable and did not differ significantly from one another (all $t_s < 1$). The fragments ending with the implausible related and implausible unrelated previews were rated as significantly lower in acceptability than each of the plausible previews (all $t_s > 20.9$, $p_s < .001$) but did not differ significantly from one another ($t = 1.41$, $p = .16$).

A further 18 participants provided semantic relatedness judgments on a 7-point scale, rating the similarity in meaning of the target and preview words. Both related previews were rated significantly closer in meaning to the target than the two unrelated previews (all $t_s > 21.1$, $p_s < .001$). While the plausible related preview was judged to be closer in meaning to

the target than the implausible related preview ($t = 8.04, p < .001$), the two unrelated previews were judged to be equally dissimilar in meaning to the target ($t = 1.81, p = .07$).

--- INSERT TABLE 1 ABOUT HERE ---

Measures of Written Language Proficiency

Participants completed the Nelson-Denny Reading Test (Brown, Fishco, & Hanna, 1993), which includes vocabulary and passage comprehension subsections. Participants also completed two measures of spelling ability: dictation and recognition. The *spelling dictation* test consisted of a list of 20 low frequency words selected from Burt and Tate (2002). Each word, and a sentence containing the word, was read aloud by the experimenter and the participant was instructed to write down the correct spelling of the word. The *spelling recognition* test comprised a list of 44 correctly spelled words intermixed with 44 incorrectly spelled items. Participants were given unlimited time in which to circle all incorrectly spelled items.

Standardized scores on the two spelling tests were highly correlated ($r = .80$) and were averaged to form a single measure of spelling ability. Reading and spelling ability were moderately correlated ($r = .58$). The first factor of a principal components analysis, which accounted for approximately 79% of the variance in scores, was used in the analyses to index overall proficiency.³

Apparatus

An EyeLink 1000 system was used to record participants' eye movements as they read sentences on a ViewSonic 225fb CRT monitor, which had a refresh rate of 150 Hz. The sentences occupied a single line and were presented in black monospaced font on a gray background. Viewing was binocular but fixation position was monitored from the right eye. Participants were seated 60 cm from the monitor and a chin and forehead rest was used to

minimize head movements. At this distance 2.5 characters subtended 1 degree of visual angle.

Procedure

Participants were instructed to read the sentences for meaning and that a comprehension question would follow some sentences. A three-point calibration procedure was followed by three practice trials and the 80 experimental trials presented in a random order. At the beginning of each trial, a fixation point appeared at the location of the first letter of the sentence. Once the participant made a stable fixation on this point, the sentence was displayed or a new calibration procedure was performed if necessary. Mean calibration error was less than 0.3 degrees of visual angle. The participant pressed a key when s/he finished reading the sentence. On all practice trials and approximately 25% of experimental trials, the sentence was followed by a three-option multiple-choice comprehension question that required a moderate understanding of the meaning of the sentence. The battery of proficiency measures was administered either before or after the eye movement task and the whole experiment lasted approximately 60 minutes.

RESULTS

Fixations below 80 ms that were within one letter space of an adjacent fixation were merged and remaining fixations below 80 ms or above 1000 ms were eliminated (4.5% of total fixations). Trials were eliminated if the participant made a blink immediately before or after fixating the target word (2.5% of trials) or the display change completed more than 10 ms into a fixation or was triggered by a saccade that landed to the left of the boundary (10.5% of trials). Target gaze durations above 1200 ms (8 trials) were also excluded. These exclusions left 6739 trials (86.6% of the data) available for analysis. Mean comprehension accuracy was high (94%; range 80-100%), indicating that participants read for meaning.

The following first-pass reading measures were analyzed: *first fixation duration* (the duration of the first fixation on the target word regardless of the number of first-pass fixations it receives), *single fixation duration* (the fixation duration in cases when only one first-pass fixation is made on the target word), *gaze duration* (the sum of all first-pass fixations on the target word). We also analyzed two late measures of reading: *go-past duration* (the sum of all fixations from the first fixation on the target word until a word to the right is fixated, i.e., this measure includes fixations on the target and any subsequent fixations on words earlier in the sentence), and *total duration* (the sum of all fixations on the target word including regressions from later in the sentence). Measures of the probability of the reader making a *first-pass fixation* on the target; the probability of *regressions out* of the target to words earlier in the sentence; and *regressions in* to the target from words later in the sentence were also analysed. Means for the five preview conditions on each of these measures are presented in Table 2.

--- INSERT TABLE 2 ABOUT HERE ---

The duration measures were analyzed by linear mixed-effects models (LMM) and fixation probability measures were analyzed with generalized LMM (GLMM) using the *lme4* package (Version 1.1-7; Bates, Maechler, Bolker, & Walker, 2014) in R (Version 3.2.0; R Core Team, 2015). The models included subject and item random intercepts and preview condition random slopes.⁴ The following planned contrasts were tested: (i) *identical* preview benefit: the average of the non-identical previews vs. identical preview; (ii) a *plausibility* effect: the average of the two implausible conditions vs. the average of the two plausible (non-identical) conditions; (iii) a *relatedness* effect: the average of the two unrelated conditions vs. the average of the two related conditions; and (iv) the *plausibility* × *relatedness* interaction: the difference in the effect of relatedness for implausible and plausible previews. This contrast coding scheme was implemented by transposing the inverse

of a matrix of contrast coefficients (see Appendix). The LMM intercepts, therefore, represent the grand mean of each dependent measure. Proficiency was included in the models as a continuous, mean-centered predictor. The models also included interactions between proficiency and the preview contrasts. Estimates 1.96 times larger than their standard errors were interpreted as significant at the .05 alpha level because, given the number of observations, the t statistic in LMMs effectively corresponds to the z statistic. The (G)LMM estimates for coefficients, standard errors, and t/z values for the fixed effects are reported in Tables 3 and 4. The random effects for each of the models are included in the Appendix. Figures were constructed using *ggplot2* (Wickham, 2009). We first summarize the preview effects in the average data and then consider the effects of proficiency on preview benefit.

--- INSERT TABLE 3 & 4 ABOUT HERE ---

Average Preview Effects

There was no difference in first-pass fixation probability across the preview conditions [all $|z|s < 1.6$].⁵ On target fixation duration, there was a significant *identical* preview benefit across all reading measures [all $|t|s > 5.6$]. Readers were also less likely to make a regression from and to the target after an identical preview [both $|z|s > 5.3$].

The effect of preview *plausibility* was significant for all duration measures [all $|t|s > 2.0$]. Readers were also more likely to regress from the target after an implausible preview [$z = 2.8$]. However, regressions into the target were more likely after a plausible preview [$z = -3.1$]. These results imply that implausible previews resulted in immediate integration difficulties, leading to inflated fixations on the target and regressions to earlier words in the sentence. In contrast, non-identical but locally plausible previews resulted in later integration difficulty that produced regressions back to the target from later words in the sentence. The late cost from plausible previews contrasts with the results of Yang et al. (2014) who found no effect of preview plausibility on late measures in Chinese. The present data therefore

suggest that the plausible preview was encoded on first-pass reading and integrated with the sentence context.

