Readers can identify the meanings of words without looking at them: Evidence from regressive eye movements

Elizabeth R. Schotter[1] · Anna Marie Fennell[1]

Abstract

Previewing words prior to fixating them leads to faster reading, but does it lead to word identification (i.e., semantic encoding)? We tested this with a gaze-contingent display change study and a subsequent plausibility manipulation. Both the preview and the target words were plausible when encountered, and we manipulated the end of the sentence so that the different preview was rendered implausible (in critical sentences) or remained plausible (in neutral sentences). Regressive saccades from the end of the sentence increased when the preview was rendered implausible compared to when it was plausible, especially when the preview was high frequency. These data add to a growing body of research suggesting that linguistic information can be obtained during preview, to the point where word meaning is accessed. In addition, these findings suggest that the meaning of the fixated target does not always override the semantic information obtained during preview.

Keywords Eye movements and reading · Reading · Word meaning · Word recognition

Introduction

Readers move their eyes from word to word to take advantage of high visual acuity in central vision, but also obtain visual information from non-central vision (i.e., during *parafoveal preview*). Although research suggests that readers obtain enough information during preview to initiate eye-movement plans (see Schotter, 2018; Schotter, Angele, & Rayner, 2012), does the preview also influence what they understand from the text? If preview is used to identify words, it means the reading system is rather risky by performing word recognition based on low-quality visual information.

Preview effects on fixation behavior Initial theories about parafoveal preview suggested that the preview was used for *trans-saccadic integration* (i.e., information from the preview

was merged or compared with information from the target once it was fixated; Pollatsek, Lesch, Morris, & Rayner, 1992; Rayner, 1975; see Cutter, Drieghe, & Liversedge, 2015). These theories were based on an eye-tracking paradigm (i.e., the *gaze-contingent boundary paradigm*; Rayner, 1975) that dissociates the preview from the directly fixated word by showing one stimulus during preview, which changes to a different word (i.e., the target) once the reader makes an eye movement to it. This paradigm yields two primary findings: (1) a *preview validity effect* whereby readers fixate for less time when the preview had been *valid* (e.g., identical or similar to the target) than *invalid* (e.g., a different stimulus; see Schotter et al., 2012; Vasilev & Angele, 2017), and (2) a *preview plausibility effect* whereby readers fixate for less time when the preview had been semantically *plausible* compared to *implausible*, regardless of whether it was orthographically or semantically related to the target (Schotter & Jia, 2016; Veldre & Andrews, 2016). The preview plausibility effect, and a related N400 effect in response to parafoveally presented words in fixed-gaze event-related potentials (ERP) studies (e.g., Barber, Doñamayor, Kutas, & Münte, 2010; see Schotter, 2018), suggests that readers can obtain semantic information from the preview. Moreover, studies finding *shorter* fixation durations on target words following invalid higher-frequency plausible previews than valid low-frequency previews suggest that higher-frequency words are more likely

✉ Elizabeth R. Schotter
eschotter@usf.edu

[1] Department of Psychology, University of South Florida, 4202 E. Fowler Ave. PCD 4118G, Tampa, FL 33620, USA

⧉ Springer

to be identified during parafoveal preview (e.g., Risse & Kliegl, 2014; Schotter & Leinenger, 2016). In contrast, the preview validity effect suggests that readers may discard preview information when they fixate a different stimulus (i.e., because trans-saccadic integration had failed).

The presence of multiple preview effects has led to detailed theories about how preview is used to plan eye movements (see Schotter, 2018), but less is known about the extent to which readers *complete* word identification based on the preview. Preview plausibility effects suggest that readers can identify the meaning of the preview, at least when it is high frequency (Schotter & Leinenger, 2016), and potentially ignore the visual information from the fixated target (Morrison, 1984). However, first-pass eye-movement measures (the canonical measure in these studies) do not indicate complete recognition because eye-movement plans are initiated prior to the completion of word identification (Henderson & Ferreira, 1990; Morrison, 1984; Reingold & Rayner, 2006). Therefore, to directly test whether the preview was identified requires other manipulations.

