Edinburgh Research Explorer

Readers detect an low-level phonological violation between two parafoveal words

1

2

3

4

Readers Detect an Low-Level Phonological Violation between Two Parafoveal Words

Michael G. Cutter[a,b], Andrea E. Martin[c,d], Patrick Sturt[a]

[a] The University of Edinburgh

[b] The University of Nottingham

[c] Max Planck Institute for Psycholinguistics

[d] Donders Centre for Cognitive Neuroimaging, Radboud University

Word count: 14,287

12

13

14

15

Author Note

Michael G. Cutter, Department of Psychology, University of Edinburgh, Edinburgh, Scotland & School of Psychology, University of Nottingham, Nottingham, United Kingdom; Patrick Sturt, Department of Psychology, University of Edinburgh, Edinburgh, Scotland; Andrea E. Martin, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands & Donders Centre for Cognitive Neuroimaging, Radboud University, Nijmegen, The Netherlands.

Correspondence regarding this article should be addressed to Michael G. Cutter, School of Psychology, University of Nottingham, University Park Campus, Nottingham, NG7 2RD. E-mail: michael.cutter@nottingham.ac.uk.

28

1   Abstract

2   In two eye-tracking studies we investigated whether readers can detect a violation of the

3   phonological-grammatical convention for the indefinite article *an* to be followed by a word

4   beginning with a vowel when these two words appear in the parafovea. Across two

5   experiments participants read sentences in which the word *an* was followed by a parafoveal

6   preview that was either correct (e.g. *Icelandic*), incorrect and represented a phonological

7   violation (e.g. *Mongolian*), or incorrect without representing a phonological violation (e.g.

8   *Ethiopian*), with this parafoveal preview changing to the target word as participants made a

9   saccade into the space preceding *an*. Our data suggests that participants detected the

10   phonological violation while the target word was still two words to the right of fixation, with

11   participants making more regressions from the previewed word and having longer go-past

12   times on this word when they received a violation preview as opposed to a non-violation

13   preview. We argue that participants were attempting to perform aspects of sentence

14   integration on the basis of low-level orthographic information from the previewed word.

15   *Keywords*; parafoveal processing; *n*+2 preview benefit; phonological processing;

16   contextual fit effects; reading.

1   **1. Introduction**

2   It has long been known that when processing written language, readers do not just

3   make use of visual information from the currently fixated word (e.g., see McConkie &

4   Rayner, 1975; Rayner, 1975). Rather, they also process at least one upcoming word

5   parafoveally, prior to direct fixation. In recent years there has been an increased focus on

6   both whether a word's contextual fit becomes apparent on the basis of parafoveal processing

7   (e.g. Angele & Rayner, 2013; Brothers & Traxler, 2016; Schotter & Jia, 2016; Veldre &

8   Andrews, 2016, 2018, Wakeford & Murray, 2013) and whether readers extract information

9   from a word further into the parafovea during reading (i.e. from word *n*+2; see Angele &

10   Rayner, 2011; Angele, Slattery, Yang, Kliegl, & Rayner, 2008; Cutter, Drieghe, &

11   Liversedge, 2014; Cutter, Drieghe, & Liversedge, 2017; Kliegl, Risse, & Laubrock, 2007;

12   Radach, Inhoff, Glover & Vorstius, 2013; Rayner, Juhasz, & Brown, 2007; Risse & Kliegl,

13   2012). In the current paper we examine these two issues simultaneously, by examining

14   whether participants' eye movement behaviour is influenced by a low-level phono-

15   grammatical violation between word *n*+1 and word *n*+2 in the parafovea, such that word *n*+1

16   was always the indefinite article in the form *an*, while word *n*+2 was sometimes manipulated

17   to begin with a consonant (e.g. *Spanish*).

18   **1.1. Parafoveal processing of word *n*+1**

19   A vast body of research has made use of the boundary paradigm (Rayner, 1975) to

20   investigate the extent to which people parafoveally process the word directly after fixation

21   (i.e. word *n*+1). In the boundary paradigm readers are given a preview of an upcoming word

22   prior to direct fixation which is either correct, or incorrect with a varying level of overlap to

23   the target word. As a reader's eyes cross an invisible boundary (typically located at the start

24   of the space before the previewed word) the preview changes into the actual target word. By

1   examining whether fixations on the target word are shorter following correct rather than

2   incorrect previews it is possible to determine whether readers have partially processed this

3   word parafoveally. A substantial number of studies have shown this is the case (see Cutter,

4   Drieghe, & Liversedge, 2015 for a review and Vasilev & Angele, 2017, for a Bayesian meta-

5   analysis). More interestingly, it is possible to determine what types of information readers

6   extract from parafoveal words by comparing reading times on the word dependent upon

7   whether the preview shares certain characteristics with the target word; for example, do

8   readers fixate the target word for less time given a preview sharing phonology with the target

9   word (e.g. *beech* as a preview for *beach*) than a preview that does not share this information

10   (e.g. *bench*)?

11   In recent years, starting with Schotter (2013), our understanding of parafoveal

12   processing has been greatly altered. Until recently it was generally assumed that readers only

13   extracted relatively low-level information from a parafoveal word, such as orthographic (e.g.

14   Briihl & Inhoff, 1995; Hyönä, Bertram, & Pollatsek, 2004; Johnson, Perea, & Rayner, 2007)

15   and phonological codes (e.g. Ashby & Rayner, 2004; Pollatsek, Lesch, Morris, & Rayner,

16   1992).[1] One way these effects were typically interpreted was in the framework of a trans-

17   saccadic activation account,[2] whereby readers would activate a candidate set of words on the

18   basis of those words sharing abstract characteristics with the parafoveal preview (e.g., see

19   Cutter, Drieghe, & Liversedge, 2015). As such, target words would typically be easier to

20   identify when preceded by a related preview than an unrelated preview.

---

[1] It should be noted that there is currently a controversy emerging over whether phonological preview effects are as large as previously assumed (Vasilev, Yates, & Slattery, 2019). For now we proceed under the assumption that these effects are real, although will return to this issue later due to its relevance to our own manipulation.

[2] This account has actually more typically been referred to as trans-saccadic integration, rather than activation. However, in the current paper we adopt the term activation so as to avoid confusion with the idea of people integrating an identified word into the sentence representation, another issue discussed in the current manuscript.

1   However, recent work examining whether participants process semantic information

2   from a parafoveal word (e.g. Hohenstein & Kliegl, 2014; Rayner & Schotter, 2014; Schotter,

3   2013), and, furthermore, whether readers are sensitive to a previewed word's fit into the

4   preceding sentence context (Schotter & Jia, 2016;Veldre & Andrews, 2016; Veldre &

5   Andrews, 2018; Wakeford & Murray, 2013; Yang, Li, Wang, Slattery, & Rayner, 2014) has

6   necessitated a reconsideration of the mechanisms underlying preview effects. In studies

7   assessing whether participants are sensitive to a word's fit in the sentence, readers have been

8   given previews of an upcoming word which was either plausible within the sentence context

9   (e.g. *modest* as a preview for *psycho* in *Melanie thought the man was really psycho…*) or

10   implausible within the sentence context (e.g. *circus*). These studies have typically found

11   longer fixation durations on the target word given an implausible preview than a plausible

12   preview. Furthermore, Veldre and Andrews (2018) found that these effects were graded with

13   the level of the unacceptability of the preview in the sentence context, such that fixations

14   would be longer on the target word given a preview that violated both semantic and syntactic

15   constraints of the sentence context (e.g. *begin* as a preview of *stool* in *She eventually found a*

16   *spare stool…*) as opposed to a preview which violated just semantic constraints (e.g. *uncle*) or

17   just syntactic constraints (e.g. *landed* as a preview of *refuel* in *Her plane will probably*

18   *refuel…*). These findings are problematic for a trans-saccadic activation account of preview

19   effects, since readers are gaining a preview benefit from a preview that should not be

20   activating the target word to a greater extent than another preview.

21   As well as finding evidence for preview effects driven by the syntactic acceptability

22   and plausibility of the previewed word in the sentence, recent studies have also observed such

23   effects upon the probability of readers skipping a word. In their study, Veldre and Andrews

24   (2018) found that participants were less likely to skip a word which was either implausible or

25   a poor syntactic fit in their sentences than words that fit into the sentence well. Trends in this

1   direction were also present in earlier work by Veldre and Andrews (2016) and Schotter and

2   Jia (2016). Furthermore, Brothers and Traxler (2017) found that a word of the incorrect

3   syntactic class is less likely to be skipped than a word of an appropriate syntactic class.

4   Finally, Cutter, Martin, and Sturt (in press) recently found evidence suggesting that

5   predictability may exert less of an effect on the skipping of words in syntactically illegal

6   positions than legal positions. It should be noted that these effects of contextual acceptability

7   upon word skipping are not strong enough to override the influence of all other factors, such

8   that readers will still skip a high-frequency syntactically illegal word more than a low-

9   frequency syntactically illegal word (Angele, Laishley, Liversedge, & Rayner, 2014; Veldre,

10 Recihle, Wong, & Andrews, 2020).

11   It should be clear from the above discussion that there is an increasing level of

12 evidence that readers are sensitive to the contextual fit of a parafoveal word within the

13 sentence context, both in terms of preview effects and word skipping. We will defer

14 discussing theoretical explanations of these effects until further below. In our own study we

15 investigate whether similar (but distinct) effects are observed for word *n*+2 in the parafovea.

16 As such, it is necessary to briefly survey the existing literature of these effects.

17 **1.2. Word *n*+2 preview effects**

18   A great deal of recent research has focussed upon the extent to which word *n*+2 is

19 processed prior to direct fixation and the factors affecting the extraction of information from

20 word *n*+2 (e.g. Angele & Rayner, 2011; Angele, Slattery, Yang, Kliegl, & Rayner, 2008;

21 Cutter, Drieghe, & Liversedge, 2014; Cutter, Drieghe, & Liversedge, 2017; Kliegl, Risse, &

22 Laubrock, 2007; Radach, Inhoff, Glover & Vorstius, 2013; Rayner, Juhasz, & Brown, 2007;

23 Risse & Kliegl, 2012). Studies of this issue initially presented an unclear pattern of results,

24 with some studies finding reliable *n*+2 preview effects, and others failing to find these

1   effects. A recent Bayesian meta-analysis by Vasilev and Angele (2017) clarified the extent to

2   which word $n+2$ is processed parafoveally. There are several findings from this Bayesian

3   meta-analysis worth focusing on. First, it was found that the average size of the $n+2$ preview

4   effect for studies using an alphabetic language was 3.9ms in first fixation duration and 1.8ms

5   in gaze duration on word $n+2$ itself, with effects of 6.7ms and 7.4ms on word $n+1$. The effect

6   on fixations on $n+2$ was 2.3ms and 6.9ms when analyses were restricted to studies using a

7   three-letter word $n+1$, as opposed to longer words. Thus, the average effect size across

8   studies is fairly small, suggesting that parafoveal processing of word $n+2$ is fairly minimal.