In contrast to the robust effects of plausibility, the effect of preview relatedness did not approach significance on first-pass reading [all $|t|s < 1.3$]. There was also no significant effect of preview relatedness on regressions out of the target [$z = 1.5$] or into the target word [$z < 1$]. However, total duration showed a significant relatedness effect [$t = 2.2$], which was qualified by a significant plausibility $\times$ relatedness interaction [$t = -2.1$]. Follow-up comparisons, separately for the plausible and implausible previews, showed that there was no effect of relatedness for implausible previews [$|t| < 1$] but there was a significant effect of relatedness for plausible previews [$b = 24.79$, $SE = 7.96$, $t = 3.11$] because regressions back to the target were more likely after a plausible unrelated word. This late effect suggests that readers initially integrated plausible preview words with the sentence but plausible *unrelated* previews, which were always incongruous with the post-target text, were more likely to be detected during late integration processes and lead to re-reading of the target word. The plausibility $\times$ relatedness interaction did not approach significance on any other measure [all $|t|s < 1.6$ and all $|z|s < 1$].

Individual Differences

Higher overall proficiency was associated with lower first-pass fixation probability, i.e. more skipping [$z = -4.5$], but this did not differ across preview conditions [all $|z|s < 1$]. For trials on which the reader made a first-pass fixation on the target, higher proficiency was associated with shorter fixation durations [all $|t|s > 3.5$] and a lower likelihood of regressions both out of the target [$z = -2.6$] and into the target [$z = -3.0$].

--- INSERT FIGURE 2 ABOUT HERE ---

Proficiency did not modulate *identical* preview benefit on any measure [all $|t|s$ and all $|z|s < 1$]. However, proficiency significantly modulated the *plausibility* preview effect on all

first-pass measures [all $|t|s > 2.1$]. To determine the basis of these effects, Figure 2 displays the partial effects of overall proficiency on gaze duration for each preview condition. These reveal that the plausibility effect was largest for lower proficiency readers and that high proficiency readers showed an identity preview benefit but no effect of preview plausibility for. To confirm that the interaction between proficiency and plausibility was not restricted to the model estimates and could also be observed in the raw data, Figure 3 presents the mean gaze duration for the plausible and implausible conditions separately for low and high proficiency readers, based on a median split of the continuous proficiency measure.

--- INSERT FIGURE 3 ABOUT HERE ---

Proficiency and plausibility did not yield any significant interactions on late reading measures [go-past duration: $t = -1.2$; total duration: $t = -1.3$] or regressions [both $zs < 1$]. There were also no significant interactions of proficiency with relatedness [all $|t|s < 1.7$ and all $|z|s < 1.9$] or the plausibility $\times$ relatedness interaction [all $|t|s < 1.6$ and all $|z|s < 1.5$] effect on any measure.

Supplementary Analyses

If semantic preview benefit depends on the reader integrating the semantic features of the preview and target word, it may be diluted by including trials in which the reader regresses out of the target word to earlier words in the sentence, and then refixates the target. To assess whether such early regressions were obscuring effects of preview-target relatedness on first-pass reading, we conducted an additional set of analyses based only on the subset of trials in which no regression was made out of the target word on first-pass reading (84.8% of total trials). These analyses revealed a significant effect of relatedness on single fixation duration [$b = 7.61$, $SE = 3.48$, $t = 2.18$] but not on first fixation duration [$b = 3.33$, $SE = 2.84$, $t = 1.17$] or gaze duration [$b = 6.19$, $SE = 3.77$, $t = 1.64$]. There were no other differences from the pattern of significant effects observed in the unrestricted analyses. Thus there was

limited evidence that readers benefited from the shared semantic features between the preview and target, but only for trials in which the reader made a single fixation during first-pass reading and did not regress from the target. Although caution is required when interpreting a small, isolated effect, this finding is consistent with Yang et al.'s (2012) Chinese data in which semantic preview benefit was restricted to single fixation duration when preview plausibility was controlled.

DISCUSSION

The aim of the present study was to provide insight into the source of semantic preview benefit in English by independently manipulating the semantic relationship between the preview and target and the contextual plausibility of the preview word in unconstrained sentences. We also investigated the role of individual differences to add to previous evidence indicating that reading proficiency modulates parafoveal processing. The average data showed that a locally plausible preview provided a benefit to first-pass reading relative to an implausible preview. However, readers obtained no additional benefit from a semantically related preview over and above an unrelated preview when both were equally plausible. Thus, when the plausibility of the preview was controlled, there was no evidence that readers (even highly proficient readers) benefited from the semantic relationship between the preview and target.

It is important to note that the apparent lack of semantic facilitation cannot be attributed to a weak manipulation of semantic relatedness. The plausible related words were selected because they shared semantic features with the target and were rated as highly similar in meaning to the target word. We have previously found that the same items produced a significant semantic preview benefit on first-pass reading relative to an unrelated word when preview plausibility was uncontrolled (Veldre & Andrews, in press). The lack of semantic preview benefit in the present study can also not be explained by low power.

Because of our focus on individual differences in reading proficiency, we tested a much larger sample of participants than is typical for eye movement studies. The present study was therefore well-equipped to detect even a small effect. However, despite the large sample and robust effects of preview plausibility, the only evidence of a semantic relatedness effect was an 8 ms benefit on single fixation duration that was only significant when trials on which the reader regressed out of the target were excluded; it did not approach significance in the full dataset.

Although there was only limited evidence that readers benefited from the semantic relationship between the preview and target, the significant benefit from plausible previews observed in the average data implies that the majority of readers extracted semantic information from the parafoveal preview before fixating the target word. This plausibility preview benefit suggests that, rather than affecting target processing, the semantic information activated by the preview was at least partially integrated with the *sentence context*. First-pass fixations on the target were shorter following a preview that was an acceptable continuation of the sentence and longer after a preview that was incongruous with the developing sentence representation. This occurred independently of the semantic relationship between the preview and target. Somewhat surprisingly, the plausibility preview benefit was restricted to low proficiency readers, suggesting that their eye movements during first-pass reading were more likely to reflect properties of the preview than the target. Before discussing this counter-intuitive relationship between proficiency and semantic preview benefits, we will first consider whether the plausibility effects observed in the average data are compatible with E-Z Reader's account of preview benefit that was outlined in the Introduction.

Plausibility Preview Effects in E-Z Reader

A key assumption of E-Z Reader's account is that oculomotor decisions based on the completion of the cursory L_1 stage of processing are "dumb": saccades are programmed when the evidence extracted during this early stage of processing predicts that lexical retrieval for the currently attended word is imminent (Reichle et al., 2012). Essentially, the reading system makes a "hedged bet" (Schotter et al., 2015) that the correct meaning will be successfully retrieved, and initiates planning of the next saccade, before the information required to judge the word's semantic and/or syntactic acceptability in the sentence is available. This is supported by evidence that word skipping is determined by the ease of lexical processing and is independent of semantic and/or syntactic fit (e.g., Abbott, Angele, Ahn, & Rayner, 2015; Angele, Laishley, Rayner, & Liversedge, 2014; Angele & Rayner, 2013). The skipping rates in the present study are also consistent with this claim because readers were equally likely to skip implausible words as they were plausible words.

