

ACCEPTED MANUSCRIPT

This is a post-peer-review, pre-copyedit version of an article published in *Psychonomic*

Bulletin & Review. The final authenticated version is available online at:

<https://doi.org/10.3758/s13423-016-1120-8>

Parafoveal preview benefit in sentence reading: Independent effects of plausibility and
orthographic relatedness

Aaron Veldre & Sally Andrews

University of Sydney

Short title: Plausibility and orthographic preview benefit

Contact details

Aaron Veldre

School of Psychology, University of Sydney

Sydney, NSW 2006

Australia

aaron.veldre@sydney.edu.au

+61 2 9351 8990

Abstract

Recent evidence from studies using the gaze-contingent *boundary paradigm* suggests that parafoveal preview benefit is contingent on the fit between a preview word and the sentence context. We investigated whether this *plausibility preview benefit* is modulated by preview/target orthographic relatedness. Participants' eye movements were recorded as they read sentences in which the parafoveal preview of a target word was manipulated. Non-identical previews were plausible or implausible continuations of the sentence that were either an orthographic neighbor of the target or unrelated to the target. First-pass reading measures showed a strong plausibility preview benefit. There was also a benefit from preview/target orthographic relatedness across reading measures. These two preview effects did not interact on any fixation measure. There was also no evidence that the relatedness effect was caused by misperception of an orthographically similar preview as the target word. These data highlight the existence of two independent mechanisms underlying preview effects: a benefit from the contextual fit of the preview word in the sentence and a benefit from the sublexical overlap between the preview and target word.

Word count

4555 (excluding Abstract, Tables, and Figures)

Parafoveal processing underpins fluent reading. Studies utilizing the *gaze-contingent boundary paradigm* (Rayner, 1975) have shown that reading is facilitated when a valid preview of the upcoming word is available in the parafovea. While it is well established that this *preview benefit* also extends to items that share orthographic and/or phonological features with the target word (see Schotter, Angele, & Rayner, 2012), more recent studies suggest that parafoveal words are sometimes processed to the semantic level (Hohenstein & Kliegl, 2014; Schotter, 2013; Schotter, Lee, Reiderman, & Rayner, 2015; Veldre & Andrews, 2016a). This evidence has raised questions about the precise mechanisms underlying preview effects.

Semantic preview effects are theoretically significant because they implicate deep parafoveal processing that was, until recently, thought not to occur in English (e.g., Rayner, Schotter, & Drieghe, 2014). To provide insight into semantic parafoveal processing, Veldre and Andrews (2016b) investigated whether semantic preview benefit is caused by facilitation from the overlapping semantic features shared by the preview and target word or the contextual plausibility of the preview word, two factors that have previously been confounded. The results showed that, when the plausibility of the preview word in the sentence was controlled, there was no effect of preview/target semantic relatedness on fixation duration. There was, however, a strong facilitative effect of preview plausibility on first-pass reading measures of the target word (see also Schotter & Jia, 2016; Yang, Li, Wang, Slattery, & Rayner, 2014). These data demonstrate that preview benefit depends, in part, on the compatibility of the preview with the context.

These results imply that plausible parafoveal words are at least partially incorporated into the reader's developing representation of the sentence meaning, perhaps in place of the target word. Such evidence has been interpreted as challenging the standard view that preview effects are due to the integration of preview information with the target word (Schotter & Jia, 2016). However, the absence of preview/target integration may be specific to

semantic relationships. Because the preview/target word pairs used to assess semantic preview benefit are selected to have minimal orthographic overlap, interference caused by the orthographic discrepancy between a preview and target may counteract any benefit of shared semantic features. It is nevertheless possible that abstract letter identities *are* integrated across saccades, even if this is not the case for semantic information. Effects of orthographic preview/target relatedness might therefore co-occur with effects of preview/context compatibility. The present study investigated this by extending Veldre and Andrews' (2016b) approach of factorially manipulating preview plausibility and relatedness to orthographic preview benefits.

If plausibility preview benefit and orthographic preview benefit are due to separate processes, reflecting contextual facilitation from the match between preview and *sentence* and integration of letter information between preview and *target*, respectively, they should produce additive effects on fixation duration (Sternberg, 1969). Alternatively, preview plausibility and relatedness might yield interactive effects on fixation duration, suggesting that the two factors affect the same stage of processing.

An interaction between plausibility and relatedness could also arise if orthographically related previews are more likely to be 'misperceived' as the target word (e.g., Slattery, 2009). That is, rather than activating the lexical entry of the preview word, preview benefit may occur because the target word is occasionally directly activated by the preview. Reduced fixation duration on these trials would then be functionally due to an identity preview benefit. Critically, if the benefit of an orthographic preview relies on misperception, it should be unaffected by the plausibility of the actual preview word because the preview is misread as the target, which is always a plausible continuation of the sentence. This account therefore predicts a reduced plausibility effect for orthographically similar

previews. The present study was designed to directly assess whether preview effects are due to word misperception.