9   As well as presenting average effect sizes, Vasilev and Angele presented 95% credible

10   intervals on the effect size across all studies, and across each individual study they examined.

11   From these credible intervals, they also calculated the probability that the $n+2$ preview effect

12   was greater than a certain magnitude. It is noteworthy that neither the credible interval for all

13   studies combined nor for any individual study excluded an effect size of 0ms. Furthermore,

14   the probability that the effect size in each measure was greater than 1ms was only .70 and .53

15   for first fixation duration and gaze duration.

16   It should be clear from Vasilev and Angele's (2017) meta-analysis that $n+2$ preview

17   effects are fairly modest in size, if they exist at all. It is also worth noting that these $n+2$

18   preview effects have generally been assessed by comparing conditions in which participants

19   received a valid preview of word $n+2$ with conditions in which they received a preview

20   consisting of a random letter string. In their meta-analysis Vasilev and Angele also examined

21   the size of $n+1$ preview effects as a function of the type of parafoveal mask used, finding that

22   for word $n+1$ these random letter masks tended to result in longer fixations on the target word

23   than almost any other preview type (e.g. unrelated words; semantically, phonologically, or

24   orthographically related words). As such, $n+2$ effects are generally very small despite

25   researchers almost exclusively using false previews that should result in greater disruption

1   than any other type of preview. Furthermore, it could well be the case that such small *n*+2

2   preview effects in prior studies could arise from methodological issues as opposed to genuine

3   lexical processing of the parafoveal word. For example, Angele, Slattery, and Rayner (2016)

4   demonstrated that readers are more likely to consciously detect that a word has changed

5   during a saccade when the preview was non-word-like rather than word-like, and,

6   furthermore, that this detection of the change resulted in greater preview effects. As such,

7   even the effects reported in Vasilev and Angele's meta-analysis may exaggerate the extent to

8   which *n*+2 is processed parafoveally. It could be argued that this makes it highly unlikely

9   that participants should process word *n*+2 to an extent where its fit in the sentence is

10   assessed, and that any effect of our manipulation would be highly surprising.

11   Two prior studies have examined the effect of the contextual fit of word *n*+2 upon

12   preview effects, albeit in a different way to our own study, and the studies discussed above

13   examining word *n*+1 preview effects. The first of these studies was conducted by Radach,

14   Inhoff, Glover, and Vorstius (2013), who presented subjects with sentences (e.g. *Ashley*

15   *quickly vacuumed the…*) that either included a highly predictable target word (e.g. *carpet*) or

16   an unpredictable target word (e.g. *stairs*). They used the boundary paradigm to manipulate

17   the preview of this word until readers made a saccade beyond *vacuum*, with this preview

18   either being a non-word (e.g. *cwoyok*) or the word of the opposite predictability (e.g. *carpet*

19   as a preview of *stairs*), finding that reading times were longer on word *n*+2 when a highly

20   predictable target was previewed with an unpredictable word or non-word mask, while

21   reading times on an unpredictable target were longer when replaced with a highly predictable

22   preview but not a non-word mask. These findings suggest that readers process word *n*+2 to a

23   greater extent when they have a strong expectation of this word's identity, with subversions

24   of these expectations leading to processing difficulty. Similarly, Cutter, Drieghe, and

25   Liversedge (2014) found a larger than typical *n*+2 preview effect when word *n*+2 was the

1   second constituent of a lexicalised spaced compound (e.g. *bear* in *teddy bear*) and was

2   replaced with a non-word (e.g. *hocu*), and thus readers had a very strong expectation for the

3   identity of this word. Together these studies suggest greater *n*+2 preview effects are observed

4   when the preview of *n*+2 does not meet the expectations from the prior sentence context or

5   word; however, these expectations have generally been for one specific word in prior studies,

6   as opposed to just any word which meets one particular constraint.

7   **1.3. The current study**

8   In summary, readers process the first word to the right of fixation (i.e. word *n*+1)

9   substantially while it is still in the parafovea, to the extent that it is possible for readers to

10   assess that word's fit into the sentential context. Furthermore, these effects of sentential fit

11   are not based upon any particular word being predictable, but rather on the extent to which a

12   parafoveal word is consistent with a number of different constraints (i.e. syntactic and

13   semantic). In contrast, readers seem to process very little information from word *n*+2, with

14   existing studies suggesting that for the most part readers only extract low-level graphemic

15   information from a word further into parafoveal vision. However, the effect of denying an

16   accurate preview of word *n*+2 does seem to be larger in cases when the target word is

17   particularly predictable (e.g. Cutter et al., 2014; Radach et al., 2013).

18   In the current study we examine whether readers process word *n*+2 in the parafovea to

19   the extent that its compatibility with word *n*+1 is assessed at a phonological level. In our

20   study participants read sentences in which the indefinite article *an* was followed by a proper

21   noun (e.g. *He thought losing against an Icelandic team was very embarrassing*). We used the

22   boundary paradigm to manipulate the preview of the proper noun until readers made a

23   saccade beyond *an*, such that the parafoveal preview was visible as word *n*+2. In Experiment

24   1 we merely manipulated the preview to be correct (i.e. *Icelandic*) or incorrect and a violation

1   of English phonological conventions, due to starting with a consonant (e.g. *Mongolian*). In

2   Experiment 2 we included a third preview condition, in which the preview was incorrect but

3   did not represent a violation of English phonological conventions (e.g. *Ethiopian*).[3]

4   To the best of our knowledge, only one prior study has examined eye movement

5   behaviour in relation to this phono-grammatical rule (Slattery, Pollatsek, & Rayner, 2006). In

6   Slattery et al.'s study acronyms were used in which phonology was consistent with

7   orthography (e.g. *CIA* is pronounced with an initial consonant; *IRS* is pronounced with a

8   vowel) or inconsistent with orthography (e.g. *FBI* is pronounced with an initial vowel; *USA* is

9   pronounced with an initial consonant). These acronyms followed either *an* or *a*, with the form

10   of the article causing either a phonological violation (e.g. *an USA*, *a FBI*), an orthographic

11   violation (e.g. *an FBI*, *a USA)*, an orthographic and phonological violation (e.g. *an CIA*, *a*

12   *IRS*) or no violation (e.g. *an IRS*, *a CIA*) between the two words. The effects of a

13   phonological violation appeared in early fixation measures on the acronym (e.g. first fixation

14   and gaze duration) while effects of an orthographic violation only appeared in later measures

15   (e.g. go-past time, total time).Slattery et al. took this as evidence that readers rapidly detected

16   a mismatch between the phonology of the acronym and preceding article. It should be noted

17   that these findings may be specific to acronyms, due to the ambiguity between orthography

18   and phonology. Indeed, Slattery et al. proposed that their finding may have been due to

19   readers using the article to determine how to pronounce the acronym, leading to a cost when

20   this cue was incorrect. Thus, while readers did indeed detect a violation early in this study, it

21   is unclear whether the processes that resulted in this effect could be considered typical of the

22   reading system, or if they only occur in acronyms in which there is a large level of ambiguity

---

[3] It might seem strange that we did not simply include this condition in Experiment 1. Frankly, given the size of prior *n*+2 preview effects, we did not expect it to be possible to observe reliably different *n*+2 preview effects for different *n*+2 preview types, and instead intended to focus upon the skipping of *an* depending on the following word's initial character. However, upon examination of the reading time data in Experiment 1 we concluded that it would be worth conducting a further experiment, including this third condition.

1   regarding pronunciation. Furthermore, the violations in this study were presented foveally,

2   and so it is impossible to determine whether they were detected much earlier, in the

3   parafovea, or only once the target acronyms were directly fixated.

4   The phonological convention for the article *an* to be followed by a word beginning

5   with a vowel as opposed to a consonant represents an interesting and potentially understudied

6   way in which to consider a word's sentential fit.[4] In the case of the studies discussed above a

7   word's sentential fit was generally determined by constraints generated on the basis of the

8   sentence's global meaning and structure. On the other hand, using our manipulation, word

9   *n*+2's lack of fit into the sentence is entirely determined by only the preceding word.

10 Secondly, in order for readers to determine whether a word is implausible or violates

11 syntactic expectations it is necessary for readers to reach a reasonably advanced stage of

12 lexically processing this word, since they must retrieve its syntactic class and meaning from

13 the mental lexicon. With our manipulation this is not the case; readers can determine a

14 violation of the grammatical convention purely on the basis of the orthographic/phonological

15 form of the following noun. Due to these two factors, it might be possible for participants to

16 detect a violation of this phonological convention rapidly enough that it appears as an *n*+2

17 preview effect in the eye-movement record. Such an effect would have important

18 implications for the extent to which people process information in the parafovea, and

19 furthermore, the time-course across which they attempt to integrate words into a larger

20 sentence representation. Both of these issues hold relevance for current computational models

21 of eye-movement control. In the current paper we focus on the E-Z Reader model (e.g.

22 Reichle, Pollatsek, & Rayner, 2006; Reichle, Warren, & McConnell, 2009) due to the fact

---

[4] To the best of our knowledge, beyond the work of Slattery, Pollatsek, and Rayner (2006) the relationship between *an* and the initial letter of the following word has primarily been examined in relation to the issue of form prediction in electrophysiological studies of language processing (e.g. see DeLong, Urbach, & Kutas, 2005; Ito, Martin, & Nieuwland, 2016).

1   that it contains parameters relating to both parafoveal processing and the integration of

2   identified words into a sentence representation, something in which it is currently unique.