Even though L_1 processing of the preview may rarely complete quickly enough to yield skipping of the target, it may still complete well before the reader begins to receive visual input from the target. Preview processing continues during the 150 ms required to implement and execute the planned saccade and the 50 ms from target fixation onset until the updated visual information from the target word enters the reading system (E-Z Reader's preattentive visual processing stage: V). This may be sufficient for L_1 processing of the preview word to complete, leading to initiation of a new saccadic program away from the location of the target word.⁶ Schotter et al. (2015) suggested that, in these conditions, a display change from a semantically related preview to the target word may not even register with the reading system until the eyes have moved on, particularly when the preview is a synonym of the target and therefore completely congruent with information from the preceding text. The eye movement record will then show a 'semantic preview benefit' but it

arises from the compatibility of the coarse information from the preview with previous context, rather than because of the relationship between the preview and the target.

Assuming that saccades away from the target word are programmed on the basis of processing of the *preview* explains why our data showed no effect of the relationship between the preview and the (not yet presented) target. The duration of L₁ processing for these preview words, which were matched on frequency and predictability should have been equivalent, on average. However, the fact that there were significant differences between plausible and implausible previews in the average data shows that semantic information was extracted from the preview and that eye movements were affected by the compatibility of that information with the sentence context.

Within the framework of E-Z Reader, this may reflect an increased likelihood that a planned saccade away from the target word is canceled, if the integration of an implausible preview word stalls during the labile stage (M₁) of saccadic programming. As outlined in the Introduction, E-Z Reader 10 introduced a mechanism for accommodating effects of postlexical processing by assuming that the completion of lexical processing for each word is followed by an integration process that makes a coarse assessment of whether it fits with the incrementally constructed representation of sentence meaning (Reichle et al., 2013). Delays or failure in integration lead to the cancellation of all planned saccades, reflected in “a pause and/or movement of the eyes and attention back to the source of processing difficulty” (Reichle et al., 2013, p. 124).

The present data are consistent with these predictions. Readers were more likely to regress out of the target to earlier words in the sentence after an implausible preview, consistent with the view that contextual implausibility often led to outright integration failures, reflected in immediate regressions back to earlier words in the sentence. However, there was still a significant effect of preview plausibility on target fixation duration for the

subset of trials without regressions, suggesting that integration difficulties were sometimes resolved without regressing, reflected in inflated fixation durations on the target. The single significant effect of relatedness on single fixation duration trials excluding regressions may indicate that previews that were both semantically related and plausible were less likely to fail integration and trigger regressions, resulting in the small benefit of preview relatedness observed on these trials. The late costs for locally plausible completions, which only became implausible later in the sentence, are also consistent with the assumption that these words have been integrated into the developing sentence representation.

It is important to point out that, although we have interpreted these data within the framework of E-Z Reader, early semantic effects are also potentially compatible with the assumptions of processing gradient models, such as SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005) in which multiple words are processed in parallel. Unlike E-Z Reader, however, SWIFT presently has no mechanism for postlexical processing that could account for the early effects of plausibility observed in the present data. There are also other aspects of the data that would appear to be inconsistent with a parallel lexical processing assumption, such as the lack of parafoveal-on-foveal effects. Furthermore, no model currently accounts for individual differences among skilled readers (but see Mancheva et al., 2015; Reichle et al., 2013 for recent attempts to account for individual differences among developing readers in E-Z Reader). The individual difference results in the present study provide further insight into the source of the early plausibility effects.

Individual Differences in Plausibility Preview Effects

Lower proficiency readers showed the largest difference between the implausible and plausible preview conditions on first-pass fixations. Within the E-Z Reader framework, this suggests that low proficiency readers were more likely to execute a saccade on the basis of L_1 processing of the preview, and perhaps also more likely to subsequently cancel this saccade,

when they encountered difficulty in integrating an implausible word with the previous context. Evidence for such early integration difficulty was also provided by the higher rate of regressions out of the target word to earlier words in the sentence among lower proficiency readers. In contrast, high proficiency readers showed no effect of preview plausibility. Importantly, however, the lack of semantic/plausibility effect for high proficiency readers cannot be interpreted as indicating that they were less likely to process parafoveal words to the semantic level. High proficiency readers had higher skipping rates overall, implying that they completed L₁ processing of parafoveal words earlier than lower proficiency readers, on average. This means that, on (the majority of) trials in which L₁ processing did not complete quickly enough to yield skipping of the target, high proficiency readers would be *more* likely than lower proficiency readers to complete L₁ processing of the parafoveal preview before the display change occurred, triggering programming of a saccade. However, they did not show either the preview benefit on gaze duration expected if that saccade was executed, or the integration costs for implausible previews expected if they completed lexical retrieval and post-lexical processing of the preview before the target was presented. Higher proficiency readers were less likely to regress from the target word to earlier words in the sentence, and less likely to regress back to the target word, suggesting that they did not misidentify the target as the preview. In combination, these findings suggest that higher proficiency readers were more likely to have resolved any processing difficulties arising from the preview before moving their eyes on.

Rather than reflecting differences in the likelihood of extracting parafoveal semantic information, the present findings suggest that the major source of individual differences in semantic preview benefit lies in how quickly and effectively readers recover from processing a contextually implausible preview word. In combination with previous evidence that highly proficient readers process parafoveal orthographic information more quickly and deeply than

average readers (Veldre & Andrews, 2014, 2015a, 2015b), the reduced effect of preview plausibility among high proficiency readers implies that they were quicker and more effective at using the updated orthographic information from the target, when it became available, to terminate any saccades programmed on the basis of preview information. The fact that the highest proficiency readers showed a strong effect of preview identity, but no influence of either plausibility or relatedness suggests that this decision is based on orthographic discrepancies between the preview and target. All non-identical previews were equally orthographically different from the target, so the costs to first-pass fixations associated with identifying the discrepancy and canceling planned saccades was equivalent, and only an identical preview effect was observed.

In contrast, low proficiency readers were slower to update their representation of the target word after crossing the boundary so their first-pass saccade planning was more likely to reflect properties of the preview. This was initially beneficial for plausible previews that were as easily integrated with the sentence context as the target word. However, preview words that were implausible caused interference that resulted in inflated fixations on the target word due to postlexical integration difficulty. These preview effects may have been enhanced by increased reliance on sentence context among lower proficiency readers (Andrews & Bond, 2009; Ashby, Rayner, & Clifton, 2005). Lower proficiency readers' increased sensitivity to preview plausibility may also reflect imprecise encoding of orthographic features of the preview, perhaps reflecting lower quality lexical representations. These readers showed plausibility effects on early integration processes and late integration effects for locally plausible words that were revealed to be implausible by later context, demonstrating that they successfully retrieved semantic information from the preview and integrated it with the evolving sentence representation. However, they appeared to be insensitive to the orthographic discrepancy between the preview and target suggesting that, in

addition to less efficient processing/updating of the new target information, they may either extract or retrieve less detailed orthographic information about the preview.

This conclusion that orthographic precision modulates the manifestation of semantic preview effects is consistent with our recent evidence (Veldre & Andrews, in press) that the preview benefit from a word that was both plausible and semantically related to the target was attenuated among readers with high spelling ability.⁷ This reduced preview effect for good spellers could not be attributed to a lack of semantic pre-processing because high spelling ability was associated with higher skipping rates and a greater sensitivity to semantic relatedness on regressions out of the target. Because the effect was specific to spelling ability and was independent of overall reading ability, it was argued to reflect competition between the orthographic features of the preview and the *target* that eliminated the benefit of the shared semantic features with the related preview. The fact that this effect was restricted to good spellers suggests that readers with precise lexical knowledge were more likely to integrate the orthographic information from the preview and target during first-pass reading. The present findings converge with these findings by showing that readers who are high in both reading and spelling ability do not show the preview plausibility effects shown by lower proficiency readers, suggesting that they are more likely to ‘discard’ information from the preview based on orthographic discrepancies between the preview and target (Schotter, 2013).