We conducted two sub-experiments that used the boundary paradigm to assess the same five preview conditions: an identity preview and four conditions that factorially manipulated the contextual plausibility of the preview in the sentence and whether or not the preview was a one-letter different neighbor of the target.

To assess the contribution of misperception of preview words, the target and preview stimuli were based on sets of matched ‘animal’, and ‘non-animal’ words. Experiment 1A included a comprehension probe after each experimental sentence that asked whether or not the sentence contained an animal. Evidence that target detection was more accurate when the preview was an orthographic neighbor of the target rather than an unrelated word would imply that readers occasionally misperceive a neighbor word preview as the target word. Sentences were constructed to be equally plausible for both animal and non-animal words at the location of the target word. In approximately one third of the sentences the target word was an animal. In a further third of the sentences, while the target was a non-animal word, the plausible neighbor preview was an animal word. In the remaining sentences the target was a non-animal but the plausible unrelated preview was an animal word (see Figure 1).

This manipulation also allowed us to probe whether participants encoded the preview or the target word into their sentence representation. Schotter and Jia (2016) found that participants were more likely to incorrectly report having read the preview word when it was a plausible continuation of the sentence, but only when the target was never directly fixated. False detection rates for plausible animal previews for non-animal targets, or vice-versa, provide similar insight into whether the plausibility preview benefit is due to readers encoding a plausible parafoveal word into their representation of the sentence meaning in place of the target word.

The animal detection task used in Experiment 1A required a response to every question. Such comprehension demands have been shown to yield a more cautious reading strategy that may modulate parafoveal processing (Wotschack & Kliegl, 2013). To confirm that the results of Experiment 1A were not specific to the high comprehension demands, a separate group of participants completed Experiment 1B under typical task requirements to respond to occasional comprehension questions.

-
- 1a) The cupboard was full of filthy; rats because nobody had bothered to clean it.
 1b) The cupboard was full of filthy; rags because nobody had bothered to clean it.
 1c) The cupboard was full of filthy; rate because nobody had bothered to clean it.
 1d) The cupboard was full of filthy; junk because nobody had bothered to clean it.
 1e) The cupboard was full of filthy; sigh because nobody had bothered to clean it.
- 2a) There was a large brown; stake near the tomato vines.
 2b) There was a large brown; snake near the tomato vines.
 2c) There was a large brown; stare near the tomato vines.
 2d) There was a large brown; fence near the tomato vines.
 2e) There was a large brown; atoms near the tomato vines.
- 3a) The king valued his favourite; wine more than anything else in his kingdom.
 3b) The king valued his favourite; wife more than anything else in his kingdom.
 3c) The king valued his favourite; wink more than anything else in his kingdom.
 3d) The king valued his favourite; hawk more than anything else in his kingdom.
 3e) The king valued his favourite; seek more than anything else in his kingdom.

Figure 1. Examples of the preview conditions used in the experiment: (a) identical, (b) plausible neighbor, (c) implausible neighbor, (d) plausible unrelated, (e) implausible 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. The three example items show the animal word probe manipulation: (1) animal word as target, (2) animal word as plausible neighbor preview, (3) animal word as plausible unrelated preview.

METHOD

Participants

Seventy-eight students (mean age 19.5 years) from The University of Sydney participated in exchange for course credit (38 in Experiment 1A; 40 in Experiment 1B). All

participants had normal or corrected-to-normal vision and reported that English was their first language.

Materials and Design

The critical stimuli were 80 sentences ($M=13.3$ words) in which the preview of a 4-6 letter target word was manipulated (see Table 1 for stimulus characteristics). Prior to the reader making a saccade across an invisible boundary at the end of the pre-target word ($M=5.5$ letters), one of five preview words occupied the target location: (1) *Identical* to the target; (2) *Plausible Neighbor*: an acceptable continuation of the sentence that differed from the target by a single non-initial letter¹; (3) *Implausible Neighbor*: a semantically and/or syntactically implausible word given the prior context that was a neighbor of the target; (4) *Plausible Unrelated*: an acceptable sentence continuation that was orthographically unrelated to the target; or (5) *Implausible Unrelated*: an implausible sentence continuation that was unrelated to the target. The plausible previews were selected to be compatible with both the pre- and post-target context. All sentences appeared in all preview conditions across five counterbalanced lists, interspersed with 30 filler sentences.

Stimulus norming. A separate group of 20 participants completed a cloze norming task in which they were given each sentence frame up to and including the pre-target word and asked to generate the most likely continuation of the sentence. The results confirmed that the target/preview words were not predictable (see Table 1).