3   **1.4. The E-Z Reader model of eye movement control**

4   An important framework in which parafoveal processing and our study can be viewed

5   is the E-Z Reader model of eye movement control. The E-Z Reader model explains eye

6   movement behaviour using a relatively parsimonious set of parameters, focusing on how

7   readers process individual words, how attention is directed to each visually available word,

8   how readers make and execute saccade programming decisions, and how words are

9   integrated into a higher level sentence representation. Within E-Z Reader it is assumed that

10 readers lexically process one word at a time, with the processing of each word split into two

11 phases (L_1 and L_2). The durations of these stages are modulated by the frequency and

12 predictability of each word. Upon completing L_1, readers begin programming a saccade to the

13 next word, while continuing to process the fixated word in the L_2 stage. Saccade

14 programming is divided into two stages, with a labile stage (during which the saccade can be

15 cancelled) lasting ~125ms, and a non-labile stage (during which the saccade cannot be

16 cancelled) lasting ~25ms. During saccadic programming, several other processes are

17 ongoing. First, the L_2 stage of processing the fixated word will be completed, resulting in

18 access to semantic information about the target word. At this point, the fixated word is

19 integrated into the sentence representation while readers begin processing the parafoveal

20 word in the L_1 stage; thus, according to E-Z Reader integration should only occur once words

21 are fully identified. Two things can happen at this point. First, the labile stage of

22 programming a saccade to the parafoveal word may complete prior to L_1 for this word; in this

23 case, the parafoveal word is directly fixated. Alternatively, the L_1 stage can complete for the

24 parafoveal word prior to the saccade program becoming non-labile; at this point the saccade

25 to word $n+1$ is cancelled, and a saccade is instead programmed to word $n+2$ – in other words,

readers will skip the parafoveal word. Furthermore, if the processing of word $n+1$ completes

from word n then attention will move to word $n+2$ prior to a saccade being executed. Once

readers fully program a saccade its execution takes approximately 25ms to complete, and

there is then a further 50ms lag before new visual information propagates to the brain;

consequently, readers continue processing the visual information extracted during a particular

fixation for ~75ms after the eye leaves that position.

E-Z Reader explains basic preview effects through the fact that attention shifts to the

parafoveal word before a saccade programme is executed, and before visual information from

the post-boundary fixation becomes available for processing. While the model does not make

any explicit predictions about exactly why certain types of incorrect previews result in shorter

fixation durations than others, the assumption has generally been based around a trans-

saccadic activation account. As mentioned, such an account cannot explain effects of the

contextual fit of a previewed word on subsequent fixations on an unrelated target word.

Instead, these effects may require an alternative mechanism, which has been put forward by

Schotter and colleagues in the form of *forced fixations* (Schotter, 2018; Schotter &

Leinenger, 2016; Schotter, Leinenger, & von der Malsburg, 2018; Schotter, von der

Malsburg, & Leinenger, 2019). Forced fixations refer to instances in which readers complete

the L_1 stage of processing a word parafoveally, but after the saccade program towards this

word becomes non-labile. In such cases it would make sense for readers to cancel the saccade

to word $n+1$, and instead program one to word $n+2$. Unfortunately, this is not possible.

Consequently, readers are forced to fixate word $n+1$. Schotter and colleagues argue that in

these cases readers will pre-initiate a saccade from word $n+1$ to word $n+2$ prior to the eye

landing on word $n+1$. In such cases the fixation on word $n+1$ will be entirely driven by the

characteristics of the parafoveally previewed word as opposed to the foveally viewed target

1   word, with readers not extracting new information from the now fixated word on the

2   assumption that they have already identified it.

3   The contextual fit preview effects discussed above can be explained on the basis of

4   forced fixations and, in our view, E-Z Reader's sentential integration parameters. In cases

5   when readers make a forced fixation due to identifying a parafoveal preview of word $n+1$ as

6   word $n+1$ itself, they should, during the course of the forced fixation, attempt to integrate this

7   previewed word into their sentential representation. In cases when this previewed word is

8   plausible in the sentential context this integration should be unproblematic, and as such the

9   pre-programmed saccade away from this word should not be cancelled, resulting in relatively

10   short fixations on the word. On the other hand, when the previewed word is either

11   implausible or syntactically illegal in the sentence context, integrating that word into the

12   sentence representation is likely to rapidly fail, resulting in the cancellation of the saccade

13   away from this word, and further time spent fixating on the word, as well as potential

14   regressions to earlier portions of the text. Hypothetically, this can explain contextual fit

15   preview effects, although simulations with E-Z Reader may be necessary to assess this. It is

16   worth noting that efforts have also been made to explain contextual fit preview effects

17   through the idea that plausible words are easier to lexically process than implausible words,

18   resulting in a greater number of forced fixations on plausible words, with these fixation being

19   shorter than normal fixations. Due to this, average reading times are shorter on plausible than

20   implausible words, due to lexical processing rather than integration processes.

21   E-Z Reader is able to explain $n+2$ preview effects through the idea that lexical

22   processing will sometimes complete on word $n+1$ prior to the eye making a saccade across

23   the invisible boundary, and the time it then takes for visual information from a post-boundary

24   fixation to then propagate through the visual system. While this should be a relatively rare

25   occurrence, it does still happen often enough to explain some level of $n+2$ preview effects.

1   In terms of how often both $n+2$ processing and forced fixations should occur

2   according to E-Z Reader, a set of simulations run by Schotter, Reichle, and Rayner (2014) are

3   informative. Schotter et al. tested how often readers should commence the L_2 stage of

4   processing word $n+1$ and $n+2$ parafoveally, and how often readers should process word $n+2$

5   at all. These simulations suggested that participants should proceed to the L_2 stage of

6   parafoveally processing word $n+1$ in 8% of instances,[5] that they should process word $n+2$ in

7   about 20% of trials, and that they should never proceed to the L_2 stage of parafoveally

8   processing word $n+2$. While these simulations do not give exact estimates of the prevalence

9   of forced fixations, they do give an upper estimate, in that such fixations should not occur

10   unless participants proceed beyond L_1 in processing the parafoveal word. One thing that

11   drops out of these simulations is that forced fixations should not occur for word $n+2$ and that,

12   as such, contextual fit preview effects should not be observed for these words.

13   Returning to our own manipulation, the above consideration of E-Z Reader should

14   make several things clear. First of all, due to simulations suggesting that readers should never

15   proceed beyond the L_1 stage of processing word $n+2$ any effect of our manipulation would be

16   highly surprising in current frameworks, due to the lack of potential forced fixations on word

17   $n+2$. Secondly, it is interesting to consider more generally at what stage of processing within

18   the model one would expect participants to detect the type of phonological violation used in

19   the current study. Arguably, the most logical point for this to occur would be at the point in

20   which readers attempt to integrate the word into the representation of the sentence. However,

21   as mentioned above it is also possible for readers to detect our violation much earlier than

22   typical syntactic or semantic violations, with ours being based upon a low-level characteristic

---

[5] This outcome was obtained when the length of word $n+1$ was set to five letters and its frequency per million to 2019. An alternative simulation with frequency set to one per million and length to eight letters resulted in estimate of participants proceeding to L_2 on 2% of trials; we report the simulation we do due to it seeming to represent more realistic lexical characteristics.

1   of the target word that does not require substantial lexical processing. As such, whether or not

2   we observe an effect of our manipulation will be informative as to whether readers delay the

3   phonological integration of an article with the following word until after lexical processing is

4   completed for both words, or if such effects occur in an earlier, pre-lexical integration stage.

5   We return to this issue in the General Discussion.

6   **2. Experiment 1**

7   In Experiment 1 we presented participants with sentences in which the indefinite

8   article *an* was followed by a proper noun (e.g. *English*). We used the boundary paradigm to

9   present readers with either a correct preview of this noun, or an incorrect preview that began

10   with a consonant (e.g. *Spanish*) up until readers made a saccade into the space preceding *an*.

11   Hence, in our design word *n*+1 was always *an*, and word *n*+2 was always a proper noun

12   which was either compatible or incompatible with the preceding article until readers had

13   begun to execute a saccade to either fixate or skip *an*. We chose to use proper nouns as our

14   target word in order to increase the salience of the information representing the grammatical

15   violation (i.e., the identity of the first letter of the word following *an*).

16   The main issue we were interested in examining in Experiment 1 was whether the

17   low-level violation of an English phonological convention resulted in the reduced skipping of

18   the article *an*. As discussed above, several researchers (e.g. Brothers & Traxler, 2016; Veldre

19   & Andrews, 2018) have recently found that word *n*+1 is less likely to be skipped when it is

20   either syntactically illegal or implausible in the preceding sentence context. While readers

21   should be no less likely to begin programming a saccade to skip *an* when it is followed by a

22   consonant-initial word, it may be the case that upon shifting attention to word *n*+2 and

23   detecting a violation this saccade may be cancelled. It actually seems fairly unlikely that

24   participants would become aware of our manipulation quickly enough for this saccade to be

1 cancelled regularly enough to result in reliable effects upon skipping. None-the-less, such a

2 finding seems potentially interesting enough for it to be worth testing this possibility.

3 More generally, we were also interested in testing whether our manipulation resulted

4 in *n+2* preview effects, and whether these effects are larger than in prior studies. Given the

5 use of a shorter word *n+1* than in other studies, as well as a capitalised word *n+2*, it seems

6 likely that we should observe such effects.

7 **2.1 Method.**

8 **2.1.1. Participants.**

9 Sixty native speakers of English (39 females; mean age = 21.50) with normal or

10 corrected to normal vision participated for £6.

11 **2.1.2. Apparatus.**

12 Eye movements were tracked using a tower-mounted SR Research Eyelink 1000.

13 Sentences were displayed on a single line of a CRT monitor running at 120hz. Viewing

14 distance was 70cm, with 1° of visual angle containing 2.7 characters of courier new font.

15 **2.1.3. Materials and design.**

16 We constructed thirty sentences containing the indefinite article in the form *an*

17 followed by a proper noun beginning with a vowel (e.g. *English*). Each proper noun was

18 paired with an alternative proper noun beginning with a consonant sound (e.g. *Spanish*),

19 rendering *an* the inappropriate form of the indefinite article. The two nouns were matched for

20 length on an item-by-item basis. In addition, paired t-tests showed no significance differences

21 between the two nouns in mean log bigram frequency ($t(29)$=-1.24, $p = 0.22$; consonant-

22 initial mean = 2.03 , vowel-initial mean = 2.16), mean log trigram frequency ($t(29)$=0.04, p =

1   0.97; consonant-initial mean = 0.83, vowel-initial mean = 0.83). Stimuli for both Experiment

2   1 and Experiment 2 can be found at https://osf.io/5sq8g/.

3   The boundary paradigm was used to present participants with the vowel-initial noun

4   on 50% of trials and the consonant-initial noun on 50% of trials. The boundary was placed at

5   the beginning of the space preceding *an*. As the eye crossed the invisible boundary the noun

6   would always change to the vowel-initial version, such that participants always read a legal

7   sentence.