Although the results of the present study suggest that preview-target relatedness has a limited role in semantic preview effects, it is important to note that the relative contribution of plausibility and relatedness may depend on sentence constraint. Schotter et al. (2015) found that semantic associates, which were often implausible continuations of the sentence, produced a preview benefit in moderately constrained sentences but not in neutral sentences. The preview benefit from a synonym was also larger in constrained sentences than in neutral

sentences. The authors attributed the differences to the pre-activation of semantically related words that are consistent with the sentence context. This semantic activation appeared to be graded, leading to a small benefit from a related (but often implausible) word that was counteracted by later re-reading resulting from integration failure. The present study shows that preview benefits also depend on preview plausibility, and that these effects can outweigh effects of preview-target relatedness, at least in unconstrained sentences. Nevertheless, as our supplementary analysis showed, there may be conditions under which it is possible to observe effects of semantic relatedness over and above the effect of plausibility.

In conclusion, the results of the present study suggest that the source of semantic preview benefit in neutral sentences in English may lie in the extent to which preview information is integrated with the sentence context. There was little evidence, even among highly proficient readers, that semantic relatedness between the preview and target benefited first-pass reading. In contrast, the present study provides clear evidence of a *plausibility preview benefit* in English. Low proficiency readers showed a benefit to first-pass reading from a plausible preview word over an implausible preview, regardless of the semantic relatedness to the target word. This effect was not evident among highly proficient readers, suggesting that they were more sensitive to orthographic discrepancies between the preview and target word. These results suggest that the relative elusiveness of semantic preview benefit in English compared to other languages is not because skilled readers fail to activate semantic attributes of parafoveal words, but rather because mismatches between the preview and either the sentence context, or the target word, often disrupt processing early enough to affect first-pass reading.

ACKNOWLEDGEMENTS

Portions of these data were presented at the 18th European Conference on Eye Movements, Vienna, Austria, 2015. We thank Erik Reichle and two anonymous reviewers for their helpful comments on an earlier version of this article.

FOOTNOTES

¹ Some of Schotter's (2013) semantic associates were also implausible continuations of the sentence: 17% were semantic anomalies and 13% were syntactic anomalies. It is therefore unclear whether the observed difference between the semantic associate and synonym conditions was due to differences in semantic overlap with the target or differences in contextual plausibility.

² One participant's data was excluded from the analyses because their comprehension accuracy in the eye-tracking task (70%) was more than three standard deviations below the mean.

³ Inspection of models including the reading and spelling scores as separate, continuous predictors indicated that collinearity was obscuring independent effects of the two individual differences measures. Specifically, there were no significant main effects of either reading or spelling ability when both predictors were included in the models, but both variables yielded significant main effects when only one of the predictors was included in the models. We therefore report analyses including the first principal component, which captured their combined influence.

⁴ To reduce model complexity, the correlations between the random effects were set to zero. Some models failed to converge with subject and item random slopes for the preview condition. Item random slopes were removed from these models (see Appendix).

⁵ There was also no evidence of any effects of preview type on fixation durations on the pretarget word, i.e. no parafoveal-on-foveal effects [all t s < 1.1].

⁶ E-Z Reader simulations reported by Schotter et al. (2014) found that this sequence of events occurred approximately 8% of the time, suggesting that effects of preview plausibility should

be observed relatively rarely. Further research will be necessary to determine whether this estimate is adequate to account for the robust effects of plausibility obtained in the present data.

⁷ There was reason to believe that the results in the present study were also mainly attributable to differences in spelling ability, rather than reading ability. As noted in the Results, the collinearity between reading and spelling ability in the present sample precluded clear separation of their effects. However, models that included only spelling ability produced the identical pattern of significant interactions as the results of the reported models with the combined proficiency measure while models that included only reading ability did not produce significant interactions with plausibility or relatedness.

REFERENCES

- Abbott, M. J., Angele, B., Ahn, Y. D., & Rayner, K. (2015). Skipping syntactically illegal *the* previews: The role of predictability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Abbott, M. J., & Staub, A. (2015). The effect of plausibility on eye movements in reading: Testing E-Z Reader's null predictions. *Journal of Memory and Language*.
- Angele, B., Laishley, A. E., Rayner, K., & Liversedge, S. P. (2014). The effect of high- and low-frequency previews and sentential fit on word skipping during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 1181-1203.
- Angele, B., & Rayner, K. (2013). Processing *the* in the parafovea: Are articles skipped automatically? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 649-662.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4 (Version R package version 1.1-6).
- Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *Nelson-Denny Reading Test*. Rolling Meadows, IL: Riverside Publishing Company.
- Burt, J. S., & Tate, H. (2002). Does a reading lexicon provide orthographic representations for spelling? *Journal of Memory and Language*, 46, 518-543.
- Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, 112, 777-813.
- Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 166-190.
- Hohenstein, S., & Kliegl, R. (2015). remef: Remove Partial Effects. R package version 1.0.6.9000. <https://github.com/hohenstein/remef/>

- Kim, Y.-S., Radach, R., & Vorstius, C. (2012). Eye movements and parafoveal processing during reading in Korean. *Reading and Writing*, 25, 1053-1078.
- Laubrock, J., & Hohenstein, S. (2012). Orthographic consistency and parafoveal preview benefit: A resource-sharing account of language differences in processing of phonological and semantic codes. *Behavioral and Brain Sciences*, 35, 292-293.
- Mancheva, L., Reichle, E. D., Lemaire, B., Valdois, S., Ecalle, J., & Guérin-Dugué, A. (2015). An analysis of reading skill development using E-Z Reader. *Journal of Cognitive Psychology*, 27, 657-676.
- Matsuki, K., Chow, T., Hare, M., Elman, J. L., Scheepers, C., & McRae, K. (2011). Event-based plausibility immediately influences on-line language comprehension. *Journal of Experimental Psychology: Learning Memory and Cognition*, 37, 913-934.
- Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264-336). Hillsdale, NJ: Lawrence Erlbaum.
- Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri & R. Treiman (Eds.), *Reading Acquisition* (pp. 145-174). Hillsdale, NJ: Lawrence Erlbaum.
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11, 357-383.
- R Core Team (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 7, 65-81.

- Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Experimental Psychology*, 40, 473-483.
- Rayner, K., & Schotter, E. R. (2014). Semantic preview benefit in reading English: The effect of initial letter capitalization. *Journal of Experimental Psychology: Human Perception and Performance*, 40, 1617-1628.
- Rayner, K., Schotter, E. R., & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit in reading revisited. *Psychonomic Bulletin & Review*, 21, 1067-1072.
- Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 1290-1301.
- Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review*, 33, 110-149.
- Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, 105, 125-157.
- Reichle, E. D., Pollatsek, A., & Rayner, K. (2012). Using E-Z Reader to simulate eye movements in nonreading tasks: A unified framework for understanding the eye-mind link. *Psychological Review*, 119, 155-185.
- Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin and Review*, 16, 1-21.
- Schotter, E. R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language*, 69, 619-633.

- Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception and Psychophysics*, 74, 5-35.
- Schotter, E. R., Lee, M., Reiderman, M., & Rayner, K. (2015). The effect of contextual constraint on parafoveal processing in reading. *Journal of Memory and Language*, 83, 118-139.
- Schotter, E. R., Reichle, E. D., & Rayner, K. (2014). Rethinking parafoveal processing in reading: Serial-attention models can explain semantic preview benefit and N+ 2 preview effects. *Visual Cognition*, 22, 309-333.
- Staub, A., Rayner, K., Pollatsek, A., Hyönä, J., & Majewski, H. (2007). The time course of plausibility effects on eye movements in reading: Evidence from noun-noun compounds. *Journal of Experimental Psychology: Learning Memory and Cognition*, 33, 1162-1169.
- Tsai, J.-L., Kliegl, R., & Yan, M. (2012). Parafoveal semantic information extraction in traditional Chinese reading. *Acta Psychologica*, 141, 17-23.
- Veldre, A., & Andrews, S. (2014). Lexical quality and eye movements: Individual differences in the perceptual span of skilled adult readers. *The Quarterly Journal of Experimental Psychology*, 67, 703-727.
- Veldre, A., & Andrews, S. (2015a). Parafoveal lexical activation depends on skilled reading proficiency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41, 586-595.
- Veldre, A., & Andrews, S. (2015b). Parafoveal preview benefit is modulated by the precision of skilled readers' lexical representations. *Journal of Experimental Psychology: Human Perception and Performance*, 41, 219-232.

- Veldre, A., & Andrews, S. (in press). Semantic preview benefit in English: Individual differences in the extraction and use of parafoveal semantic information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review*, 14, 770-775.
- Wickham, H. (2009). ggplot2: Elegant graphics for data analysis. New York: Springer.
- Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Chinese readers extract semantic information from parafoveal words during reading. *Psychonomic Bulletin & Review*, 16, 561-566.
- Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2012). Lexical and sublexical semantic preview benefits in Chinese reading. *Journal of Experimental Psychology: Learning Memory and Cognition*, 38, 1069-1075.
- Yang, J., Li, N., Wang, S., Slattery, T. J., & Rayner, K. (2014). Encoding the target or the plausible preview word? The nature of the plausibility preview benefit in reading Chinese. *Visual Cognition*, 22, 193-213.
- Yang, J., Wang, S., Tong, X., & Rayner, K. (2012). Semantic and plausibility effects on preview benefit during eye fixations in Chinese reading. *Reading & Writing*, 25, 1031-1052.

FIGURE CAPTIONS

Figure 1. Examples of the preview conditions used in the experiment: (a) identical, (b) plausible and semantically related, (c) implausible and semantically related, (d) plausible and unrelated, (e) implausible and unrelated. The invisible boundary is represented by the dashed line. In all conditions, the identical target word was displayed when the reader's eye crossed the boundary.

Figure 2. Partial effect of the continuous reading proficiency variable on gaze duration on the target word for each of the preview conditions in the experiment. Lines are regression estimates over data points from which subject and item random variance has been removed with the *remef* function (Hohenstein & Kliegl, 2015) and error bands represent 90% confidence intervals. Solid lines represent related previews; dotted lines are unrelated previews. From the top, the blue lines represent implausible previews, the gold lines are plausible previews, and the red line is the identical condition. Please see online version of article for the color version of this figure.

Figure 3. Mean gaze duration (aggregated by subjects) for the plausible (average of the plausible related and unrelated previews) and implausible conditions (average of the implausible related and unrelated previews) based on a median split of reading proficiency. Error bars represent 95% confidence intervals.

APPENDIX A

The sentence materials used in the experiment are listed below. The five preview conditions appear in parentheses in the following order: identical, plausible/related, implausible/related, plausible/unrelated, and implausible/unrelated.

Melanie thought that the man was really [psycho, insane, murder, modest, circus] after learning of his horrific crimes.

He needed to buy some suitable [boots, shoes, hiker, games, check] because he was a keen bushwalker.

They would probably need [eight, three, count, money, house] bottles of wine for the dinner party.

Patty needed to buy some [glue, tape, join, rice, jury] because she had to mend the ripped photo.

David began to prepare his [supper, dinner, hungry, speech, wonder] with the ingredients he bought at the supermarket.

He recommended a lovely [perch, trout, foods, clock, crazy] that was cooked with a lemon sauce.

She stared at the green [toad, frog, damp, hose, defy] swimming in the backyard pond.

The sharp tooth [ache, pain, sore, edge, wood] caused her to miss work all week.

The basement contained [brown, green, shade, pipes, sound] carpet that was ugly and worn.

She gave the dog a quick [wash, bath, suds, meal, vote] after returning from the beach.

He hoped that the heavy [armour, helmet, defend, donkey, angles] would protect him in battle.

She hoped she would be a good [parent, mother, adults, farmer, better] because this was her first child.

The chef decided he would [grill, roast, stove, punch, pupil] some chicken for tonight's special.

The little girl had pretty [blonde, yellow, bleach, wooden, scheme] hair and resembled her grandmother.

The filthy and rusty [pans, pots, cook, keys, rays] made the food taste terrible.

They thought that Joel was a real [geek, nerd, book, hoot, volt] because he was always studying.

Toby kept his money in a large [barn, shed, calf, pile, woke] because he lived on a farm.

They would probably need [eleven, twelve, amount, advice, forest] years for the economy to fully recover.

There was a fantastic [view, look, eyes, home, left] from every window in the house.

Suzy really loved [candy, sugar, taste, birds, match] because she was never allowed to eat it as a kid.

They often played [poker, cards, chips, sport, rapid] after work at the casino.

She remembered the exact [date, year, past, fact, give] when man first walked on the moon.

She found her mother's [skirt, dress, trend, horse, basis] hanging at the back of the wardrobe.

After reading the newspaper, Matthew felt [shame, sorry, abuse, below, drink] about Australia's treatment of refugees.

She approached the heavy [gate, door, shut, girl, sure] before realising it was locked.

She remembered that Kathy [jogs, runs, move, kept, wife] around the creek every afternoon.

They used large [saws, axes, chop, tags, sobs] when they cleared the land for farming.

The clerk shouted for us to quickly [move, come, away, work, must] towards the counter.

They got inside the cramped [buggy, wagon, dunes, booth, mouse] before driving to the next stop.

They recorded the number of white [vans, cars, road, dogs, rent] during the vehicle audit.

He carefully placed his large [sword, knife, stabs, sheep, skill] down after the fight.

They noticed the young [fawn, deer, wild, folk, lump] eating acorns in the forest.

The child had a large [face, head, skin, room, less] with big, expressive eyes.

They slowly filled four [mugs, cups, hold, rows, noon] with coffee for the guests.

There was a light [mist, rain, fogs, wine, sick] when they left the cinema.

They carefully toured the destroyed [ship, boat, ahoy, farm, diet] that had washed ashore.

Tom asked if I would [lend, loan, owes, kick, lamb] some money to him for his car.

She picked up the small [turnip, potato, garden, bullet, critic] from the counter and began to peel it.

They visited the small [chapel, church, preach, prison, effect] before booking their wedding.