Preview effects on word encoding Prior research has used comprehension questions with the preview and target as response options to test identification of the preview. One study found that readers were likely to select the preview if they skipped over or briefly fixated the word and it was plausible (i.e., 70%; Schotter, Leinenger, & von der Malsburg, 2018). Another found a lower base rate for reporting the preview (i.e., 20%; Schotter & Jia, 2016, Exp 2), which decreased with implausibility (i.e., < 10% for implausible previews) and increased with semantic relatedness to the target (i.e., ~35% for antonym previews).

Another approach manipulated the end of the sentence so that the preview or target was rendered implausible and measured regressions from this region (Schotter, von der Malsburg, & Leinenger, 2019). Regressions only increased in response to the implausibility of the target, which suggests that readers did not identify the preview; however, the presence of implausible target trials may have overwhelmed the potentially more subtle effects of the preview. Indeed, the magnitude of the effects of linguistic variables (e.g., word frequency) on reading behavior changes in response to global experiment parameters like comprehension question difficulty (Wotschack & Kliegl, 2013) or reading task (e.g., Schotter, Bicknell, Howard, Levy, & Rayner, 2014). Schotter et al. (2019) also found that regressions out of a *buffer region* between the target and the end of the sentence, which maintained the plausibility of both words, were more likely for invalid compared to valid preview conditions, and this effect was stronger for high-frequency previews (see Schotter et al., 2019, for a detailed discussion), indicating a failure of *trans-saccadic integration* (Cutter et al., 2015; Pollatsek et al., 1992; Rayner, 1975).

In the current study, to reduce the salience of the plausibility manipulation and to more directly assess whether the preview was encoded, we manipulated the plausibility of only the preview in the region at the end of the sentence. We modified the stimuli from Schotter et al. (2019) so that (1) both the preview and the target word were plausible following the preceding sentence context, (2) the remainder of the sentence always maintained the plausibility of the target, (3) the words following the target (i.e., the *buffer region*) always maintained the plausibility of both words, and (4) the end of the sentence was manipulated such that the preview either remained plausible (i.e., in *neutral sentences*) or became implausible (i.e., in *critical sentences*). We predicted that, if the preview was semantically encoded, regressions should increase when the different preview became implausible (i.e., in critical sentences) but not when it remained plausible (i.e., in neutral sentences).

Method

Participants

Eighty college students from the University of South Florida participated in the experiment for course credit. They had normal or corrected-to-normal vision, were native English speakers, and had no history of language or cognitive impairments. Sixteen participants were excluded due to (1) experimental program errors (n = 5), (2) not qualifying for the experiment or failing to follow instructions (n = 5), (3) excessive data loss due to incorrectly timed display changes (e.g., more than 29% of trials; n = 2), or (4) noticing too many display changes (e.g., more than 9; n = 4). This study was approved by the University of South Florida Institutional Review Board and we followed all ethical guidelines with regard to the treatment of human subjects.

Apparatus

Participants were seated approximately 60 cm away from an HP p1230 CRT monitor (1,024 × 768 resolution, 150-Hz refresh rate). Viewing was binocular, but movements of only the right eye were recorded via an SR Research Ltd. Eyelink 1000 eye tracker (sampling rate = 1,000 Hz) in a tower setup (i.e., head movements were restrained with padded forehead and chin rests). Text was displayed on a white background, in the vertical center of the screen in one line of black text (Courier New, 12-point; 2.65 characters subtended 1° of visual angle). Display changes were completed, on average, within 4 ms of the tracker detecting a saccade crossing the invisible boundary, which was located at the beginning of the space preceding the target word.

Materials

Each participant read 120 experimental sentences (Table 1; see Supplemental Materials) in which each member of a high-low frequency word pair served as the target word for a set of sentences (see examples 1 and 2 below) that included critical endings that rendered one of the words implausible (version a) or neutral (endings in which both words remained plausible; version b). In addition to the sentence manipulation, there was a gaze-contingent display-change manipulation, in which the different preview replaced the target word prior to the reader fixating or skipping over it and this was compared to an identical preview condition in which the word did not change. The critical and neutral sentence versions for each target word were matched for number of words and differed in number of characters by an average of 4.68 characters (SD = 3.79, range = 0–19).

