Author preprint of a manuscript accepted for publication in Psychophysiology on June-10th, 2015

Parafoveal processing in reading Chinese sentences: Evidence from event-related brain potentials

Nan Li[1*] & Florian Niefind[2*], Suiping Wang[1], Werner Sommer[2], Olaf Dimigen[2]

[1] South China Normal University, Guang Zhou
[2] Humboldt University at Berlin

* These authors contributed equally to this work.

Correspondence concerning this article should be addressed to:
Suiping Wang, School of Psychology, South China Normal University, Guangzhou, 510631, China, Phone: (86) 20-85216510, Fax: (86) 20-85216412, E-mail: suipingscnu@gmail.com.
Or to:
Werner Sommer, Department of Psychology, Humboldt University at Berlin, Rudower Chaussee 18, 12489 Berlin, Germany. Tel.: 0049-30-2093-99127, Fax: 0049-30-2093-4910, E-mail: Werner.Sommer@cms.hu-berlin.de

Author note:
This work was supported by grants from the Natural Science Foundation of China (NSF 31271086), Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme (GDUPS 2011) to Suiping Wang. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank Wenjia Zhang and Fang Wang for material preparation, Niels Krause and Linda Gerresheim for their help in data collection, Rainer Kniesche for technical support.

Abstract

Natural reading involves the preprocessing of upcoming words, resulting in shorter fixations on words visible in the parafovea during preceding fixations. While this preview benefit is established in behavior, its brain-electric correlates have only recently been investigated. Using fixation-related potentials, Dimigen et al. (2012) demonstrated an attenuation of the occipito-temporal N1 component for words that were parafoveally visible during preceding fixations. In contrast, Barber et al. (2013), using an RSVP paradigm with parafoveal flanker words, observed no such general preview benefit in ERPs, but instead reported N400 effects triggered by semantically incongruous parafoveal words. To follow up on these discrepant findings and to extend them to a non-alphabetic writing system, we conducted two ERP experiments with Chinese readers using Barber et al.'s paradigm and rigorous fixation control via eye-tracking. We replicate robust parafoveal N400 semantic congruency effects in Chinese participants. Additionally, we found that once a word was directly looked at, words after a valid preview elicited a smaller N1 and a weaker N400 than those after an invalid preview. Results underline the importance of considering parafoveal vision in ERP studies on reading.

Keywords: parafovea, reading, EEG, N400, semantic processing, preview positivity, RSVP-with-flankers

During reading, the eyes move in a series of rapid, jerk-like saccades in order to bring new words into the fovea, that is, the central 1–2 degrees of the visual field. Visual input is obtained mainly during the intervening periods of stable gaze (fixations). Due to the drop-off in visual acuity from the fovea to the parafoveal (2-5°) and peripheral (>5°) visual field, the amount of information gleaned from these outer regions is limited (Rayner, 1998). Nevertheless, readers clearly do use parafoveal information to make the reading process more efficient (McConkie, & Rayner, 1975; Rayner, 1975). Two key research questions, extensively investigated for four decades with eye-tracking, concern the types of information that readers obtain from not-yet-fixated words and how this preprocessing influences their subsequent recognition (Schotter, Angele, & Rayner, 2012).

Electro-encephalography (EEG) is a promising method in this regard. Unlike fixation durations, which only represent the end products of information processing, the EEG provides a continuous measure of neurocognitive processes at high temporal resolution and an approximation of the brain areas involved in – and affected by – parafoveal processing. However, as we will explain below, the few existing EEG studies on parafoveal processing (Barber, Doñamayor, Kutas, & Münte, 2010; Barber, van der Meij, & Kutas, 2013; Dimigen, Kliegl, & Sommer, 2012; Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011; Kretzschmar, Bornkessel-Schlesewsky, & Schlesewsky, 2009) have not only produced partially conflicting results, but have all focused on Western alphabetic languages. Aim of the present study was to follow up on the findings of two recent studies (Barber et al., 2013; Dimigen et al., 2012), which have investigated the depth of parafoveal processing and its brain-electric correlates. Specifically, we were interested whether the results of these studies – obtained with Western alphabetic languages (Spanish, German) extend to readers of Chinese, a non-alphabetic writing system in which parafoveal preview effects, even at a semantic level, have been stably found in eye movement research, indicating that Chinese may be a suitable language to investigate parafoveal processing effects in ERPs. Furthermore, we investigated the role of involuntary lateral eye movements and the influence of stimulus timing for observing parafoveal effects in ERPs.