She enjoyed the fresh [ocean, beach, waves, bread, birth] breeze on her vacation in Fiji.

He could not even [limp, step, foot, plan, nose] over the finish line after falling during the race.

I did not think that it would [smash, break, glass, begin, older] when I dropped it.

We had some [fries, chips, eaten, coats, stole] while we watched the big game.

She would often [chat, talk, says, keep, week] with Evan about his many brothers and sisters.

Hannah would often [sniff, smell, noses, clean, moral] some of the flowers that grew outside her window.

Tina wanted a huge [lawn, yard, land, cake, fail] because she loved lying on the grass in the sun.

She thought she would [retch, vomit, spews, stall, monks] when she smelled the sour milk.

They stepped into the quiet [cellar, garage, bottle, avenue, injury] that had the tall wooden shelves.

Angel could only [laugh, smile, happy, reach, worse] after hearing the good news.

They decided they would [dive, swim, deep, rely, ours] under the fallen branch that lay across the river.

He stepped back from the dangerous [ledge, cliff, climb, snake, bacon] before admiring the splendid view.

He could not believe how many [bugs, ants, wing, tins, trim] were living under the pot plant.

The kids would never [chew, bite, eats, drag, sums] their food properly.

She eventually found an empty [stool, chair, relax, glass, usual] that she could sit on.

The jacket was stored [below, under, depth, until, thing] some other items of clothing.

The teacher said that they should [print, write, words, watch, mouth] their names at the top of the page.

He bought a brown [vest, coat, wear, sofa, moon] that he planned to wear to the party.

We watched the large [hawk, bird, soar, lake, blow] swoop down to get the food.

The little boy kept his precious [coin, cash, cent, suit, sold] inside his leather wallet.

She decided to learn [waltz, dance, party, maths, solid] because she had always loved the music.

Around the house were several [brick, stone, layer, hills, reach] columns that were painted white.

It was quite an unusual [town, city, maps, idea, miss] but it was located on the beautiful coastline.

The violent thug's [rival, enemy, match, teeth, exist] started a massive brawl.

They saw the other [kids, boys, play, army, deep] after their class had finished.

She looked at the small [shack, house, build, women, today] before planning the renovation.

There was a terrible [noise, sound, heard, story, maybe] when the waiter dropped the tray of drinks.

Tom stood by his very [tidy, neat, sort, loud, aunt] desk and waited for his boss to speak to him.

Pete decided he should [ring, call, says, stop, west] his bank about the unusual transactions on his account.

A lot of people like [toast, bread, plate, trees, month] with their morning coffee at breakfast.

He was told that both [legs, feet, walk, days, full] would need to be amputated.

The wooden crate was filled with [jars, cans, keep, mice, pubs] that she was donating to charity.

When she saw the yellow [cabs, taxi, fare, corn, self] she knew she was in New York City.

He would often [choke, cough, lungs, float, ratio] when he ate Indian food because of the intense spices.

He placed the small [budgie, canary, chirps, peanut, cliché] back into its cage and topped up its water dish.

The woman's black [shawl, scarf, wears, olive, skies] draped over her shoulders.

The kids brought a small [torch, flame, glows, spoon, fatal] with them to illuminate the caves.

Bill was an excellent [partner, husband, married, student, chapter] because he was always willing to compromise.

They were ready for the short [exam, test, mark, film, safe] because they had studied all night.

Hector used a small [chisel, hammer, carves, pistol, scores] when he created the stone statue.

We saw a large [cheetah, leopard, safaris, diamond, analogy] during our trip to the zoo.

APPENDIX B

Random effects for the (G)LMM analyses

Group	Random effect	Linear mixed models					Generalized linear mixed models		
		First fixation	Single fixation	Gaze duration	Go-past duration	Total duration	First-pass fixation	Regressions out	Regressions in
Subject	Intercept	38.79	42.63	54.64	87.14	100.91	0.837	0.892	1.039
	Identical preview benefit	24.31	29.83	47.07	63.93	77.18	0.145	0.643	0.250
	Plausibility	13.97	20.07	10.76	32.19	18.30	0.239	0.304	0.086
	Relatedness	2.91	6.43	8.39	12.76	7.06	0.095	0.234	0.194
	Plausibility $\times$ Relatedness	1.94	6.61	11.91	1.10	23.38	0.318	0.201	0.074
Item	Intercept	13.09	15.17	23.42	34.25	44.13	0.504	0.524	0.421
	Identical preview benefit	-	3.90	1.57	-	-	0.196	0.383	-
	Plausibility	-	16.16	9.75	-	-	0.205	0.375	-
	Relatedness	-	7.86	18.54	-	-	0.189	0.137	-
	Plausibility $\times$ Relatedness	-	15.31	4.05	-	-	0.160	0.373	-
Residual		87.09	86.36	109.34	215.43	184.08	-	-	-

Note. Values are standard deviations. Missing values indicate the model did not include the random effect due to non-convergence.

APPENDIX C

Contrast matrix used in the (G)LMM analyses

Condition	Identical PB	Plausibility effect	Relatedness effect	Plausibility × Relatedness
Identical	-0.8	-	-	-
Plausible/Related	0.2	-0.5	-0.5	-0.5
Implausible/Related	0.2	0.5	-0.5	0.5
Plausible/Unrelated	0.2	-0.5	0.5	0.5
Implausible/Unrelated	0.2	0.5	0.5	-0.5

Table 1

Mean (and Standard Deviation) Stimulus Characteristics and Norming Data

Variable	Preview condition				
	Identical	Plausible Related	Implausible Related	Plausible Unrelated	Implausible Unrelated
<i>Stimulus characteristics</i>					
Length	4.73 (0.80)	4.73 (0.80)	4.73 (0.80)	4.73 (0.80)	4.73 (0.80)
Log frequency (HAL)	8.29 (1.87)	9.71 (1.87)	9.16 (2.06)	9.69 (1.45)	9.85 (1.89)
Neighborhood size (N)	5.04 (4.31)	5.19 (4.27)	5.41 (4.06)	5.49 (4.90)	5.25 (3.85)
Bigram frequency	30.04 (20.82)	32.53 (21.58)	37.22 (32.79)	32.33 (21.01)	32.60 (26.85)
Trigram frequency	5.26 (3.28)	5.09 (2.90)	6.24 (4.50)	5.62 (3.60)	5.64 (5.19)
Shared letters with target ^a	4.73 (0.00)	0.59 (0.76)	0.44 (0.69)	0.45 (0.57)	0.38 (0.58)
Letter overlap with target ^b	1.00 (0.00)	0.13 (0.17)	0.09 (0.14)	0.10 (0.13)	0.08 (0.12)
<i>Norming data</i>					
Cloze predictability	.01 (.02)	.02 (.05)	.00 (.00)	.02 (.05)	.00 (.00)
Sentence fragment plausibility (1-7 scale)	6.21 (0.91)	6.20 (0.83)	2.84 (1.04)	6.13 (1.06)	2.57 (0.98)
Semantic relatedness to target (1-7 scale)	-	5.68 (0.91)	4.41 (1.12)	1.64 (0.56)	1.48 (0.55)

^a Mean number of letters shared with the target in the same position.

^b Mean proportion of letters shared with the target in the same position, relative to word length.