Example sentences. Previews and targets are presented in italics, different previews are presented in parentheses, and the end regions are underlined:

(1a) The boy found a red *(phone) scarf* and then he wrapped it around his neck for warmth.
(1b) The boy found a red *(phone) scarf* and then he dropped it on his way to school.
(2a) Danielle unfortunately forgot her new *(scarf) phone* so she couldn't call her mom after her class.
(2b) Danielle unfortunately forgot her new *(scarf) phone* so she was sad when she left this morning.

Normative data Twenty-six participants, who were not in the eye-tracking study, rated the sentences for plausibility using a 7-point scale with endpoints marked with verbal labels (i.e., extremely likely and extremely unlikely). Each participant rated half of the stimuli, counterbalanced across conditions (i.e., the preview/target version), so that each sentence version was rated by approximately half the participants. For the stimuli used in the experiment, the plausibility of the target was higher than the preview in critical sentences, and the plausibility of both words was high and similar in neutral sentences. Ten separate participants provided cloze responses to the fragments of the sentences preceding the preview/target, which indicated that the words were generally not predictable.

Procedure

The eye tracker was calibrated using a three-point calibration scheme at the beginning of the experiment, after any participant-initiated breaks, after every ~60 trials, or if the calibration accuracy dropped below .3° of visual angle. At the beginning of each trial, a fixation point appeared in the center of the screen and, if calibration was accurate, the experimenter initiated the trial. A black box appeared on the left side of the screen at the location of the beginning of the sentence. Once the reader made a fixation within the box, the sentence appeared, which they read silently for comprehension. Once the participant was done reading, they looked at a bullseye to the right of the screen and pressed a button on a response pad to indicate they were done. The response pad was used to answer occasional yes/no questions that followed 62 filler sentences (34% of total trials). Sentences were counterbalanced in a latin-square design with four list versions, and randomized in a unique order for each participant. After the experiment, participants were debriefed and asked how many display changes they noticed.

Results

All fixations remained in the dataset, except for fixations shorter than 81 ms and within one character space of another fixation, in which case the two fixations were combined (i.e., summed), and fixations longer than 800 ms, which were

Table 1 Descriptive statistics of experimental stimuli measured in word length, word frequency (obtained from the English Lexicon Project; Balota et al., 2007), cloze probability, and plausibility

	High-frequency word				Low-frequency word			
	M	SD	Min	Max	M	SD	Min	Max
Word length	5.88	0.92	5	8	5.88	0.92	5	8
Frequency per million	127.35	132.01	10	786	4.40	4.42	0	18
Cloze probability (0–10)	0.29	0.98	0	8	0.36	1.29	0	9
Plausibility (1–7):								
Target (neutral sentence)	4.76	0.74	3.27	6.36	4.70	0.79	2.75	6.55
Different preview (neutral sentence)	4.67	0.80	2.45	6.55	4.74	0.74	3.45	6.27
Target (critical sentence)	4.98	0.73	3.50	6.55	4.99	0.80	2.83	6.36
Different preview (critical sentence)	3.08	0.90	1.45	5.36	3.02	0.74	1.58	4.73

Table 2 Results of linear mixed-effects models (LMMs) for raw and log-transformed single-fixation duration on the target word. Significant effects are indicated in bold

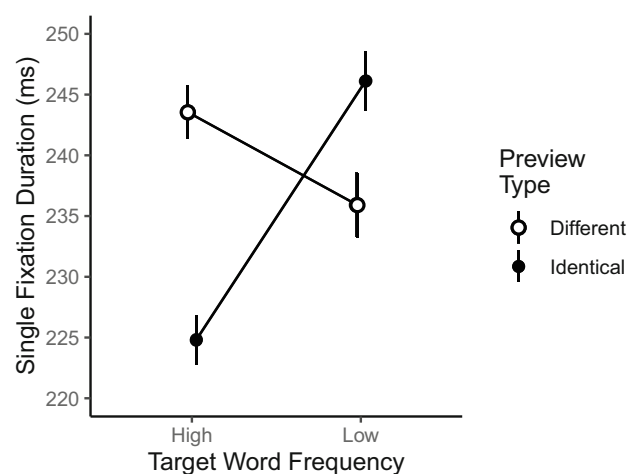|	Raw data			Log-transformed data		
	b	SE	|t|	B	SE	|t|
Intercept	**238.68**	**4.67**	**51.12**	**5.42367**	**0.01937**	**279.95**
Target frequency	**9.00**	**3.51**	**2.57**	0.02633	0.01334	1.97
Preview type (High-Frequency Target)	**17.87**	**3.35**	**5.33**	**0.07690**	**0.01389**	**5.53**
Preview type (Low-Frequency Target)	**-11.02**	**4.84**	**2.28**	**-0.05445**	**0.02030**	**2.68**