Much of what is known about parafoveal processing has been learned from eye-tracking experiments using the boundary paradigm (Rayner, 1975). In this paradigm, an upcoming word in a sentence is masked while in parafoveal vision and only unmasked once it receives a direct fixation. By varying the degree of valid information that the parafoveal mask contains about the word later seen in foveal vision, one can study how much of this information is extracted parafoveally. For example, while in the parafovea, the target word ”chair” could be masked either by the word itself (“chair”), by an orthographically similar word (“chase”), or by some completely unrelated word serving as an invalid preview (e.g., “light”). The classic result obtained with this paradigm is the preview benefit: Fixations are significantly shorter after informative/valid than after uninformative/invalid previews, showing that at least some of the parafoveal information had been processed by the reader. Relative to a situation with an invalid preview, the strongest facilitation is observed after a fully valid, identical preview (“chair”). This difference is called the *identity preview benefit*. However, readers of English text have been found to also benefit from masks that are orthographically or phonologically similar to the target (e.g. Balota, Pollatsek, & Rayner, 1985; see Schotter et al., 2012 for a review) and from masks consisting of the identical word written in a different case (“CHAIR”). This suggests that the identity preview benefit is not primarily due to the preprocessing of visual low-level features but based on more abstract orthographical or phonological information.

For word-masks containing semantic information about the target word (e.g. ”table” for ”chair”) results are more controversial. Some studies on German reported such a *semantic preview benefit* (Hohenstein, Laubrock, & Kliegl, 2010; Hohenstein & Kliegl, 2013), while the effect was not found with English and Spanish readers (Rayner, Balota, & Pollatsek, 1986; White, Bertram, & Hyönä, 2008; Rayner, Schotter, & Drieghe, 2014; but see also Schotter, 2013). Interestingly, readers of Chinese (a non-alphabetic script) seem to use parafoveal information to a much larger degree. Semantic preview benefits in Chinese reading have been found consistently (Tsai, Kliegl, & Yan, 2012; Yan, Richter, Shu, & Kliegl, 2009; Yan, Risse, Zhou, & Kliegl, 2012; Yan, Zhou, Shu, & Kliegl, 2012; Yang, Wang, Tong, & Rayner, 2012; Yen, Tsai, Tzeng, & Hung, 2008).

Preview benefits and parafoveal semantic processing have also been investigated with EEG (Baccino & Manunta, 2005; Dimigen et al., 2011, 2012; Kretzschmar et al., 2009; Barber et al., 2010; 2013; Simola, Holmqvist, & Lindgren, 2009). Most of these studies used fixation-related brain potentials (FRPs) in paradigms using isolated word pairs (Baccino & Manunta, 2005; Simola et al. 2009), lists of nouns (Dimigen et al. 2012), or natural sentences (Kretzschmar et al. 2009; Dimigen et al., 2011).

Here, we will focus on a design by Barber and colleagues (2013), which differs from the other studies in two regards. First, the authors recorded ERPs in a variant of the rapid serial visual presentation (RSVP) paradigm, in which words in a sentence are usually presented one by one in the screen center to avoid eye movement artifacts in the EEG. In their paradigm, Barber et al. added the preceding and following words in the sentence as flanking words to the left and right of the word presented at fixation, thereby allowing for parafoveal preview of the upcoming word. Although this paradigm is obviously different from natural reading, it allows examining the effect of various factors modulating parafoveal processing while keeping the location of the parafoveal information constant and minimizing the influence of eye movements (Barber et al., 2013). Second, Barber et al. used a design in which contextual congruency of a target word was manipulated independently during its parafoveal and foveal presentation.

Each sentence contained a target that was either semantically congruent or incongruent with regard to the preceding context. Like the other words of the sentence, this target first appeared in the right parafoveal location and only afterwards in the foveal location. Both during its initial parafoveal presentation and during the subsequent foveal presentation, either the congruent or the incongruent version of the target was presented, that is, parafoveal and foveal congruency were manipulated orthogonally within the same trial. ERP waveforms were analyzed relative to both the parafoveal and the foveal target presentation. As additional factors, Barber and colleagues manipulated the contextual constraint of the sentence and the speed of sentence presentation. Sentences were either highly constraining, rendering the target highly predictable, or they were of low constraint. Reading speed was manipulated by changing the duration of the inter-stimulus blank interval (ISI) inserted between words. This resulted in a stimulus-onset asynchrony (SOA) that was either long (450 ms) or short (250 ms), resembling the speed of normal saccadic reading.