Table 2

Mean (and Standard Error) Reading Measures on the Target across Conditions

Measure	Preview condition				
	Identical	Plausible Related	Implausible Related	Plausible Unrelated	Implausible Unrelated
<i>Fixation duration measures</i>					
First fixation duration	231 (2.5)	244 (3.0)	264 (3.1)	248 (3.3)	266 (3.0)
Single fixation duration	233 (2.7)	245 (3.2)	270 (3.5)	251 (3.7)	274 (3.4)
Gaze duration	258 (3.5)	276 (3.8)	308 (4.0)	285 (4.1)	305 (3.7)
Go-past duration	295 (5.6)	334 (7.9)	391 (8.8)	347 (7.1)	390 (7.3)
Total viewing duration	321 (6.1)	373 (6.2)	400 (6.9)	401 (7.1)	403 (6.1)
<i>Fixation probability measures</i>					
First-pass fixation	.83 (.01)	.82 (.01)	.83 (.01)	.82 (.01)	.85 (.01)
Regressions-out	.10 (.01)	.13 (.01)	.18 (.01)	.16 (.01)	.19 (.01)
Regressions-in	.19 (.01)	.28 (.01)	.25 (.01)	.30 (.01)	.26 (.01)

Table 3

Results of the Linear Mixed-Effects Models for Fixation Duration Measures. Significant

Effects are Indicated in Bold.

Measure	Fixed effect	<i>b</i>	<i>SE</i>	<i>t</i>
First fixation duration	Intercept	248.28	4.37	56.67
	Identical preview benefit (PB)	23.80	3.87	6.15
	Plausibility	17.31	3.02	5.72
	Relatedness	1.87	2.68	0.70
	Plausibility × Relatedness	-1.15	2.67	-0.43
	Overall proficiency	-13.29	3.39	-3.92
	Identical PB × Proficiency	-1.49	3.21	-0.46
	Plausibility × Proficiency	-5.95	2.53	-2.35
	Relatedness × Proficiency	-0.22	2.25	-0.10
	Plausibility × Relatedness × Proficiency	0.07	2.24	0.03
Single fixation duration	Intercept	254.18	4.83	52.60
	Identical preview benefit (PB)	28.58	4.43	6.45
	Plausibility	22.25	4.00	5.56
	Relatedness	4.05	3.12	1.30
	Plausibility × Relatedness	1.61	3.46	0.46
	Overall proficiency	-14.52	3.72	-3.91
	Identical PB × Proficiency	-1.70	3.64	-0.47
	Plausibility × Proficiency	-6.33	2.97	-2.13
	Relatedness × Proficiency	-0.52	2.50	-0.21
	Plausibility × Relatedness × Proficiency	-1.34	2.51	-0.53
Gaze duration	Intercept	281.82	6.32	44.59
	Identical preview benefit (PB)	34.07	6.09	5.60
	Plausibility	24.35	3.69	6.60
	Relatedness	1.62	4.04	0.40
	Plausibility × Relatedness	5.57	3.60	1.55
	Overall proficiency	-16.67	4.73	-3.53
	Identical PB × Proficiency	-1.56	5.03	-0.31
	Plausibility × Proficiency	-7.73	2.96	-2.62
	Relatedness × Proficiency	-1.30	2.90	-0.45
	Plausibility × Relatedness × Proficiency	-2.23	2.99	-0.75
Go-past duration	Intercept	345.79	10.09	34.26
	Identical preview benefit (PB)	67.51	9.82	6.88
	Plausibility	46.64	7.37	6.33
	Relatedness	3.84	6.72	0.57
	Plausibility × Relatedness	-6.81	6.59	-1.03
	Overall proficiency	-35.76	7.68	-4.66

	Identical PB × Proficiency	-5.93	8.15	-0.73
	Plausibility × Proficiency	-7.22	6.16	-1.17
	Relatedness × Proficiency	8.23	5.63	1.46
	Plausibility × Relatedness × Proficiency	8.55	5.53	1.55
Total duration	Intercept	373.06	11.66	32.00
	Identical preview benefit (PB)	71.52	10.08	7.10
	Plausibility	11.80	5.94	1.99
	Relatedness	12.27	5.68	2.16
	Plausibility × Relatedness	-13.03	6.13	-2.13
	Overall proficiency	-36.53	8.67	-4.21
	Identical PB × Proficiency	-3.85	8.34	-0.46
	Plausibility × Proficiency	-6.24	4.98	-1.25
	Relatedness × Proficiency	-8.24	4.76	-1.73
	Plausibility × Relatedness × Proficiency	-3.03	5.14	-0.59

Note. PB = preview benefit.

Table 4

Results of the Generalized Linear Mixed-Effects Models for Fixation Probability Measures.

Significant Effects are Indicated in Bold.

Measure	Fixed effect	<i>b</i>	<i>SE</i>	<i>z</i>
First-pass fixation	Intercept	1.86	0.11	17.13
	Identical preview benefit (PB)	0.01	0.09	0.10
	Plausibility	0.13	0.08	1.57
	Relatedness	0.11	0.08	1.38
	Plausibility × Relatedness	0.04	0.08	0.42
	Overall proficiency	-0.34	0.08	-4.51
	Identical PB × Proficiency	-0.06	0.07	-0.85
	Plausibility × Proficiency	-0.03	0.06	-0.46
	Relatedness × Proficiency	-0.01	0.06	-0.18
	Plausibility × Relatedness × Proficiency	-0.02	0.07	-0.30
Regressions-out	Intercept	-2.03	0.12	-17.20
	Identical preview benefit (PB)	0.80	0.15	5.31
	Plausibility	0.29	0.10	2.76
	Relatedness	0.14	0.09	1.53
	Plausibility × Relatedness	0.05	0.10	0.48
	Overall proficiency	-0.22	0.09	-2.57
	Identical PB × Proficiency	0.04	0.12	0.29
	Plausibility × Proficiency	0.00	0.08	0.02
	Relatedness × Proficiency	0.13	0.08	1.63
	Plausibility × Relatedness × Proficiency	-0.05	0.08	-0.66
Regressions-in	Intercept	-1.34	0.12	-10.92
	Identical preview benefit (PB)	0.63	0.10	6.36
	Plausibility	-0.23	0.08	-3.05
	Relatedness	0.07	0.08	0.84
	Plausibility × Relatedness	-0.07	0.08	-0.90
	Overall proficiency	-0.28	0.09	-2.95
	Identical PB × Proficiency	0.03	0.09	0.33
	Plausibility × Proficiency	0.01	0.07	0.14
	Relatedness × Proficiency	-0.13	0.07	-1.90
	Plausibility × Relatedness × Proficiency	-0.10	0.07	-1.54

Note. PB = preview benefit

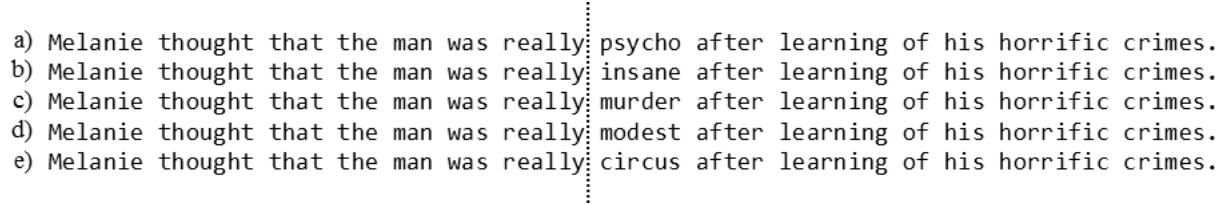
- 
- a) Melanie thought that the man was really: psycho after learning of his horrific crimes.
 - b) Melanie thought that the man was really: insane after learning of his horrific crimes.
 - c) Melanie thought that the man was really: murder after learning of his horrific crimes.
 - d) Melanie thought that the man was really: modest after learning of his horrific crimes.
 - e) Melanie thought that the man was really: circus after learning of his horrific crimes.