¹ In 57.5% of items, the replacement letter occurred in the same location for the plausible and implausible neighbors (i.e., either both medial or both final letter replacements). The proportion of medial and final replacements was also similar for Plausible and Implausible neighbor previews (75% vs. 65% medial replacements respectively).

Table 1

Mean (and Standard Deviation) of Stimulus Characteristics and Norming Data

Variable	Preview condition				
	Identical	Plausible Neighbor	Implausible Neighbor	Plausible Unrelated	Implausible Unrelated
Length	4.50 (0.53)	4.50 (0.53)	4.50 (0.53)	4.50 (0.53)	4.50 (0.53)
Log frequency (HAL)	8.21 (1.52)	7.86 (1.67)	8.18 (2.09)	8.02 (1.38)	7.91 (1.30)
Proportion of letter overlap with target ^a	1.00 (0.00)	0.77 (0.03)	0.77 (0.03)	0.09 (0.13)	0.05 (0.10)
Cloze predictability	0.01 (0.02)	0.01 (0.03)	0.00 (0.00)	0.00 (0.02)	0.00 (0.00)
Sentence fragment plausibility (1-5 scale) ^b	4.46 (0.62)	4.33 (0.60)	2.06 (0.81)	4.45 (0.65)	1.70 (0.52)
Complete sentence plausibility (1-5 scale) ^c	4.42 (0.63)	4.34 (0.73)	1.60 (0.55)	4.39 (0.68)	1.30 (0.49)

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

^b Plausibility rating for sentence up to and including the target/preview word.

^c Plausibility rating for the whole sentence including the target/preview word.

Two additional groups of participants provided plausibility ratings on a 5-point scale for the sentence frames up to and including the target/preview word ($n=22$) or for the whole sentence including each of the target/preview words ($n=18$). The fragments/sentences containing the target, plausible neighbor, and plausible unrelated previews were rated as highly acceptable and did not differ significantly from one another (all $t_s < 1.17$, $p_s > .245$). Both implausible previews were rated significantly lower in acceptability than each of the plausible previews (all $t_s > 18.45$, $p_s < .001$), but the small differences in mean plausibility rating between the implausible neighbor and implausible unrelated previews were significant (fragment: $t(79)=3.23$, $p=.001$; sentence: $t(79)=3.62$, $p=.001$).

Apparatus and Procedure

Participants' eye movements were recorded by an EyeLink 1000 as they read sentences on a ViewSonic 225fb CRT monitor with 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. A chin and forehead rest minimized participants' head movements at a distance of 60 cm from the monitor. At this distance 2.5 characters subtended 1 degree of visual angle.

Participants were instructed to read the sentences for meaning and to respond to comprehension questions when they appeared. Participants in Experiment 1A were told they would be asked whether or not sentences contained an animal. At the beginning of the experiment a 3-point calibration procedure was followed by three practice trials with comprehension questions (and a further three practice trials with animal word probes in Experiment 1A). The 80 experimental and 30 filler sentences were then presented in random order. At the beginning of each trial, a fixation point occupied the location of the first letter of the sentence. The sentence was displayed when the participant made a stable fixation on this point or a new calibration procedure was performed if necessary. Mean calibration error was less than 0.3 degrees of visual angle. On all filler trials in both experiments, the sentence was followed by a multiple-choice comprehension question that required a moderate understanding of the meaning of the sentence. In Experiment 1A the experimental sentences were followed by the question: "Did this sentence contain an animal(s)?"

RESULTS

Fixations below 80 ms that were within one letter space of an adjacent fixation were merged with that fixation and remaining fixations below 80 ms or above 1000 ms were eliminated (6.3% of fixations). Trials were discarded if there was a blink immediately before or after a fixation on the target (3.7% of trials) or if the display change completed more than

10 ms into a fixation or was triggered by a saccade landing to the left of the boundary (12.2% of trials). Gaze and go-past durations above 2000 ms and total durations above 4000 ms were also excluded (9 trials). These exclusions left 5240 trials (84.0% of the data) available for analysis.

The following reading measures were analyzed: *first fixation duration* (the duration of the first fixation on the target regardless of the number of fixations it received), *single fixation duration* (the fixation duration when only one first-pass fixation was made on the target), *gaze duration* (the sum of all first-pass fixations on the target), *go-past duration* (the sum of all fixations from the first fixation on the target until a word to the right was fixated, including regressions to words earlier in the sentence), and *total duration* (the sum of all fixations on the target including regressions from later in the sentence). Measures of the probability of 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 analyzed. Means for each of these measures are presented in Table 2.

The data were analyzed by (generalized) linear mixed-effects models (LMM) using the *lme4* package (Version 1.1-10; Bates, Maechler, Bolker, & Walker, 2015) in *R* (Version 3.2.2; R Core Team, 2015). The models included subject and item random intercepts and random slopes for the preview effects. Four planned contrasts were tested: (i) *Identical* preview benefit: identical preview vs. average of the non-identical previews; (ii) *Plausibility* effect: the average of the two plausible (non-identical) vs. the average of the two implausible previews; (iii) *Relatedness* effect: the average of the two related vs. the average of the two unrelated previews; and (iv) *Plausibility* \times *Relatedness* interaction.