8   **2.1.4. Procedure.**

9   Participants provided informed consent upon arrival. Readers' heads were stabilized

10   using a head rest. A three-point horizontal calibration grid was used, with the calibration

11   being accepted if average error was below 0.30 degrees. Before each trial, there was a drift

12   check in the centre of the screen, then a drift check on the left of the screen, and then a gaze

13   contingent box in the same position as the sentence's first character. Participants were

14   recalibrated if either drift check returned a value greater than 0.40 on two consecutive trials

15   or they failed to trigger the gaze contingent box. Participants pressed a button after reading

16   each sentence. Yes/no comprehension questions were presented following one third of the

17   sentences; participants answered using the left or right mouse button. The experiment lasted

18   approximately 30 minutes.

19   **2.2. Results.**

20   We analysed reading behaviour across two interest areas. The first was the word *an*,

21   for which we examined skipping probability only. The second interest area was the

22   previewed noun, on which we examined first fixation duration (i.e., the duration of the first

23   fixation on this word), gaze duration (i.e., the total time between first fixating the word and

24   making a saccade to another region), go-past time (i.e., the total time between first fixating

1   the word and exiting the word to the right), and the probability of participants making a first-

2   pass regression out. For fixation time measures, values more than 3 standard deviations from

3   the grand mean of the untransformed data were removed, resulting in a maximum of 28

4   observations being removed from the data. Outlier removal was performed on a measure by

5   measure basis, such that we did not necessarily remove a data point which was an outlier in

6   one measure for all other measures. Means for these measures are shown in Table 1.

7   Statistical analyses were conducted on log-transformed measures.

8   Prior to analysis we used SR-Research Data Viewer to 1) remove fixations above

9   800ms; 2) merge fixations below 80ms with fixations less than 0.5 degrees of visual angle

10   away; and 3) merge fixations below 40ms with fixations less than 1.25 degrees away. These

11   values represent the default settings within Data Viewer for fixation cleaning in reading

12   studies. 0.75% of fixations across the whole experimental session were affected by the

13   merging operation. We removed trials in which participants blinked on either the word *an* or

14   the critical noun. We also removed trials in which the display change triggered early for our

15   analysis of reading behaviour on *an* and the critical noun, and trials in which the display

16   change completed more than 10ms late for fixations on the previewed noun.[6] These

17   exclusions accounted for 11% of data for *an* and 13% of data for the previewed noun.

18   We constructed Bayesian linear mixed models to analyse our data. There are two

19   main reasons we took a Bayesian approach in the current study. First, one of our main points

20   of interest in Experiment 1 was the skipping of the article *an*. It seemed likely at the outset

21   that we may have to argue in favour of a null hypothesis for this measure, and as such an

22   analysis that allows us to do so was an obvious choice. The Bayesian LMMs had a prior of

---

[6] Adopting a stricter exclusion criteria and excluding any trials in which the display change occurred during a fixation made no difference to our pattern of effects, nor what was considered reliable and unreliable in either Experiment 1 or Experiment 2.

1   *Normal*(μ= 0, σ= 10) for the model intercept and *Normal*(0, 1) for the effect of preview type,

2   with a regularization of 2 on the covariance matrix of random effects. These priors are

3   weakly informative, and take account of the fact we analysed log-transformed data. Each

4   model was run with two chains of 2000 iterations each. The Rhat value for each parameter in

5   each model was equal to 1, indicating successful model convergence. For each analysed

6   measure we report a mean effect estimate, the upper and lower bounds of a 95% credible

7   interval, and the probability that the effect of our preview manipulation was greater than 0.

8   This output is displayed in Table 2. Data files and R Scripts for analysing them can be found

9   at https://osf.io/5sq8g/.

10   Our manipulation had very little effect on the skipping of *an*. On average, participants

11   would skip this word on 67% of trials given a vowel preview, and 68% of trials given a

12   consonant preview. The output from our model suggests that the probability of our effect

13   being larger than a value of zero is only .772; as such we have very little evidence that the

14   skipping of *an* was influenced by the initial letter of the following word. Furthermore, the

15   trend was actually in the opposite direction to our initial hypothesis.

16   Table 1

17   Mean (Standard Error) Skipping Probability of "an" and Fixation Times and Regression
18   Probabilities on Word *n*+2.

	Identity/vowel	Violation/consonant
An skipping	0.67 (0.02)	0.68 (0.02)
First Fixation Duration	225 (2.34)	235 (2.38)
Gaze Duration	275 (4.49)	289 (4.42)
Go-Past Time	321 (6.25)	343 (6.12)
Regressions Out	0.13 (0.01)	0.17 (0.01)

19

20   Our statistical models suggested that we had evidence in favour of an effect of our

21   preview manipulation upon fixation behaviour on the target noun, with mean effects

22   estimates of 9ms in first fixation duration, 16ms in gaze duration, 25ms in go-past time, and a

1 4% increase in the probability of participants making a regression back to earlier parts of the

2 text.

3 Table 2

4 Output from our Bayesian mixed effects models. Mean effects estimates (*b*), 95% confidence
5 intervals, and the probability of our effect being in the more likely direction are reported.
6 Log-units output by the model are reported, followed by values translated into fixation times
7 and probabilities in parentheses.

	Intercept	*B*	L95%	U95%	P(*b*>/<0)
An skipping	0.77 (0.68)	0.10 (2%)	-0.16 (-4%)	0.37 (7.4%)	0.772
FFD	5.40 (221)	0.04 (9)	0.01 (2)	0.08 (18)	0.994
GD	5.56 (260)	0.06 (16)	0.00 (0)	0.13 (36)	0.977
GP	5.69 (296)	0.08 (25)	0.01 (3)	0.15 (48)	0.991
Regs. Out	-2.00 (0.12)	0.34 (4%)	-0.01 (-0%)	0.70 (9%)	0.973

8

9 **2.3. Discussion.**

10   In Experiment 1 we tested whether participants a) extracted information from a

11 parafoveal preview of word *n*+2 during fixations on word *n* and b) whether the extraction of

12 this information affected the skipping of the intervening article *an* due to a grammatical

13 violation. We did this by presenting participants with previews of word *n*+2 which were

14 either correct, or which were incorrect and violated the convention for the indefinite article

15 *an* to be followed by a word beginning with a vowel sound. Readers were clearly extracting

16 some information from word *n*+2 in the parafovea, with our Bayesian models showing

17 preview effects across all measures examined on word *n*+2. In contrast, our manipulation had

18 minimal effects on the skipping of word *n*+1.

19   It is clearly not the case that our lack of skipping effect was due to participants failing

20 to extract information from word *n*+2 while still fixated on word *n*. Had this been the case

21 then our manipulation should also have had no effect on reading times on word *n*+2. Rather,

22 we suspect that word *n*+2 was simply processed too late for it to have any effect on the

23 programming of a saccade which was supposed to either skip or land on *an*. We will leave

1   any further discussion of this effect until our General Discussion, after we present additional

2   data relating to this issue in Experiment 2.

3   One surprising aspect of our data was the size of the $n+2$ preview effect on fixation

4   durations on word $n+2$ itself. Our effects were relatively large compared to the estimated

5   effect size from Vasilev and Angele's (2017) meta-analysis. Specifically, our effect was 9ms

6   in first fixation durations and 16ms in gaze durations, whereas when Vasilev and Angele

7   focused on studies using a shorter word $n+1$ they found effects of 2.3ms and 6.9ms in these

8   measures. Thus, there is presumably something about our manipulation that resulted in these

9   larger effects. We can think of three potential aspects of our stimuli and manipulation that

10   drove this effect.

11   First, we used a two-letter word $n+1$, which is shorter than word $n+1$ has typically

12   been in prior $n+2$ preview studies. Arguably, this should increase the probability of observing

13   $n+2$ preview effects due to the low processing load of this word,[7] and the fact that word $n+2$

14   will be in higher acuity vision and thus processed more efficiently. Second, we used a

15   capitalised word $n+2$ in our experiment– prior research (e.g. Rayner & Schotter, 2014) has

16   shown that readers gain a larger preview effect from word $n+1$ when it is capitalised, and it

17   would be unsurprising if the same was true of word $n+2$.

18   A final factor driving our larger preview effects, and the focus of Experiment 2, is the

19   possibility that participants were in fact processing word $n+2$ to an extent that the violation

20   preview was being detected by the reading system, leading to processing difficulties which

21   did not appear until direct fixation on word $n+2$. Unfortunately, it is not possible to assess

22   this possibility without data testing whether smaller $n+2$ effects are observed with a false

---

[7] It should be noted that this explanation relies on the same logic as findings showing that foveal load affects the size of $n+1$ preview effects, which recent evidence (Veldre & Andrews, 2018b) suggests may not be as robust of an effect as often assumed.

1   preview which does not violate English inter-word phonological conventions. As such, it was

2   necessary to run a further experiment.

3   **3. Experiment 2**

4   In order to further investigate our surprisingly large *n*+2 preview benefit from

5   Experiment 1 we conducted Experiment 2. Specifically, we were interested in examining

6   whether our large effect in Experiment 1 was in part due to the fact that the incorrect preview

7   violated the typical use of the indefinite article *an*. To do so we presented readers with

8   sentences in which they received either 1) an identity preview which followed the standard

9   English convention (e.g. *an Icelandic*), an incorrect preview which violated this convention

10   (e.g. *an Mongolian*), or an incorrect preview which followed standard grammatical

11   conventions (e.g. *an Ethiopian*). Once again an invisible boundary was set before the

12   indefinite article *an* such that the preview would change to the target word as readers' eyes

13   moved towards it.

14   As well as wanting to test whether readers gain differential preview effects to these

15   two preview types we wished to further investigate our (lack of) skipping effect from

16   Experiment 1. We suspect that the non-reliable 2% estimated difference we observed in this

17   experiment was most likely just noise, as opposed to a real effect which we lacked the power

18   to detect in Experiment 1. We will be able to use the data from Experiment 2 to test this

19   possibility in two ways. First of all, we can examine whether we observe a similar trend in

20   Experiment 2, and then combine the data from our two experiments to test for an effect with

21   considerably more statistical power than in either individual experiment.

22   **3.1. Method.**

23   **3.1.1. Participants.**

1   78 native speakers of English (61 female; mean age = 21.64) with normal or corrected

2   to normal vision participated for £6. An additional 6 participants were tested, but their data

3   was excluded due to them either noticing more than five display changes or poor tracking.

4   **3.1.2. Apparatus.**

5   Apparatus were identical to Experiment 1. A slightly smaller font was used to

6   accommodate items from an experiment run concurrently; 1° of visual angle contained 3.4

7   characters of courier new font.