eliminated. Trials were excluded if there was a blink or track loss on the target during first-pass reading, or if the display change was triggered by a j-hook or completed after fixation on the target. These exclusions left 6,294 trials available for analysis (88% of the original data).[1] Data were analyzed with (generalized) linear mixed-effects models (G)LMMs using the lmer and glmer functions from the lme4 package (version 1.1-12; Bates, Maechler, Bolker, & Walker, 2015) within the R Environment for Statistical Computing (version 3.3.2; R Development Core Team, 2016). Fixed-effects structures varied by dependent measure and are described in individual sections below. Subjects and items were entered as crossed random effects[2] (see Baayen, Davidson, & Bates, 2008), using the maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013).

Initial reading time on the target

To determine whether properties of the preview had an effect on saccade planning, we analyzed *single fixations*, in which the reader fixated the target once before moving on (71% of the included trials).[3] We used an LMM with custom contrasts that allowed us to directly estimate the magnitude of the preview effect (effect coded as -.5, .5) by nesting it within the effect of target frequency[4] (effect coded as -.5, .5; see Supplemental Materials for other model structures with similar results). For the analyses on both raw and log-transformed data, there was a significant effect of target frequency (both ts > 1.96), a significant standard preview validity effect for the high-frequency target (both ts > 5.32), and a significant

reversed preview validity effect for the low-frequency target (both ts > 2.27; Table 2 and Fig. 1).

Regressions out of the buffer region We analyzed regressions out of the buffer region with a logistic regression with preview type nested within preview frequency, which was nested within sentence type, all entered as effect-coded fixed effects. The model produced seven contrasts: one comparing critical and neutral sentences collapsed across preview, one for the effect of preview type in the neutral sentences, one for the effect of preview type in the critical sentences, and four that represented the effect of preview frequency for every sentence and preview type combination. We only included trials where the target was not skipped and there was not already a regression out of the target region (which would indicate that the reader had detected and responded to the display change already). The analysis revealed no significant effects of sentence type or preview frequency (all ps > .43), but an effect of preview type that was significant for every condition (all ps < .005) except for the low-frequency preview in critical sentences (p = .61; Table 3 and Fig. 2). Because this condition is functionally the same as the low-frequency preview neutral sentence condition (i.e., the manipulation at the end of the sentence had not

---

[1] A few stimuli (n = 13) had minor typos, which occurred after the target. Rerunning the analyses on the regression data with these items excluded did not change the patterns of significance.

[2] For regressions out of the end and buffer regions, item was a specific sentence (i.e., for which target/preview frequency was not manipulated) because the sentence up to the end region was completely identical. For single-fixation duration, item was a pair of sentences (i.e., across which target/preview frequency was manipulated) so that we could include the slope for preview frequency in the random effects for item.

[3] See Supplemental Materials for analyses of other measures.

[4] We thank Reinhold Kliegl for the suggestion of this parsimonious analysis approach.