Figure 1. Examples of the preview conditions used in the experiment: (a) identical, (b) plausible and semantically related, (c) implausible and semantically related, (d) plausible and unrelated, (e) implausible and unrelated. The invisible boundary is represented by the dashed line. In all conditions, the identical target word was displayed when the reader's eye crossed the boundary.

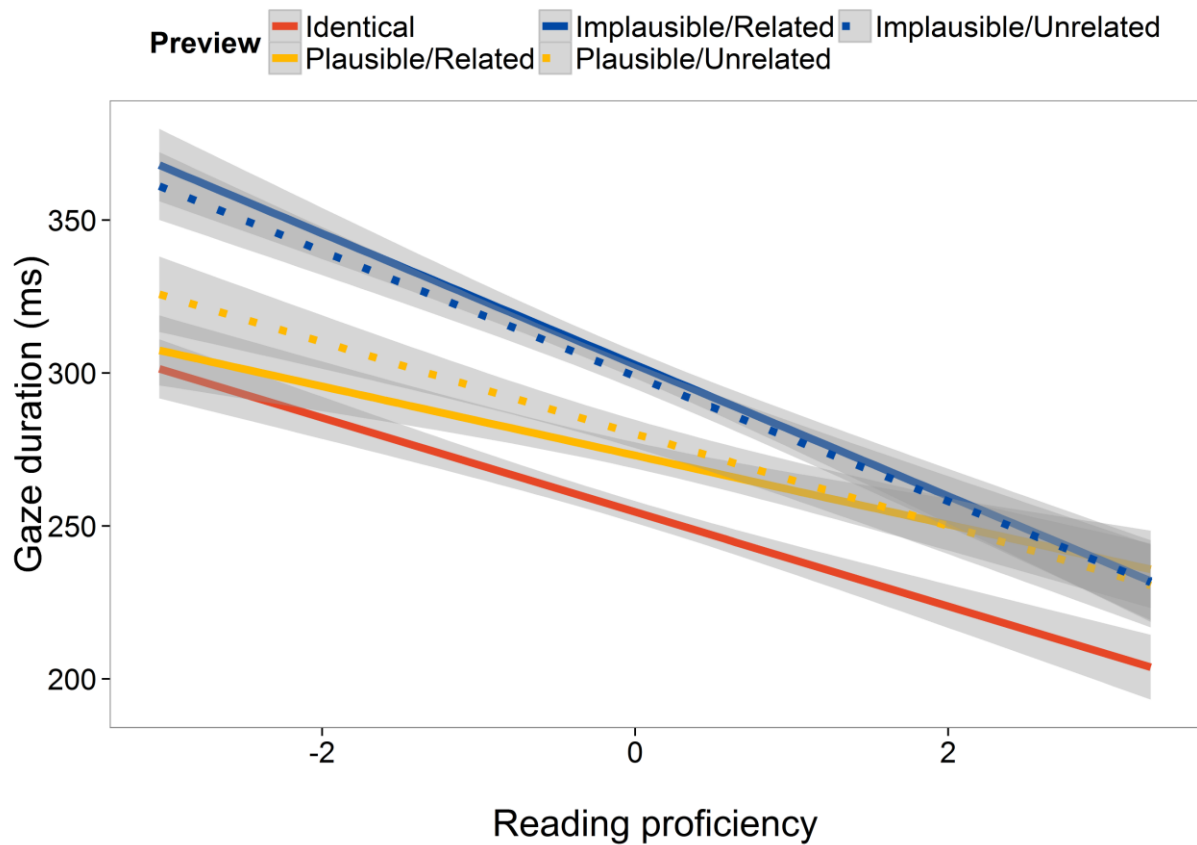


Figure 2. Partial effect of the continuous reading proficiency variable on gaze duration on the target word for each of the preview conditions in the experiment. Lines are regression estimates over data points from which subject and item random variance has been removed with the *remef* function (Hohenstein & Kliegl, 2015) and error bands represent 90% confidence intervals. Solid lines represent related previews; dotted lines are unrelated previews. From the top, the blue lines represent implausible previews, the gold lines are plausible previews, and the red line is the identical condition. Please see online version of article for the color version of this figure.

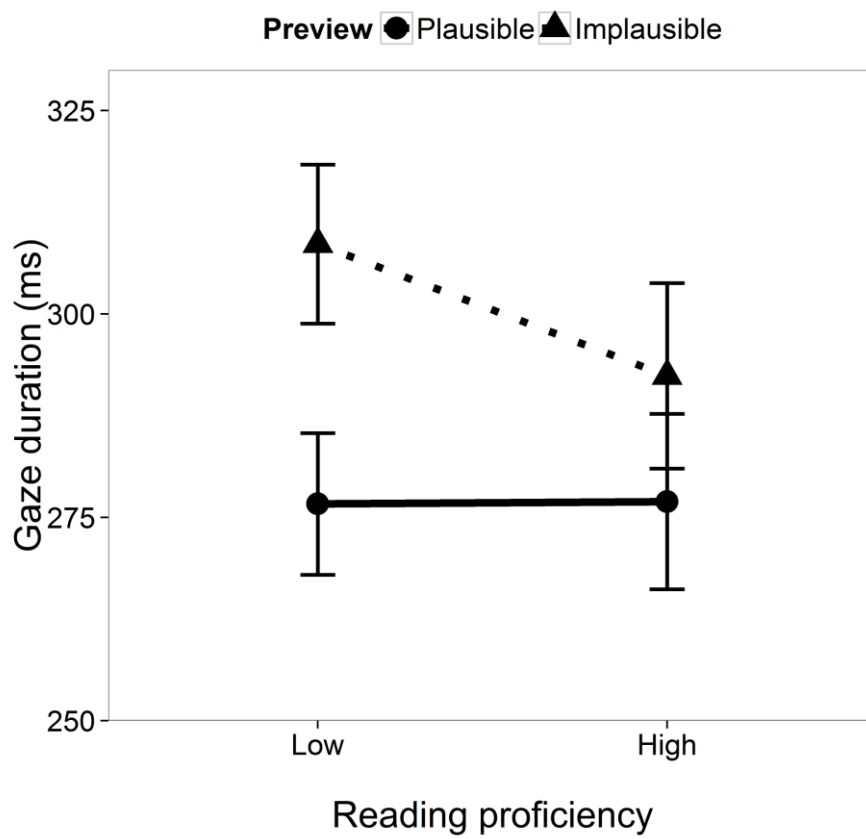


Figure 3. Mean gaze duration (aggregated by subjects) for the plausible (average of the plausible related and unrelated previews) and implausible conditions (average of the implausible related and unrelated previews) based on a median split of reading proficiency. Error bars represent 95% confidence intervals.