8   **3.1.3. Materials and design.**

9   We constructed forty-two sentences containing the indefinite article in the form *an*

10   followed by a proper noun beginning with a vowel (e.g. *Icelandic*). Fourteen of these

11   sentences were identical to those used in Experiment 1. Each proper noun was paired with

12   two alternative proper nouns, one beginning with a consonant sound (e.g. *Mongolian*) and the

13   second beginning with a vowel sound (e.g. *Ethiopian*). The three words were matched for

14   length on an item-by-item basis.

15   In addition, we also matched the words on average on a number of variables. First, we

16   ensured that the two false previews did not differ significantly in terms of their level of

17   orthographic overlap with the target word (t(41)=0.23, p = 0.82). This was calculated in terms

18   of the Levenshtein distance between each preview word and the target word. We also ensured

19   that the incorrect vowel-initial preview never shared an initial letter with the target word.

20   There were also no significant differences between the three words in terms of mean log

21   trigram frequency (all p >.4), mean log bigram frequency (all p >.2), or log frequency per

22   million in the British National Corpus (all p >.5). Means for each of these variables can be

23   found in Table 3. It should be noted that a large number of vowel-initial proper nouns were

1   used as both correct previews and incorrect-vowel previews, due to the fact that there are a

2   limited number of such proper nouns to use.

3   The boundary paradigm was used to present participants with the vowel-initial

4   preview on one third of trials, the consonant-initial preview on one third of trials, and the

5   target word on the remaining third of trials. The boundary was placed at the beginning of the

6   space preceding *an*. As the eye crossed the invisible boundary the noun would always change

7   to the vowel-initial version, such that participants always read a legal sentence.

8   **3.1.4. Procedure.**

9   The procedure was identical to Experiment 1.

10   Table 3

11   Characteristics of the Stimuli Used in Experiment 2.

	Identity	Consonant	Vowel
Levenshthein Distance from target	0	6.14	6.11
Mean Log Trigram Frequency	0.85	0.91	0.82
Mean Log Bigram Frequency	1.98	1.97	2.07
Log Frequency Per Million (BNC)	1.59	1.72	1.61

12

13   **3.2. Results.**

14   We prepared our data for analysis in an identical way to Experiment 1, and analysed

15   the same measures in the same regions. 0.88% of fixations across the whole experimental

16   session were affected by the merging operation in Data Viewer. Exclusions due to blinks and

17   display change errors accounted for 17% of the data for *an*, and 20% of data for the

18   previewed noun. Outlier cleaning resulted in a maximum of 54 observations being removed

19   for any one measure. Means can be seen in Table 4, and model output in Table 5. The Rhat

20   value for each parameter in each model was equal to 1, indicating successful model

21   convergence.

3.2.1. Word *n*+1 skipping.

For skipping behaviour on the word *an* we used the default contrast structure in R, with the consonant preview set as the baseline. Thus, this model returned two contrasts, with one checking whether skipping differed in the identity preview condition relative to the consonant preview condition, and the second checking whether skipping differed in the vowel preview condition relative to the consonant preview condition. Numerically, there was a trend for participants to skip *an* more when it was followed by a word beginning with a consonant, similarly to Experiment 1. However, this effect was not reliable in either contrast in the model.

As mentioned above, it is possible to combine the data from the current experiment with the data from Experiment 1, in order to test whether the initial character of word *n*+2 affected the skipping of *an*. We constructed a Bayesian logistic mixed model with a fixed effect testing for the effect of whether the preview of *n*+2 started with a vowel or a consonant; we did not discriminate between cases in which word *n*+2 was a false preview beginning with a vowel and cases in which word *n*+2 was the correct preview beginning with a vowel. This model suggested that participants would skip *an* more when it was followed by a consonant-initial word than a vowel-initial word; however, while the credible interval was smaller than in either of the models examining the data for Experiment 1 and Experiment 2 separately, it still included some estimates that the effect went in the opposite direction. The probability that the effect went in the direction of increased skipping with a consonant preview was .911.

3.2.2. Word *n*+2 reading.

For reading behaviour on word *n*+2 we used the default contrast structure in R for all models, with our intercept set to be the vowel-initial incorrect preview. As such, our model

returned one comparison comparing trials in which participants received an identity preview

with an incorrect preview which did not violate phonological conventions (henceforth

referred to as the identity preview effect), and one contrast comparing trials in which

participants received an incorrect preview which did not violate grammatical conventions

with trials in which participants received an incorrect preview which did violate grammatical

conventions (henceforth referred to as violation preview effect).

There was a trend in all measures for an identity preview effect, such that participants

would read word *n*+2 faster given a correct preview. Our linear mixed models suggested that

we had clear evidence for this effect in go-past times, although the evidence in first fixation

durations and gaze durations was less clear. It should be noted that this does not constitute a

failure to replicate Experiment 1, since the relevant contrast here would be between the

identity preview and the consonant-initial preview.

More interestingly, there was also a trend for a violation preview effect in all

measures, such that participants would make shorter fixations on a word given an incorrect

vowel-initial preview than an incorrect constant-initial preview. This effect was clear in go-

past times and the probability of participants making a first pass regression out of word *n*+2,

suggesting that the effect was mainly driven by re-reading behaviour.

Table 4

Mean (Standard Error) Skipping of the Article *An*, and Reading Measures for Word *n*+2 in
each Preview Condition

	Identity	Vowel	Consonant
An skipping	0.64 (0.02)	0.64 (0.02)	0.67 (0.02)
FFD	221 (2.21)	226 (2.30)	233 (2.47)
GD	271 (4.21)	277 (4.15)	284 (4.48)
Go-Past	325 (6.42)	340 (6.67)	365 (7.22)
Regression out	0.16 (0.01)	0.17 (0.01)	0.23 (0.01)

1   Table 5

2   Output from our Bayesian Models for Experiment 2.

		Intercept	*B*	L95%	U95%	P(*b*>/<0)
An skipping	Identity Contrast	0.74 (68%)	-0.10 (-2%)	-0.33 (-8%)	0.10 (2%)	0.843
	Vowel Contrast	0.74 (68%)	-0.11 (-2%)	-0.32 (-7%)	0.11 (2%)	0.821
	Composite model	0.71 (67%)	-0.10 (-2%)	-0.26 (-6%)	0.05 (1%)	0.911
FFD	Identity Contrast	5.38 (217ms)	-0.02 (-4ms)	-0.05 (-11ms)	0.01 (2ms)	0.911
	Violation Contrast	5.38 (217ms)	0.03 (6ms)	-0.00 (-0ms)	0.06 (13ms)	0.951
GD	Identity Contrast	5.54 (255ms)	-0.03 (-8ms)	-0.07 (-17ms)	0.01 (3ms)	0.928
	Violation Contrast	5.54 (255ms)	0.01 (3ms)	-0.03 (-8ms)	0.05 (13ms)	0.648
GP	Identity Contrast	5.71 (302ms)	-0.05 (-15ms)	-0.10 (-28ms)	-0.00 (-0ms)	0.980
	Violation Contrast	5.71 (302ms)	0.05 (15ms)	0.01 (3ms)	0.10 (31ms)	0.988
Reg. out	Identity Contrast	-1.79 (14%)	-0.09 (-1%)	-0.41 (-4%)	0.22 (3%)	0.724
	Violation Contrast	-1.79 (14%)	0.40 (6%)	0.10 (1%)	0.69 (11%)	0.998

3

4   A final issue worth considering is the extent to which our effects were dependent

5   upon readers skipping the article *an*. In order to assess this we repeated our analysis, looking

6   only at trials in which readers did in fact skip *an*. In this sub-set of our data our effects were

7   numerically larger, such that reading times on word *n*+2 were 11ms, 16ms, and 33ms longer

8   given a violation preview as opposed to an incorrect vowel preview in first fixations, gaze

9   durations, and go-past times, respectively. Our Bayesian models suggested evidence in favour

10   of these effects in both first fixation duration (P (*b*>0) = 0.975) and go-past times (P (*b*>0) =

11   0.999), although, surprisingly, not gaze duration (P(*b*>0) = 0.880). Thus, our effects were

12   numerically larger when readers did skip the word *an*, although the statistical evidence for

13   these effects was still less robust in early measures as compared to later measures.[8]

14   **3.3. Discussion.**

15   In Experiment 2 we primarily set out to test whether readers obtained a preview effect

16   from word *n*+2 driven by the violation of an inter-word phonological rule. The key

---

[8] It could also be asked whether there were any earlier effects of our manipulation, which appeared either in the rare occurrences when participants did fixate word *n*+1 prior to word *n*+2, or even during fixations prior to the eye crossing our invisible boundary. A set of extra analyses conducted to test this possibility revealed very little evidence in favour of earlier effects. A more detailed discussion of these analyses can be found in the appendix.

1   comparison for assessing such an effect was in whether participants experienced greater

2   processing difficulty when given a false preview beginning with a consonant as opposed to a

3   false preview beginning with a vowel. We observed such evidence, finding that participants

4   were more likely to make a regression out of word *n*+2 following a consonant-initial preview,

5   with this resulting in longer go-past times for this word. While there were trends in a similar

6   direction on earlier measures of processing (first fixation duration; gaze durations) there was

7   not strong evidence in favour of the effect in our statistical model, although there was

8   evidence for an effect in first fixation durations when we limited our analysis to cases in

9   which *an* was skipped. Thus, it seems that any processing difficulty primarily resulted in

10   greater re-reading times of earlier portions of the sentence as opposed to identification

11   difficulty of the previewed word. This could be considered surprising; in general preview

12   effects tend to produce reliable effects in the earliest measures of word fixation time, and it

13   would be considered unusual for these effects to only emerge in measures taking re-reading

14   of the sentence into account. However, this pattern makes more sense when considering the

15   likely locus of our effect. Typically, preview effects are due to one preview activating the

16   target word to a greater extent than another preview, resulting in less time being needed to

17   identify that word. In these cases it makes sense that early measures of reading time would be

18   affected, since these measures can be viewed as more of an index of word identification.

19   However, in our study, as we discuss in more detail below, the likely locus of our effect is the

20   rapid detection of a grammatical violation, thus explaining why our manipulation seemed to

21   have more of an effect on re-reading behaviour as opposed to word fixation time.