Fig. 1 Single-fixation duration on the target word as a function of target frequency and preview type. Error bars represent ± 1 SEM based on observations rather than aggregated subject/item condition means

Table 3 Results of a logistic regression for regressions out of the buffer region. Significant effects are indicated in bold

Effect	b	SE	|z|	P
Intercept	**-2.41844**	**0.12479**	**19.38**	**< .001**
Sentence type	-0.10910	0.14145	0.77	.44
Preview frequency (Critical Sentences)	-0.01921	0.18450	0.10	.92
Preview frequency (Neutral Sentences)	-0.12410	0.16734	0.74	.46
Preview type (High-Frequency Preview, Critical Sentences)	**1.60860**	**0.33373**	**4.82**	**< .001**
Preview type (High-Frequency Preview, Neutral Sentences)	**0.99536**	**0.29817**	**3.34**	**< .001**
Preview type (Low-Frequency Preview, Critical Sentences)	0.16404	0.32394	0.51	.61
Preview type (Low-Frequency Preview, Neutral Sentences)	**0.87797**	**0.31043**	**2.83**	**< .005**

yet been encountered), the non-significant effect of preview type, which is in the same direction as the other conditions, may be a type II error.

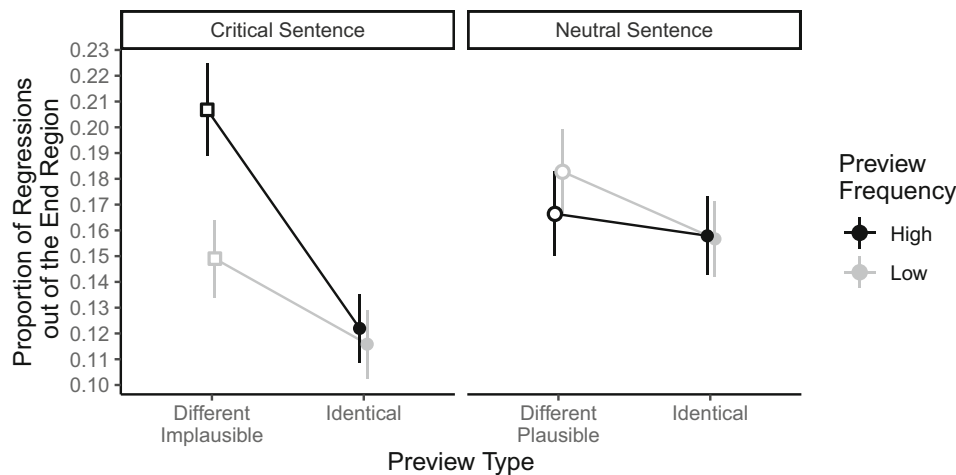**Regressions out of the end region** Because the plausibility of the different preview varied by sentence type, we expected an interaction whereby readers would make more regressions in critical sentences for implausible than identical previews, but there would be no difference in neutral sentences, and we expected these effects to be stronger for high-frequency previews because they would be more likely to be identified. We tested these hypotheses with a logistic regression with the
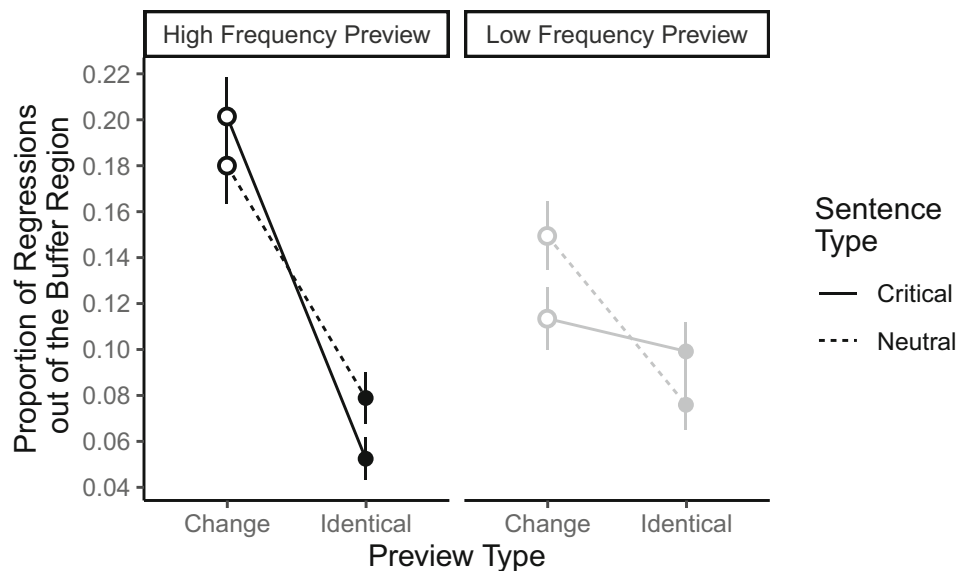
Fig. 2 Regressions out of the buffer region as a function of sentence type and preview type. Error bars represent ±1 SEM based on observations rather than aggregated subject/item condition means