Table 2

Mean (and Standard Deviation) Reading Measures on the Target for each Experiment and Animal Word Detection Accuracy

(Experiment 1A only) across Preview Conditions

Measure	Experiment 1A					Experiment 1B				
	Identical	Plausible Neighbor	Implaus. Neighbor	Plausible Unrel.	Implaus. Unrel.	Identical	Plausible Neighbor	Implaus. Neighbor	Plausible Unrel.	Implaus. Unrel.
<i>Fixation Duration Measures (ms)</i>										
First fixation	237 (28)	245 (30)	265 (35)	267 (43)	289 (38)	236 (33)	248 (31)	247 (26)	254 (28)	270 (35)
Single fixation	243 (38)	249 (32)	273 (37)	276 (50)	307 (45)	236 (42)	252 (35)	257 (32)	262 (35)	280 (45)
Gaze duration	261 (45)	278 (44)	305 (51)	316 (65)	340 (56)	262 (39)	276 (29)	287 (37)	286 (33)	310 (40)
Go-past duration	298 (61)	349 (69)	355 (65)	391 (86)	443 (81)	297 (60)	333 (53)	359 (64)	342 (66)	379 (75)
Total duration	438 (108)	504 (107)	497 (150)	558 (124)	517 (118)	360 (87)	418 (75)	399 (84)	424 (66)	409 (65)
<i>Fixation Probability Measures</i>										
First-pass fixation	.80 (.11)	.80 (.09)	.83 (.11)	.81 (.09)	.86 (.09)	.75 (.11)	.81 (.10)	.85 (.11)	.79 (.10)	.83 (.08)
Regressions out	.09 (.07)	.15 (.13)	.15 (.11)	.19 (.15)	.27 (.12)	.10 (.10)	.12 (.09)	.16 (.10)	.16 (.12)	.15 (.12)
Regressions in	.42 (.13)	.47 (.15)	.38 (.16)	.51 (.12)	.41 (.13)	.31 (.13)	.36 (.13)	.32 (.14)	.40 (.13)	.27 (.14)
Animal word detection accuracy	.94 (.07)	.91 (.07)	.93 (.07)	.92 (.07)	.92 (.07)	-	-	-	-	-

The major analyses were conducted on pooled data from Experiments 1A and 1B and included a sum-coded contrast assessing the effect of Experiment and interactions with the preview effect contrasts. Estimates yielding t/z values greater than $|1.96|$ were interpreted as significant at the .05 alpha level. Coefficients, standard errors, and t/z values from the (G)LMMs are reported in Tables 3 and 4. Separate LMMs for the two sub-experiments were also conducted to confirm that they yielded the same pattern of significant results. Unless noted, the pattern of significant effects in the combined analyses was identical to that obtained in each individual experiment.

Eye Movement Data

The preview effects are summarized in Figure 2. There was a significant *identical preview benefit* across all duration measures [all $t_s > 5.1$]. Readers were also less likely to fixate the target [$z=4.0$] and less likely to regress from the target to earlier in the sentence after an identical preview [$z=4.2$]. Regressions in did not differ significantly between identical and non-identical conditions [$z=1.8$].

Preview plausibility significant affected all first-pass measures. Fixation durations on the target were shorter after plausible than implausible previews [all $t_s > 3.5$]. When the preview was a plausible word, participants were also less likely to fixate the target [$z=3.8$]² and less likely to regress from the target [$z=2.7$]. However, there was no plausibility preview benefit on total duration [$t=-1.4$] because readers were significantly more likely to regress to the target when the preview was a plausible word [$z=-5.9$]. Thus the plausibility preview benefit was restricted to first-pass reading because it was counteracted by late re-reading of the target.

² Effects of plausibility have typically not been observed as early as skipping (see e.g., Abbott & Staub, 2015). While statistically significant in the present data, the effect was small (~4%). Supplementary analyses revealed that the effect was restricted to 4-letter targets [$b=0.41$, $SE=0.11$, $z=3.60$] and that there was no effect of plausibility on skipping of longer targets [$b=0.18$, $SE=.13$, $z=1.39$]. Further research will be necessary to establish the conditions under which parafoveal plausibility can affect skipping.