22   **4. General Discussion**

23   In the current study, we set out to test whether readers would detect a violation of the

24   English convention for the indefinite article *an* to be followed by a word beginning with a

25   vowel sound (e.g. *Icelandic*) as opposed to a consonant sound (e.g. *Mongolian*) while both of

1   these words were still in the parafovea. Across our two experiments we examined whether

2   such violations resulted in a) the reduced skipping of *an* and b) whether reading times of

3   word *n+2* were increased given a violation preview of this word. The answer to the first of

4   these question is clearly no. In all three relevant comparisons for assessing this issue we

5   actually observed a trend towards an effect in the opposite direction, such that participants

6   would skip *an* more given a violation preview than any other preview type, with our

7   combined analysis of the data from Experiment 1 and 2 suggesting that there was a 91%

8   chance that the observed effect was in the opposite direction to our hypothesis. While this

9   level of probability does not indicate strong evidence for the effect, we will return to this

10   counterintuitive trend further below. Our reading time data showed a pattern of effects more

11   in line with what we hypothesised, such that participants seemed to encounter a greater level

12   of processing difficulty when given a violation preview than a grammatical preview. This

13   effect mainly emerged in later measures of processing, such as go-past time and the

14   probability of participants making a regression back to earlier parts of the text. We will

15   primarily focus on these effects for the rest of this discussion.

16   As discussed above, one way of explaining previous contextual fit preview effects is

17   in terms of readers identifying a word on the basis of its parafoveal preview, but having to

18   fixate this word anyway due to it no longer being possible to cancel the saccade towards it,

19   and then attempting to integrate this identified word into the sentence representation during

20   the forced fixation. When this word cannot be integrated, a saccade from the parafoveal word

21   will be cancelled, resulting in increased reading times for that word. This explanation of such

22   effects for word *n+1* already relies on very few instances of eye movement behaviour (~8%

23   in Schotter et al., 2014) and some researchers have already questioned whether these limited

24   instances can explain their contextual fit preview effects (e.g. Veldre & Andrews, 2018;

25   Andrews & Veldre, 2019). These instances should be even rarer for studies of *n+2* preview

1   effects, and as such, we see very little room for explaining our effects within this framework

2   as it is, although acknowledge that future simulations with E-Z Reader may end up

3   contradicting our position. Instead, we favour an account in which, in certain cases, readers

4   will attempt to at least partially integrate a word that has not yet been fully identified into the

5   sentence, or at least with certain other words in the sentence. To be clear, we are not

6   proposing that all integrative processes happen in parallel with lexical processing; rather, we

7   merely intend to argue that when certain aspects of a word's fit within the sentence can be

8   assessed on the basis of lower-level information then readers will take advantage of this. As

9   such, it is not our intention to claim that E-Z Reader's assumptions about integrating multiple

10   words within a sentence together are fundamentally incorrect; rather, we are advancing the

11   position that as such models are further developed, a greater level of nuance will be required

12   in terms of what would be a pre-lexical as opposed to post-lexical integration process.

13   Quite how early this early integration might have occurred in the current study can be

14   viewed in terms of the framework provided by E-Z Reader. We consider the sequence of

15   events to have proceeded as follows. First, while fixated on word *n* readers would often

16   manage to fully identify both word *n* and word *n*+1, and subsequently integrate these

17   identified words into the sentence. At this point attention would have moved onto word *n*+2

18   as a saccade continued to be programmed towards this word. It seems unlikely that the

19   reading system detected the grammatical violation prior to this saccade program becoming

20   non-labile, otherwise we would have observed decreased skipping of *an*. Rather, we believe

21   that our manipulation must have affected the reading system at some point between the time

22   at which the saccade towards word *n*+2 became non-labile, and when participants gained

23   access to visual information extracted upon direct fixation on word *n*+2. As participants

24   detected the grammatical violation they would instantly have begun programming a

25   regression out of the target word on the basis of the preview information on a proportion of

1   trials, resulting in both higher regression probabilities and increased go-past times for our

2   target words, in the absence of reliable first-pass reading effects.

3   It is briefly worth considering whether our manipulation could actually be explained

4   in a trans-saccadic activation account of preview effects, as opposed to a contextual fit

5   account. It is possible that such an account could explain our effect at a basic level, in that

6   participants may be experiencing a phonological preview effect (Ashby & Rayner, 2004;

7   Ashby, Treiman, Kessler, & Rayner; Chace, Rayner, & Well, 2005; Miellet & Sparrow,

8   2004; Pollatsek, Lesch, Morris, & Rayner, 1992). In this account, readers would have gained

9   a preview benefit on the basis of the target word being primed by a word beginning with an

10   initial phoneme which is either more (i.e. a vowel) or less (i.e. a consonant) concordant with

11   this target word. There are a number of reasons why we view this as an unlikely account of

12   our findings. First, it seems unlikely that such a small level of overlap between a preview and

13   target should result in detectable processing time differences, even if this preview was

14   presented as word $n+1$ rather than word $n+2$. Secondly, a recent Bayesian meta-analysis of

15   phonological preview effects (Vasilev, Yates, & Slattery, 2019) suggests that these effects are

16   far smaller than generally assumed, even for previews of word $n+1$. In this meta-analysis it

17   was found that in studies of English the phonological preview effect amounts to about only

18   1.9ms in first fixation duration and 4.5ms in gaze duration. Unfortunately, the meta-analysis

19   does not provide an estimate of the effect size in go-past time, for us to be able to assess

20   whether it grows in a similar way in this measure as in our own study. However, if effects of

21   previews with far higher phonological overlap with the target word are so small for word

22   $n+1$, it would be highly surprising to find effects with word $n+2$ using a weaker manipulation

23   of phonology. Finally, the evidence for our effect was strongest in measures taking

24   regressions into account (i.e. regressions out of word $n+2$; go-past time), without particularly

25   strong evidence in first-pass measures, although effects may have appeared in these measures

1   with greater statistical power. Even ignoring whether effects were reliable in certain measures

2   compared to others, there was a substantial increase in the numerical size of the effect in later

3   measures (7ms in gaze duration vs. 25ms in go-past time). As alluded to in our Discussion of

4   Experiment 2, this seems more in-line with an integration effect than a trans-saccadic

5   activation effect. While it would be possible to discriminate between these two accounts by

6   testing whether our preview effect is still observed when *of* is used in place of *an* (i.e.

7   phonological difference remains without grammaticality difference) we consider the above

8   reasons sufficient to proceed on the understanding that our effect is due to integration, while

9   acknowledging that further empirical work may be required in the future to lend further

10   weight to our arguments. This effect of integration would be consistent with theories of

11   sentence processing that conceptualize comprehension as a form of perceptual cue integration

12   based on inference wherein 'low-level' perceptual information cues the inference of 'higher-

13   level' structure and where in turn, sentence context shapes the processing of incoming lower-

14   level perceptual features (see Martin, 2016).

15   In order to further establish the time-course across which readers detect a violation

16   between the indefinite article form and the following word's form, it may be possible to

17   examine foveal as opposed to parafoveal processing of this word. For example, in order to

18   demonstrate that readers do not perform syntactic integration of a word prior to completing

19   lexical processing, Staub (2011) orthogonally varied whether a word was easy or difficult to

20   identify through a frequency manipulation, and whether it was easy or difficult to integrate

21   into the sentence through various syntactic manipulations. Staub argued in favour of the serial

22   staged-architecture built into E-Z Reader on the basis that these two variables did not interact

23   with each other. If we are correct that in the case of the *an* violation readers are attempting

24   some basic level of integration pre-lexically, then we might find some form of interaction

25   between these violations and the lexical processing difficulty of the target noun. Specifically,

1   the violation may cause participants to terminate lexical processing of the noun early, thus

2   resulting in a reduced frequency effect in violation as opposed to non-violation conditions.

3   A further issue worth considering is the extent to which our manipulation exerted its

4   influence at a phonological level, as opposed to an orthographic level. In the former account

5   the effect would have been due to readers converting the orthographic information into a

6   phonological code, and detecting the violation through that, while in the former account

7   readers would simply determine the violation on the basis of whether the letter is a vowel or

8   consonant. In our study, the effects could have been due to either factor. One way of

9   discriminating between these possibilities would be through the use of words beginning with

10   the letter *u*, which is typically considered a vowel but will often be pronounced with a

11   consonant sound. By testing whether violation effects are observed for constructions such as

12   *an Ukrainian* compared to constructions such as *an Ulsterman* it would be possible to

13   determine the extent to which phonology was used in the detection of our manipulation. It

14   may also be possible to make use of acronyms in a similar manner to Slattery, Pollatsek, and

15   Rayner (2006) to discriminate between a phonological and orthographic locus of our effect,

16   although it is unclear whether the processing of these acronyms could be considered

17   representative of other words.

18   As well as adding to a literature showing that readers assess the fit of a parafoveal

19   word with preceding words in the sentence, our study also adds to a literature suggesting that

20   people may sometime use (relatively) low-level information from words in order to assess

21   their contextual fit. For example, Cutter, Martin, and Sturt (2019) recently demonstrated that

22   people may use initial-letter capitalisation as a cue that an upcoming word is a proper noun

23   prior to completing lexical processing of that word. In our study readers were seemingly

24   assessing whether word *n*+2 was compatible with the preceding article on the basis of its

1   orthographic form, with their being little chance that readers could have processed this word

2   substantially enough for the effect to be post-lexical.

3   One perplexing aspect of our data was the tendency for *an* to be skipped more as

4   opposed to less when it was followed by a consonant-initial preview. This effect makes very

5   little sense in any framework we can think of. Technically, it is possible to view a violation

6   preview of word *n*+2 increasing the skipping of word *n*+1 as an attractor effect, whereby

7   readers' eyes are drawn towards unusual information in the parafovea. However, in the case

8   of our study the violation preview should only start to be considered unusual once

9   participants have already fully identified the preceding word *an*. Presumably, if it was still

10   possible for readers to cancel a saccade to *an* and program one towards word *n*+2 at this point

11   in time, they would chose to do so regardless of whether there is something drawing their eye

12   to word *n*+2. As such, it is not clear to us what mechanism may be responsible for this trend.

13   It could be argued that one possibility is that readers were adopting a word-grouping strategy

14   (e.g. see Radach, 1996), whereby they treated a short article and the following noun as a

15   single perceptual unit when targeting saccades. In such cases, the perceptual pattern *an*

16   followed by a consonant would be considered orthographically unusual, resulting in effects

17   on where the eyes land on this two-word region, similarly to the effect of unusual

18   orthography in a single word (e.g. White & Liversedge, 2004). To our minds, this seems

19   unlikely, especially when considering research contradicting the word-grouping hypothesis

20   (e.g. Drieghe, Pollatsek, Staub, & Rayner, 2008). None-the-less, further work may be

21   necessary to determine if our effects could be explained in this framework.