Table 4 Results of a logistic regression for regressions out of the end region when the target was not skipped and there was no regression out of the buffer region. Significant effects are indicated in bold

Effect	b	SE	|z|	P
Intercept	**-2.02358**	**0.12132**	**16.68**	**< .001**
Sentence type	-0.25672	0.13876	1.85	0.06
Preview type (Critical Sentences)	**0.38540**	**0.17649**	**2.18**	**< .05**
Preview type (Neutral Sentences)	0.12137	0.14748	0.82	.41
Preview frequency (Different Preview, Critical Sentences)	**-0.71133**	**0.26720**	**2.66**	**< .01**
Preview frequency (Different Preview, Neutral Sentences)	0.18459	0.23412	0.79	.43
Preview frequency (Identical Preview, Critical Sentences)	-0.04669	0.25415	0.18	.85
Preview frequency (Identical Preview, Neutral Sentences)	0.01974	0.23356	0.09	.93

Fig. 3 Regressions out of the end region as a function of sentence type, preview type, and preview frequency, when the target was not skipped and there was no regression out of the buffer region. Error bars represent ±1 SEM based on observations rather than aggregated subject/item condition means

same nested fixed effects as for the buffer region: preview frequency (effect coded) was nested within preview type (effect coded), which was nested within sentence type (critical vs. neutral; effect coded). We only included trials in which the target was not skipped and there was not already a regression out of the buffer region (80% of included trials) to assess whether readers had encoded the preview when they fixated a different target word. The effect of sentence type was not quite significant ($p = .06$), but was qualified by the nature of the preview: there was an effect of preview type for critical sentences ($p < .05$) but not for neutral sentences ($p = .41$), and this was further qualified by preview frequency whereby the effect of preview type in critical sentences was smaller for low-frequency previews ($p < .01$) but was unaffected by preview frequency in all other conditions (all ps $> .42$; Table 4 and Fig. 3). These data suggest that readers had sometimes encoded the meaning of the preview word because they made regressions when it became implausible (i.e., in critical sentences) and the effect was stronger for previews that were easier to identify (i.e., high frequency).

Discussion

Our study generated three key insights: (1) Readers sometimes semantically identified the preview because regressions out of the end of the sentence increased when it became implausible, especially when it was high frequency, (2) readers used the preview to plan eye movements because fixations on the target were shorter when it was high frequency, regardless of whether it was orthographically or semantically related to the target, and (3) readers sometimes attempted to integrate the preview

and target because regressions out of a semantically neutral region increased when the preview was invalid.

Increased regressions out of the end of the sentence in response to the implausibility of the preview suggests that readers (at least occasionally) encode semantic information during preview. These data align with past research reporting semantic preview benefits on fixation times (e.g., Schotter, 2013), parafoveal N400 effects (e.g., Barber et al., 2010), and that readers sometimes explicitly report having read the preview word even when they directly fixated a different target word (e.g., Schotter et al., 2018; see Schotter, 2018). The fact that this effect was larger for high-frequency previews complements the fixation duration data and suggests that readers obtain more information (including semantics) from previews that are high frequency and therefore may ignore the target information once fixated (e.g., Morrison, 1984; Schotter et al., 2018; see Schotter, 2018). In contrast, readers may not have progressed as far into word identification for low-frequency previews and therefore may be more likely to discard the information in favor of the clearer foveal target.

In addition to the predicted interaction between sentence type and preview type, regressions were more common in neutral sentences than in critical sentences for identical plausible previews.[5] This may be because sentences in which two unrelated words are plausible (e.g., sentence 3b) require somewhat vague and potentially awkward wording compared to sentences for which only one word must make sense (e.g., sentence 3a). In fact, our plausibility norms showed that on average the target words were rated slightly less plausible in neutral (M = 4.73) than critical contexts (M = 4.88).