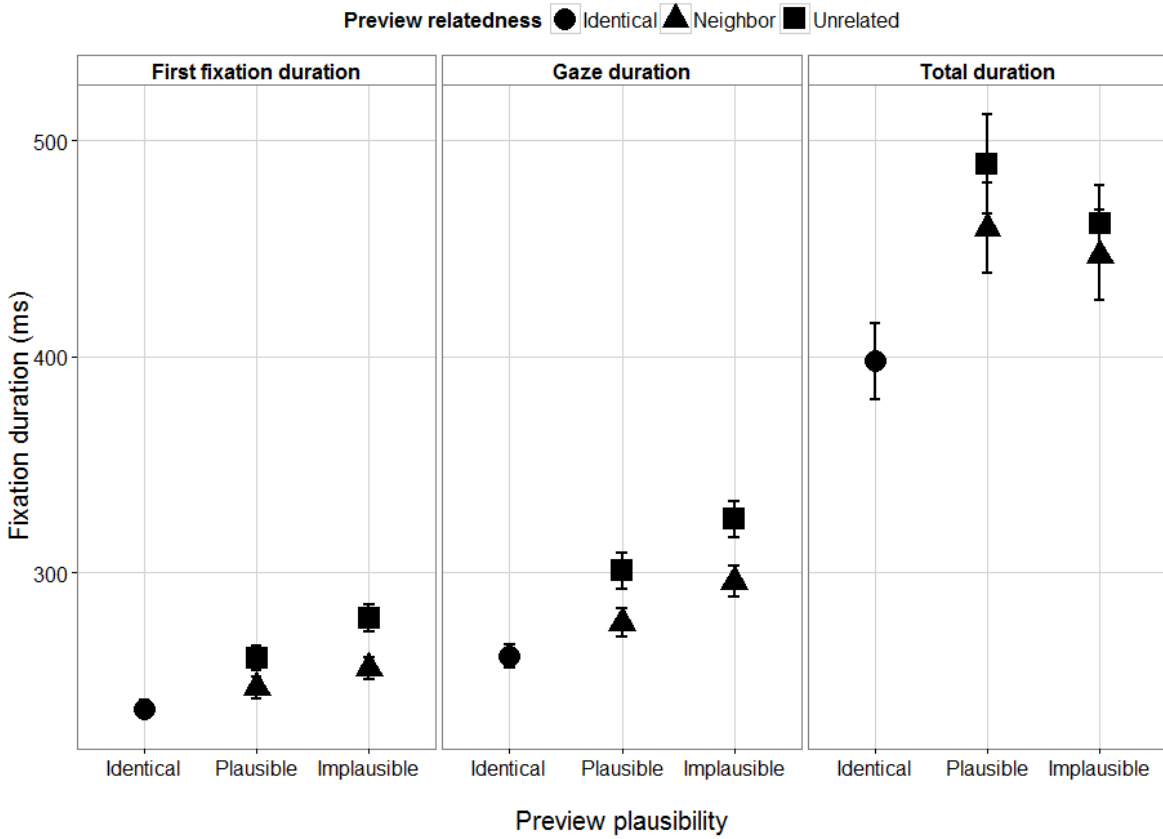


Figure 2. Mean first fixation, gaze, and total duration (aggregated by subjects) for each of the preview conditions. Data are pooled across Experiment 1A and 1B. Error bars represent 95% confidence intervals.

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	255.83	4.23	60.49
	Identical preview benefit	23.37	4.10	5.70
	Plausibility	14.96	3.97	3.77
	Relatedness	17.09	3.98	4.30
	Plausibility × Relatedness	3.55	4.04	0.88
	Experiment	10.92	7.85	1.39
	Identical preview benefit × Experiment	9.75	8.07	1.21
	Plausibility × Experiment	11.89	7.82	1.52
	Relatedness × Experiment	3.30	6.86	0.48
	Plausibility × Relatedness × Experiment	-7.07	7.86	-0.90
Single fixation duration	Intercept	262.56	4.84	54.27
	Identical preview benefit	28.04	4.73	5.93
	Plausibility	20.67	3.98	5.19
	Relatedness	21.60	4.22	5.12
	Plausibility × Relatedness	3.86	4.41	0.88
	Experiment	13.15	8.88	1.48
	Identical preview benefit × Experiment	9.76	9.45	1.03
	Plausibility × Experiment	14.28	7.97	1.79
	Relatedness × Experiment	8.80	8.45	1.04
	Plausibility × Relatedness × Experiment	-1.49	8.82	-0.17
Gaze duration	Intercept	292.03	6.11	47.80
	Identical preview benefit	39.56	5.98	6.62
	Plausibility	21.89	5.60	3.91
	Relatedness	23.61	5.43	4.35
	Plausibility × Relatedness	1.24	5.37	0.23
	Experiment	17.29	11.32	1.53
	Identical preview benefit × Experiment	20.88	11.54	1.81
	Plausibility × Experiment	9.87	10.04	0.98
	Relatedness × Experiment	14.10	9.99	1.41
	Plausibility × Relatedness × Experiment	-8.34	10.08	-0.83
Go-past duration	Intercept	354.65	9.14	38.78
	Identical preview benefit	69.57	8.77	7.93
	Plausibility	30.39	8.11	3.75
	Relatedness	41.11	8.30	4.96
	Plausibility × Relatedness	13.59	7.80	1.74
	Experiment	27.50	16.39	1.68
	Identical preview benefit × Experiment	29.98	17.53	1.71
	Plausibility × Experiment	-1.76	16.22	-0.11
	Relatedness × Experiment	44.34	16.59	2.67
	Plausibility × Relatedness × Experiment	17.05	15.59	1.09