22   Moving beyond our phonology effects, the current study also contributes to the

23   growing literature examining *n*+2 preview effects during reading. In Experiment 1 we

24   observed an *n*+2 preview effect that was larger than what has typically been observed in prior

25   research, and in Experiment 2 we observed what is – to the best of our knowledge – the first

instance of differential effects of two different incorrect *n*+2 previews, with prior studies only

finding a difference between a correct vs. incorrect preview. Recall that it may be the case

that the small *n*+2 preview effects observed in prior studies could be a consequence of

display change awareness as opposed to pre-processing of information from the parafoveal

word (e.g. Angele et al., 2016); our effects could not be explained in this way, given that

readers responded differently to different types of preview. Thus, our study may constitute

some of the strongest evidence available that readers process word *n*+2 during reading. It

should be noted that in the current study we used target words beginning with a capitalised

proper noun. It may be that this exaggerated the size of our effect relative to studies using

lower-case nouns (see Rayner & Schotter, 2014). Thus, future research may be necessary to

test whether these effects replicate when using common nouns.

4.1. Conclusion.

In closing, we investigated whether readers of English are sensitive to a relatively

low-level inter-word phonological violation when this violation only appears in the

parafovea. Our data suggest that this is the case, with longer target reading times in a

violation condition than a non-violation condition. This effect suggests that readers may

sometimes perform limited integration of an upcoming word into a sentence representation

prior to the completion of lexical processing for that word, broadly consistent with a cue-

based perceptual inference mechanism.

1   **5. Appendix**

2   In the following appendix we present an additional analyses of our data, requested by

3   a reviewer of an earlier version of our article. The purpose of this analysis was to test whether

4   our effect appeared earlier than suggested by our main a prior analysis. Specifically, we

5   examine whether preview type affected fixations made on the word prior to the invisible

6   boundary or on the article *an*. We discuss the analysis for these additional regions for both

7   Experiment 1 and Experiment 2 concurrently.

8   **5.1. Parafoveal-on-foveal effects**

9   An analysis looking at the reading of the word prior to the invisible boundary was

10 conducted to detect any potential parafoveal-on-foveal effects of our manipulation. It may be

11 that readers detected an unusual visual pattern between two parafoveal words (e.g. *an S*) and

12 as such reading times on the pre-boundary word would be affected. To test for such effects

13 we constructed Bayesian linear mixed models. In Experiment 1 these models compared the

14 identity preview to the violation preview. In Experiment 2 one contrast compared the

15 violation preview to the identity preview, while a second contrast compared the violation

16 preview to the incorrect vowel preview. Mean reading times for Experiment 1 and 2 can be

17 seen in Table A.1 and Table A.3, respectively, while the output of the Bayesian linear mixed

18 models can be seen in Table A.2 and Table A.4, respectively.

19 Table A.1

20 Mean (Standard Error) Reading Times on the Word Prior to the Invisible Boundary in
21 Experiment 1.

	Identity/vowel	Violation/consonant
First Fixation Duration	217 (2.44)	216 (2.39)
Gaze Duration	253 (3.82)	246 (3.57)
Go-Past Time	288 (5.08)	286 (5.42)
Regressions Out	0.11 (0.01)	0.14 (0.01)

22

23

1

2   Table A.2

3   Output from our Bayesian mixed effects models for Reading Times on the Word Prior to the
4   Invisible Boundary in Experiment 1. Mean effects estimates (*b*), 95% confidence intervals,
5   and the probability of our effect being larger than zero are reported. Log-units output by the
6   model are reported, followed by values translated into fixation times and probabilities in
7   parentheses.

	Intercept	*B*	L95%	U95%	P(*b*>0)
FFD	5.33 (206)	0.00 (0)	-0.03 (-6)	0.03 (6)	0.523
GD	5.44 (230)	-0.02 (-4)	-0.06 (-13)	0.02 (5)	0.124
GP	5.56 (260)	-0.00 (0)	-0.05 (-13)	0.04 (11)	0.453
Regs. Out	-2.44 (0.08)	0.30 (0.03)	-0.16 (-0.01)	0.75 (0.08)	0.910

8

9   Evidence for a parafoveal-on-foveal effect of our manipulation was limited. In

10   Experiment 1 first fixation durations and go-past times were practically identical regardless

11   of preview type. There was a numerical trend for gaze durations to be 7ms longer given an

12   identity preview, but the statistical evidence for this effect was weak (P(*b*>0) = 0.124).

13   In Experiment 2, there were certain contrasts in our data that could be viewed as

14   evidence for a parafoveal-on-foveal effect of our manipulation when considered in isolation,

15   but not when considered alongside other contrasts within Experiment 2, as well as the data

16   from Experiment 1. For example, in go-past time there was actually evidence for a difference

17   between the identity condition (*m* = 262ms) and the consonant preview condition (*m* =

18   275ms; P(b>0) = 0.976); however, there was no equivalent difference between the consonant

19   preview and the incorrect vowel preview (*m* = 272 ms; P(*b* > 0) = 0.652), which there

20   presumably would have been if the effect was due to participants detecting the violation. In

21   addition, the lack of an effect in this measure in Experiment 1, and the fact that the trend here

22   was in the opposite direction, leads us to suspect the effect in this contrast from Experiment 2

23   may have been spurious. Again, there was very little in the way of a systematic difference

24   between the conditions in either first fixation duration nor gaze duration. Thus, on balance,

1 there seems to be very little evidence for an effect of our manipulation prior to readers

2 crossing the invisible boundary.

3 Table A.3

4 Mean (Standard Error) Reading Times on the Word Prior to the Invisible Boundary in
5 Experiment 2.

	Identity	Vowel	Consonant
FFD	206 (2.28)	207 (2.20)	210 (2.16)
GD	236 (3.52)	232 (3.26)	236 (3.33)
Go-Past	262 (4.77)	272 (5.37)	275 (5.20)
Regression out	0.09 (0.01)	0.12 (0.01)	0.12 (0.01)

6

7 Table A.4

8 Output from our Bayesian mixed effects models for Reading Times on the Word Prior to the
9 Invisible Boundary in Experiment 2. Mean effects estimates (b), 95% confidence intervals,
10 and the probability of our effect being in the expected direction are reported. Log-units output
11 by the model are reported, followed by values translated into fixation times and probabilities
12 in parentheses.

		Intercept	b	L95%	U95%	P(b>/<0)
FFD	Identity Contrast	5.30 (200ms)	-0.02 (-4ms)	-0.05 (-10ms)	0.01 (2ms)	0.928
	Vowel Contrast	5.30 (200ms)	-0.01 (-2ms)	-0.04 (-8ms)	0.02 (4ms)	0.814
GD	Identity Contrast	5.38 (217ms)	-0.01 (-2ms)	-0.05 (-11ms)	0.03 (7ms)	0.673
	Vowel Contrast	5.38 (217ms)	-0.01 (-2ms)	-0.05 (-11ms)	0.03 (7ms)	0.713
GP	Identity Contrast	5.49 (242ms)	-0.04 (-9ms)	-0.09 (-21ms)	-0.00 (-0ms)	0.976
	Vowel Contrast	5.49 (242ms)	-0.01 (-2ms)	-0.06 (-14ms)	0.04 (10ms)	0.652
Reg. out	Identity Contrast	-2.40 (8.3%)	-0.39 (-2.5%)	-0.93 (-4.8%)	0.08 (0.63%)	0.946
	Vowel Contrast	-2.40 (8.3%)	-0.06 (-0.4%)	-0.48 (-3.0%)	0.35 (3%)	0.621

13

14 **5.2 Preview effects on word n+1**

15   It is also possible that, in cases where readers' eyes landed on word n+1 (i.e. the

16 article *an*) prior to word n+2, preview effects may have appeared on word n+1. Prior to

17 discussing fixation times on word n+1 it is worth noting that the eyes only landed on this

18 word in a third of trials. As such, in both of our experiments there was very little data for this

19 analysis, and so estimates of any effects were highly prone to noise. As such, finding any

20 strong evidence in favour of an effect of our manipulation was highly unlikely, even if such

21 an effect was present. None the less, we do present such an analysis below. The Bayesian

22 linear mixed models for this analysis were structured in an identical manner as the models

1   which examined reading behaviour on word *n+2*. Mean reading times on this word for

2   Experiment 1 and Experiment 2 can be seen in Table A.5 and Table A.7, while the output of

3   the Bayesian linear mixed models can be seen in Table A.6 and Table A.8.

4   Table A.5

5   Mean (Standard Error) Reading Times on the Word *an* in Experiment 1.

6

	Identity/vowel	Violation/consonant
First Fixation Duration	212 (4.46)	211 (4.44)
Gaze Duration	219 (4.99)	215 (4.89)
Go-Past Time	264 (8.52)	269 (9.87)
Regressions Out	0.14 (0.02)	0.17 (0.02)

7   Table A.6

8   Output from our Bayesian mixed effects models for Reading Times on *an* in Experiment 1.
9   Mean effects estimates (*b*), 95% confidence intervals, and the probability of our effect being
10   in the expected direction are reported. Log-units output by the model are reported, followed
11   by values translated into fixation times and probabilities in parentheses.

	Intercept	*B*	L95%	U95%	P(*b*>0)
FFD	5.30 (200)	-0.00 (-0)	-0.06 (-12)	0.05 (10)	0.460
GD	5.32 (204)	-0.01 (-2)	-0.07 (-14)	0.04 (8)	0.321
GP	5.47 (237)	0.01 (2)	-0.08 (-18)	0.09 (22)	0.557
Regs. Out	-2.30 (0.09)	0.39 (0.04)	-0.42 (-0.03)	1.20 (0.16)	0.829

12

13   In Experiment 1 fixation times on *an* were highly similar in both conditions, with any

14   numerical differences varying in direction from measure to measure. A similar lack of

15   systematic effects was present in Experiment 2, whereby numerical differences were

16   generally small, and varied in direction across measures. Thus, it seems unlikely that effects

17   of our manipulation emerged in fixation times on word *n+1*, although once again it seems

18   unlikely that we would have had adequate power to detect such effects if they were present.