---

[5] This was also true when we excluded the stimuli with typos or awkward phrasing, see footnote 1.

(3a) The bakery was known for its lovely *aroma* which was <u>always wafting around the block</u>.
(3b) The bakery was known for its lovely *aroma* which was <u>paired with their specialty coffees</u>.

Our finding of both longer fixations following invalid previews for high-frequency targets and shorter fixations following invalid previews for low-frequency targets replicates prior studies showing that initial reading time on the target word is influenced by the preview and not necessarily its relationship to the target (Risse & Kliegl, 2014; Schotter & Leinenger, 2016; Schotter et al., 2018, 2019; Veldre & Andrews, 2016; see Schotter, 2018). Thus, the preview is not only linguistically encoded, as our first finding suggests, but also has a direct influence on reading behavior. In addition, reading behavior showed evidence for trans-saccadic integration failure (Cutter et al., 2015; Pollatsek et al., 1992); after a display change, regressions increased. This suggests that readers sometimes attempted to integrate preview and target information and were more likely to have reading difficulty (i.e., make regressions) when integration was not possible, and this effect was stronger in the buffer region for high-frequency previews and stronger in the target region for low-frequency previews (see Supplemental Materials; see Schotter et al., 2019, for a discussion).

The main finding from this study (increased regressions when the preview became implausible) shows that readers activate a high level of semantic information from parafoveal vision. This adds to a growing literature showing semantic processing of the preview in eye-tracking display-change paradigms (e.g., Schotter, 2013; Schotter & Jia, 2016; Veldre & Andrews, 2016) and ERP paradigms (e.g., Barber et al., 2010; see Schotter, 2018, for a review). Importantly, the current data shed light on what happens to that semantic information once it is initially activated. Although some theories suggest that information is only briefly maintained until it can be integrated with subsequent foveal information or discarded (i.e., via trans-saccadic integration; Cutter et al., 2015), our data suggest that the meaning of (at least a high-frequency) preview may persist after a different word is fixated, and can lead to confusion and later regressions if it is subsequently rendered implausible by the context.

References

Baayen, R. H., Davidson, D. H., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390 – 412.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39,* 445– 459.

Barber, H. A., Doñamayor, N., Kutas, M., & Münte, T. (2010). Parafoveal N400 effect during sentence reading. *Neuroscience Letters, 479(2),* 152-156.

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

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1– 48.

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

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the percep- tual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 16,* 417–429.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology. Human Perception and Performance, 10,* 667–682.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology. Human Perception and Performance, 18,* 148-162.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65-81.

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the EZ Reader model. *Psychological Science, 17,* 742-746.

Risse, S., & Kliegl, R. (2014). Dissociating preview validity and preview difficulty in parafoveal processing of word n + 1 during reading. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 653–668.

Schotter, E. R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language, 69,* 619-633.

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

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics, 74,* 5-35.

Schotter, E.R., Bicknell, K., Howard, I., Levy, R., & Rayner, K. (2014). Task effects reveal cognitive flexibility responding to frequency and predictability: Evidence from eye movements in reading and proofreading. *Cognition, 131,* 1-27.

Schotter, E. R., & Jia, A. (2016). Semantic and plausibility preview benefit effects in English: Evidence from eye movements. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 42,* 1839-1866.

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

Schotter, E. R., Leinenger, M., & von der Malsburg, T. (2018). When your mind skips what your eyes fixate: How forced fixations lead to comprehension illusions in reading. *Psychonomic Bulletin & Review 25,* 1884-1890.

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

Vasilev, M. R., & Angele, B. (2017). Parafoveal preview effects from word N þ 1 and word N þ 2 during reading: A critical review of Bayesian meta-analysis. *Psychonomic Bulletin & Review, 24,* 666-689.

Veldre, A., & Andrews, S. (2016). Is semantic preview benefit due to relatedness or plausibility? *Journal of Experimental Psychology. Human Perception and Performance, 42,* 939-952.

Wotschack, C., & Kliegl, R. (2013). Reading strategy modulates parafoveal-on-foveal effects in sentence reading. *The Quarterly Journal of Experimental Psychology, 66(3),* 548-562.