Total duration	Intercept	449.30	18.81	23.89
	Identical preview benefit	67.00	11.50	5.82
	Plausibility	-18.28	10.63	-1.72
	Relatedness	17.85	9.02	1.98
	Plausibility × Relatedness	-5.81	10.97	-0.53
	Experiment	102.50	32.24	3.18
	Identical preview benefit × Experiment	40.51	23.00	1.76
	Plausibility × Experiment	-6.83	21.27	-0.32
	Relatedness × Experiment	27.95	18.04	1.55
	Plausibility × Relatedness × Experiment	-15.92	21.94	-0.73

Table 4

Results of the Generalized Linear Mixed-Effects Models for Fixation Probability Measures and Animal Word Detection Accuracy (Experiment 1A only). Significant Effects are Indicated in Bold.

Measure	Fixed effect	<i>b</i>	<i>SE</i>	<i>z</i>
First-pass fixation	Intercept	1.67	0.11	14.85
	Identical preview benefit	0.38	0.10	4.03
	Plausibility	0.37	0.10	3.84
	Relatedness	0.04	0.10	0.38
	Plausibility × Relatedness	0.09	0.09	0.99
	Experiment	0.10	0.21	0.46
	Identical preview benefit × Experiment	-0.27	0.18	-1.47
	Plausibility × Experiment	0.02	0.18	0.12
	Relatedness × Experiment	0.34	0.19	1.78
	Plausibility × Relatedness × Experiment	0.06	0.17	0.37
Regressions out	Intercept	-1.93	0.12	-16.58
	Identical preview benefit	0.56	0.13	4.23
	Plausibility	0.28	0.10	2.68
	Relatedness	0.44	0.11	3.90
	Plausibility × Relatedness	0.07	0.11	0.60
	Experiment	0.25	0.20	1.23
	Identical preview benefit × Experiment	0.39	0.26	1.49
	Plausibility × Experiment	0.13	0.21	0.63
	Relatedness × Experiment	0.41	0.21	1.93
	Plausibility × Relatedness × Experiment	0.38	0.21	1.82
Regressions in	Intercept	-0.59	0.14	-4.13
	Identical preview benefit	0.18	0.10	1.84
	Plausibility	-0.48	0.08	-5.90
	Relatedness	0.03	0.08	0.39
	Plausibility × Relatedness	-0.13	0.08	-1.62
	Experiment	0.57	0.25	2.27
	Identical preview benefit × Experiment	-0.06	0.19	-0.30
	Plausibility × Experiment	0.03	0.16	0.21
	Relatedness × Experiment	0.19	0.16	1.19
	Plausibility × Relatedness × Experiment	0.14	0.16	0.86
Animal word detection accuracy	Intercept	3.27	0.24	13.56
	Identical	-0.66	0.36	-1.80
	Plausibility	1.34	0.33	4.02
	Relatedness	-0.23	0.31	-0.75
	Plausibility × Relatedness	-0.38	0.33	-1.15
	First-pass fixation	0.14	0.27	0.51
	Regression in	-0.18	0.20	-0.89

Identical × First-pass fixation	0.21	0.72	0.29
Plausibility × First-pass fixation	-1.58	0.59	-2.66
Relatedness × First-pass fixation	1.11	0.60	1.87
Plausibility × Relatedness × First-pass fixation	1.03	0.60	1.73
Identical × Regression in	-0.04	0.52	-0.08
Plausibility × Regression in	-1.29	0.42	-3.08
Relatedness × Regression in	-0.54	0.42	-1.30
Plausibility × Relatedness × Regression in	-0.43	0.42	-1.02

Preview orthographic relatedness affected all duration measures. Relative to an unrelated preview, fixation duration was significantly shorter following a neighbor preview [all t 's > 1.98]. Relatedness did not affect first-pass fixation rate [$z < 1$] but readers were less likely to regress out of the target to earlier in the sentence when the preview was related to the target [$z = 3.9$]. There was no effect of relatedness on regressions back to the target [$z < 1$]. Thus, in contrast to the effect of preview plausibility, the orthographic preview benefit affected reading from when the target word was first fixated until the reader moved past it.

The *Plausibility × Relatedness* interaction did not approach significance on any measure [all t 's and $|z|$'s < 1.7].