19   Table A.7

20   Mean (Standard Error) Reading Times on the Word *an* in Experiment 2.

	Identity	Vowel	Consonant
FFD	208 (3.63)	213 (3.91)	205 (4.07)
GD	213 (4.14)	221 (4.44)	214 (4.75)
Go-Past	251 (6.77)	260 (7.63)	266 (8.48)
Regression out	0.15 (0.02)	0.15 (0.02)	0.18 (0.02)

1   Table A.8

2   Output from our Bayesian mixed effects models for Reading Times on *an* in Experiment 2.
3   Mean effects estimates (*b*), 95% confidence intervals, and the probability of our effect being
4   in the expected direction are reported. Log-units output by the model are reported, followed
5   by values translated into fixation times and probabilities in parentheses.

		Intercept	*B*	L95%	U95%	P(*b*>/<0)
FFD	Identity Contrast	5.31 (202ms)	-0.01 (-2ms)	-0.05 (-10ms)	0.04 (8ms)	0.636
	Violation Contrast	5.31 (202ms)	-0.03 (-6ms)	-0.08 (-16ms)	0.02 (4ms)	0.113
GD	Identity Contrast	5.33 (206ms)	-0.02 (-4ms)	-0.07 (-14ms)	0.03 (6ms)	0.782
	Violation Contrast	5.33 (206ms)	-0.03 (-6ms)	-0.08 (16ms)	0.03 (6ms)	0.168
GP	Identity Contrast	5.44 (230ms)	-0.00 (-0ms)	-0.07 (-16ms)	0.06 (14ms)	0.515
	Violation Contrast	5.44 (230ms)	0.03 (7ms)	-0.05 (-11ms)	0.10 (24ms)	0.755
Reg. out	Identity Contrast	-2.07 (8.9%)	0.01 (0.1%)	-0.56 (-4.5%)	0.57 (7%)	0.484
	Violation Contrast	-2.07(8.9%)	0.13 (1.4%)	-0.54 (-4.4%)	0.73 (9.5%)	0.685

6

7

1 <div align="center">Acknowledgements</div>

7

References

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. https://doi.org/10.1037/a0036396

Angele, B., & Rayner, K. (2011). Parafoveal processing of word n + 2 during reading: Do the preceding words matter? *Journal of Experimental Psychology: Human Perception and Performance, 37*, 1210-1220. doi:10.1037/a0023096

Angele, B., & Rayner, K. (2013). Processing in the parafovea: Are articles skipped automatically? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 649-663. http://dx.doi.org/10.1037/a0029294

Angele, B., Slattery, T. J., & Rayner, K. (2016). Two stages of parafoveal processing during reading: Evidence from a display change detection task. *Psychonomic Bulletin & Review*, *23*, 1241-1249. https://doi.org/10.3758/s13423-015-0995-0

Angele, B., Slattery, T. J., Yang, J., Kliegl, R. & Rayner, K. (2008). Parafoveal processing in reading: Manipulating n+1 and n+2 previews simultaneously. *Visual Cognition, 16*, 697-707. doi:10.1080/13506280802009704

Ashby, J., & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements, Language and Cognitive Processes, 19, 391-426.

Ashby, J., Treiman, R., Kessler, B., & Rayner, K. (2006). Vowel processing during silent reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 416-424.

1 Briihl, D., & Inhoff, A. W. (1995). Integrating information across fixations during reading:

2 The use of orthographic bodies and exterior letters. *Journal of Experimental*

3 *Psychology: Learning Memory and Cognition, 21*, 55-67.

4 Brothers, T., & Traxler, M.J. (2016). Anticipating syntax during reading: Evidence from the

5 boundary change paradigm. *Journal of Experimental Psychology: Learning Memory*

6 *and Cognition, 42*, 1894-1906. Doi: 10.1037/xlm0000257

7 Chace, K. H., Rayner, K., & Well, A. D. (2005). Eye movements and phonological preview:

8 Effects of reading skill. *Canadian Journal of Experimental Psychology, 59*, 209-217.

9 Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2014). Preview benefit in English spaced

10 compounds. *Journal of Experimental Psychology: Learning Memory and Cognition,*

11 *40*, 1778-1786. doi:10.1037/xlm0000013

12 Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2015). How is information integrated across

13 fixations in reading? In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of*

14 *reading* (pp. 245-260). New York, NY: Oxford University Press.

15 Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2017). Is orthographic information from

16 multiple parafoveal words processed in parallel: An eye-tracking study. *Journal of*

17 *Experimental Psychology: Human Perception and Performance, 43*, 1550-1567.

18 http://dx.doi.org/10.1037/xhp0000408

19 Cutter, M. G., Martin, A. E., & Sturt, P. (2019). Capitalization Interacts with Syntactic

20 Complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

21 Advance online publication. https://doi.org/10.1037/xlm0000780

Cutter, M. G., Martin, A. E., & Sturt, P. (2020). The activation of contextually predictable words in syntactically illegal positions. *Quarterly Journal of Experimental Psychology.* Advance online publication. https://doi.org/10.1177/1747021820911021

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*, 1117-1121. https://doi.org/10.1038/nn1504

Drieghe, D., Pollatsek, A., Staub, A., & Rayner, K. (2008). The word grouping hypothesis and eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1552-1560. http://dx.doi.org/10.1037/a0013017

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 166-190.

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eyemovement-contingent display change study. *Memory & Cognition, 32*, 523-532.

Ito, A., Martin, A. E., Nieuwland, M. S. (2016). How robust are prediction effects in language comprehension? Failure to replicate article-elicited N400 effects. *Language, Cognition, and Neuroscience, 32*, 954-965. https://doi.org/10.1080/23273798.2016.1242761

Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 209-229.

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word n+2. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1250-1255. doi:10.1037/0096-1523.33.5.1250

Martin, A. E. (2016). Language processing as cue integration: Grounding the psychology of language in perception and neurophysiology. *Frontiers in psychology*, *7*, 120.

McConkie, G. W., & Rayner, K. (1975). The span of effective stimulus during a fixation in reading. Perception and Psychophysics, 17, 578-586.

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

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.

Radach, R. (1996) Blickbewegungen beim Lesen: Psychologische Aspekte der Determination von Fixationspositionen (Eye Movements in Reading). Münster/New York: Waxmann; 1996.

Radach, R., Inhoff, A. W., Glover, L., & Vorstius, C. (2013). Contextual constraint and N+2 preview effects in reading. *The Quarterly Journal of Experimental Psychology, 66*, 619-633. doi:10.1080/17470218.2012.761256

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

Rayner, K., Juhasz, B. J., & Brown, S. J. (2007). Do readers obtain preview benefit from word N+2? A test of serial attention shift versus distributed lexical processing models

of eye movement control in reading. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 230-245. doi:10.1037/0096-1523.33.1.230

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. Doi:10.1037/a0036763

Reichle, E. D., Pollatsek, A., Rayner, K. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research, 7,* 4-22. doi: 10.1016/j.cogsys.2005.07.002

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using EZ Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16*, 1-21. doi:10.3758/PBR.16.1.1

Risse, S., & Kliegl, R. (2012). Evidence for delayed parafoveal-on-foveal effects from word n+2 in reading. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 1026-1042. doi: 10.1037/a0027735

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 out bets on seeing the future: Eye tacking and electrophysiology evidence for parafoveal lexical processing and saccadic control by partial word recognition. *Psychology of Learning and Motivation, 68*, 263-298. https://doi.org/10.1016/bs.plm.2018.08.011

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. http://dx.doi.org/10.1037/xlm0000281

1   Schotter, E. R., & Leinenger, M. (2016). Reversed preview benefit effects: Forced fixations

2   emphasize the importance of parafoveal vision for efficient reading. *Journal of*

3   *Experimental Psychology: Human Perception and Performance, 42*(12), 2039-2067.

4   http://dx.doi.org/10.1037/xhp0000270

5   Schotter, E. R., Leinenger, M., & von der Malsburg, T. (2018). When your mind skips what

6   your eyes fixate: How forced fixations lead to comprehension illusions in reading.

7   *Psychonomic Bulletin & Review, 25,* 1884-1890. https://doi.org/10.3758/s1342

8   Schotter, E. R., Reichle, E. D., & Rayner, K. (2014). Rethinking parafoveal processing in

9   reading: Serial-attention models can explain semantic preview benefit and N+2

10   preview effects. *Visual Cognition, 22*, 309-333. doi:10.1080/13506285.2013.873508

11   Schotter, E. R., von der Malsburg, T., & Leinenger, M. (2019). Forced fixations, trans-

12   saccadic integration, and word recognition: Evidence for a hybrid mechanism of

13   saccade triggering in reading. *Journal of Experimental Psychology: Learning,*

14   *Memory, and Cognition, 45*, 677-688. http://dx.doi.org/10.1037/xlm0000617

15   Slattery, T. J., Pollatsek, A., & Rayner, K. (2006). The time course of phonological and

16   orthographic processing of acronyms in reading: Evidence from eye movements.

17   *Psychonomic Bulletin & Review*, *13*, 412-417. https://doi.org/10.3758/BF03193862

18   Staub, A. (2011). Word recognition and syntactic attachment in reading: Evidence for a

19   staged architecture. *Journal of Experimental Psychology: General, 140*(3), 407-433.

20   http://dx.doi.org/10.1037/a0023517

21   Vasilev, M. R., & Angele, B. (2017). Parafoveal preview effects from word *n*+1 and word

22   *n*+2 during reading: A critical review and Bayesian meta-analysis. *Psychonomic*

23   *Bulletin & Review, 24*, 666-689. doi: 10.3758/s13423-016-1147-x

Vasilev, M. R., Yates, M., & Slattery, T. J. (2019). Do readers integrate phonological codes across saccades? A Bayesian meta-analysis and a survey of the unpublished literature. *Journal of Cognition*, *2*, 1-29. https://doi.org/10.5334/joc.87

Veldre, A., & Andrews, S. (2016). Parafoveal preview effects depend on both preview plausibility and target predictability. *The Quarterly Journal of Experimental Psychology,* 1-12.

Veldre, A., & Andrews, S. (2018). Beyond cloze probability: Parafoveal processing of semantic and syntactic information during reading. *Journal of Memory and Language, 100*, 1-17. https://doi.org/10.1016/j.jml.2017.12.002

Veldre, A., Reichle, E. D., Wong, R., & Andrews, S. (2020). The effect of contextual plausibility on word skipping during reading. *Cognition*, *197*. https://doi.org/10.1016/j.cognition.2020.104184

Wakeford, L., & Murray, W. (2013). Effects of parafoveal plausibility during reading. *Proceedings of the Annual Meeting of the Cognitive Science Society, 35*. https://escholarship.org/uc/item/9k45772m

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. https://doi.org/10.1080/13506285.2014.890689