In summary, there was a significant effect of preview plausibility on first-pass reading measures and a significant effect of preview/target relatedness on all fixation duration measures. However, these two factors did not interact on any measure suggesting they reflect two independent processes.

Between-experiment differences. The only significant main effect of Experiment was on late re-reading: Experiment 1A, in which participants were required to detect the presence of animal words in addition to responding to occasional comprehension questions, was associated with higher total fixation durations [$t = -3.2$] because of a higher rate of regressions back to the

target [$z=-2.3$]. These late effects presumably reflect a more cautious reading strategy in response to the increased comprehension demands of Experiment 1A. However, there was no effect of experiment on any first-pass measure [all $|t|$ s and $|z|$ s < 1.7].

The only evidence of a significant difference in preview effects between experiments was a Relatedness \times Experiment interaction on go-past duration [$t=2.7$]. Although larger in Experiment 1A, the relatedness effect was significant in both experiments [1A: $b=62.22$, $SE=10.31$, $t=6.04$; 1B: $b=21.43$, $SE=9.86$, $t=2.17$].

Thus, although the higher comprehension demands of Experiment 1A led to a somewhat more cautious reading strategy the overall pattern of results was very similar in both experiments³, providing independent replications of the critical, additive effects of plausibility and relatedness.

Comprehension Analyses

Accuracy for the filler sentence comprehension questions was high (96.1%; range 80-100%) and did not differ between Experiment 1A ($M=96.4\%$) and 1B ($M=95.8\%$), $t < 1$.

The animal detection task in Experiment 1A was designed to check whether the preview or target word was encoded into the participant's representation of the sentence meaning. The correct response was 'Yes' when the target was an animal and 'No' when the target was a non-animal. False alarms to non-animal targets preceded by animal previews (or misses for the reverse arrangement) would suggest that the preview was occasionally encoded instead of the target word.

³ There were some minor differences in the pattern of significance for individual effects between experiments. The identical preview effect on skipping was not significant in Experiment 1A [$z < 1$]. There were no preview effects on regressions out in Experiment 1B [all $|z|$ s < 1.43]. The plausibility effect on first fixation duration was marginally significant in Experiment 1B [$t=1.93$]. The relatedness effect on total duration only reached significance in the combined analysis but not in the separate LMMs for each experiment [t s < 1.92]. All other reported effects were significant in both experiments.

As summarized in Table 2, mean accuracy for detecting whether or not the target was an animal word was very high in all preview conditions. To assess how this was affected by whether or not participants had fixated the target, the GLMM model conducted on the accuracy data included first pass fixation probability, and the probability of regressing back to the target word (both coded as a sum contrasts) and their interactions with the preview contrasts. Accuracy was equally high for related and unrelated previews [$z < 1$] providing no evidence that readers misperceived the preview as the target word. However, there was a significant effect of preview plausibility on target detection accuracy suggesting that readers were more likely to encode the preview instead of the target when it was a plausible continuation of the sentence [$z = 4.0$]. However, a significant interaction with first-pass fixation probability revealed that the plausibility effect on preview encoding was restricted to cases where the reader skipped the target word [$z = -2.7$]. There was also a Plausibility \times Regression interaction because the plausibility effect was only observed when there was no regression to the target word [$z = -3.1$].

These findings converge with those of Schotter and Jia (2016) in showing that, despite the significant plausibility preview benefit on first-pass measures, participants only appeared to incorporate plausible preview words into their representation of sentence meaning in place of the target word on trials in which they never directly fixated the target.

DISCUSSION

We investigated whether preview/target orthographic relatedness modulates plausibility preview effects to shed light on the mechanisms underlying parafoveal preview benefit. The results replicated earlier evidence that a plausible preview benefits first-pass reading relative to an implausible preview (Veldre & Andrews, 2016b). There was also a robust effect of preview/target orthographic relatedness on all reading measures. These two factors produced

strictly additive effects suggesting that benefits of preview plausibility and preview/target orthographic overlap are due to independent processes: a semantic benefit (and/or cost) from the fit of the preview with the context and an orthographic effect due to overlap (or discrepancy) between the preview and target. The proposed mechanisms underlying these independent contributions to preview effects are discussed below.

As elaborated by Veldre and Andrews (2016b), effects of preview plausibility are compatible with the mechanisms that Schotter et al. (2014) proposed to account for semantic preview benefits in E-Z Reader 10 (Reichle, Warren, & McConnell, 2009). Central to this account is the assumption that the co-ordination between lexical processing and saccadic planning operates equivalently under normal reading conditions and in the boundary paradigm. Parafoveal information contributes to both lexical processing of the upcoming word and saccadic planning through adaptive re-programming of saccades to allow skipping of upcoming words that are successfully identified in the parafovea. Critically, because saccades can only be re-programmed during an early, labile stage, parafoveal words will sometimes be identified after the point at which a saccade to the target word can be modified. Unlike normal reading, however, in the boundary paradigm, the parafoveal preview provides misleading information about the upcoming word which can result in the reader programming a saccade away from the target based on information extracted from the preview rather than the target. This can produce an apparent preview benefit on target fixation that is due to a saccade planned before visual information from the target became available to the reading system. When the preview is an acceptable continuation of the sentence, the saccade out of the target will be executed. However, when the preview is implausible, failure of the postlexical integration processes that follow

identification of the preview cancels the planned saccade, leading to increased fixation duration on the target.

Consistent with previous evidence of plausibility preview benefit (Schotter & Jia, 2016; Yang et al., 2014) the effect did not extend to total fixation duration because the early benefit for plausible previews was counteracted by a higher rate of regressions to the target after a plausible than an implausible preview, even though the plausible previews were always compatible with the post-target text. Thus, plausible previews yield a trade-off between shorter fixation durations during first-pass reading, but longer second-pass reading. The likely source of the late interference is that competition due to orthographic discrepancies between the preview and target words caused a “double-take”. For plausible previews this was observed on second-pass reading because the forward saccade from the target location, planned on the basis of preview processing, was not cancelled by post-lexical processing failure. In these cases, a regression back to the target word resulted in the target word replacing the preview that had originally been encoded in the representation of sentence meaning, accounting for the lack of plausibility effect for trials with a regression in the animal detection task in Experiment 1A.

While the sequence of events described above is broadly compatible with the processing assumptions of E-Z Reader, simulations suggest that the conditions required to observe plausibility preview effects occur relatively rarely. Schotter et al.’s (2014) simulations estimated that full lexical identification of parafoveal previews occurred on only 8% of trials, and the likelihood of completing the postlexical processing required to generate contextually-based semantic preview effects would be even lower. The rarity of the sequence of events necessary to observe parafoveal semantic processing in E-Z Reader contrasts with the robust effect of preview plausibility observed by Veldre and Andrews (2016b) and in the present data. The mechanisms

accounting for postlexical effects in E-Z Reader are relatively rudimentary and unexplored. Plausibility preview effects provide a novel source of evidence that will contribute to further refinement of this, and other, eye movement models.

The other critical finding is that, in tandem with plausibility preview effects, the results also showed a robust effect of the orthographic similarity of the preview and target. We found no evidence of the interaction between plausibility and relatedness on eye movements, or increased animal detection accuracy for target words preceded by orthographically similar previews, that were expected if plausibility effects were due to misperception of the preview as the orthographically similar target. This suggests that the source of the relatedness effect lies in sublexical processing of words in the parafovea that occurs independently of any subsequent preview processing required to yield a plausibility effect. This is consistent with evidence that orthographic preview benefit is due to abstract letter codes shared by the preview and target that are activated early in the processing of a parafoveal preview, regardless of the preview's lexical status (see Schotter et al., 2012). The present data indicate that, independently of its contextual fit, activation of the sublexical components of the preview enhances the speed of processing the target word when it replaces the preview, to yield an orthographic preview benefit.

The significant orthographic relatedness effect in the present data contrasts with the lack of semantic relatedness effect in our earlier work (Veldre & Andrews, 2016b). This is consistent with the sublexical integration account described above. The lack of orthographic overlap between semantically related previews and targets in the earlier study meant there was no benefit to target word identification relative to a semantically unrelated (but equally orthographically dissimilar) preview.

Overall, the present study provides evidence of two distinct mechanisms underlying preview benefits. The plausibility preview effect indicates that parafoveal processing is sensitive to the contextual compatibility of the word in the sentence. However, readers also benefit from the orthographic information extracted from a parafoveal preview due to integration of sublexical features with the target word.

REFERENCES

- Abbott, M.J., & Staub, A. (2015). The effect of plausibility on eye movements in reading: Testing E-Z Reader's null predictions. *Journal of Memory and Language, 85*, 76-87.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1-48.
- Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 166-190.
- 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., Schotter, E.R., & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit in reading revisited. *Psychonomic Bulletin & Review, 21*, 1067-1072.
- 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., & Jia, A. (2016). Semantic and plausibility preview benefit effects in English: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

- 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.
- Slattery, T.J. (2009). Word misperception, the neighbor frequency effect and the role of sentence context: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 1969-1975.
- Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 57, 421-457.
- Veldre, A., & Andrews, S. (2016a). Semantic preview benefit in English: Individual differences in the extraction and use of parafoveal semantic information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Veldre, A., & Andrews, S. (2016b). Is semantic preview benefit due to relatedness or plausibility? *Journal of Experimental Psychology: Human Perception and Performance*.
- Wotschak, C., & Kliegl, R. (2013). Reading strategy modulates parafoveal-on-foveal effects in sentence reading. *The Quarterly Journal of Experimental Psychology*, 66, 548-562.
- 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.