Barber and colleagues (2013) already found a semantic congruency effect on the centroparietal N400 component (Kutas & Hillyard, 1980) in ERPs time-locked to the *parafoveal* presentation of the target. Specifically, incongruent targets yielded more negative ERP amplitudes around 450 ms after their onset in the parafoveal position. At the long SOA, this effect was found regardless of constraint. At the reading-like SOA, it only persisted for high-constraint sentences, but vanished for low-constraint sentences. These results suggest that readers can evaluate the contextual fit of a not-yet-fixated word, provided that this process is facilitated by strong predictability or an unnaturally slow presentation rate (see also Kretzschmar et al., 2009). In addition, regardless of constraint and SOA, the authors observed the typical – and much stronger – N400 congruency effect during the subsequent foveal target presentation.

As a side result, Barber et al. reported that the ERP time-locked to the foveal presentation was entirely unaffected by the type of word shown previously in the parafovea. In other words, in their ANOVAs of ERP amplitude across the scalp (see their pages 54 and 57), there were no interactions between parafoveal and foveal congruency. This result is surprising, because by manipulating parafoveal and foveal congruency in an orthogonal fashion, in the design of Barber and colleagues, also a third factor is implicitly manipulated: the *validity of the preview*. If a congruent word in the parafoveal position is followed by a congruent word in the foveal position (or an incongruent parafoveal word is followed by an incongruent foveal word), readers obtain a valid preview on the foveal word. However, in the other half of trials, in which a different word is shown in both positions, the reader obtains no useful preview. It should be beneficial to have a valid preview rather than an invalid preview, regardless of congruency. The absence of such a facilitation in the study of Barber et al. (2013) is puzzling, because as explained above, identity preview benefits are a benchmark finding in eye-tracking research (Schotter et al., 2012).

Importantly, brain-electric correlates of the identity preview benefit have recently been established in FRPs recorded during the fluent, left-to-right reading of word lists (Dimigen et al., 2012). Word lists consisted of five German nouns and the participants' task was to detect animal names contained in some of the lists. Using the boundary paradigm, Dimigen et al. then compared the brain responses following three types of previews: a valid preview (identical word), a semantically associated preview, or an invalid preview (unrelated word). The authors found no effect of semantically associated previews. However, upon their direct fixation, words with a valid preview yielded more positive FRP amplitudes at occipito-temporal electrode sites between 200–280 ms after fixation onset. This reduction of the N1 component, called *preview positivity,* was attributed to partial orthographic priming. In addition, a valid preview also tended to reduce the size of the following N400 component, but this *late preview effect* was only marginally significant.

Together, these findings suggest that having a valid preview on a word – the default in any natural reading situation – can have a global modulating effect on the electrophysiological response to words that is overlooked in traditional RSVP studies and therefore important to study with ERPs.

The present study pursued three goals: The first was to ensure that the parafoveal N400 effects reported in the RSVP-with-flanker paradigm are really the result of parafoveal perception. A potential problem with this paradigm is that even a small proportion of trials with a rightward-shifted eye position will also produce weak but spurious “parafoveal” N400 effects. To control for this possibility, Barber and colleagues analyzed the electro-oculogram (EOG) aligned to the parafoveal target presentation. Furthermore, they argued that word presentations of only 100 ms do not allow participants to read the central word and to move the eyes rightwards. However, there are problems with this logic. First, the EOG with a bipolar reference is not sensitive to detect microsaccades or slow drifting eye movements (Dimigen et al., 2009; Keren et al., 2011). Second, an analysis of the target-locked EOG does not exclude the possibility that participants execute rightwards eye movements at some earlier point during the trial and simply continue reading the sentence from there. To exclude this possibility, we controlled central fixation with a high-resolution eye-tracker in the present study.

Our second, closely related goal was to test whether parafoveal N400 effects generalize from alphabetic languages to Chinese script, which – despite being also read from left to right in its modern version – is believed to differ from alphabetic languages in terms of the extent of parafoveal preprocessing. We have already mentioned the consistency of semantic preview benefits in Chinese readers' fixation behavior (e.g. Yan et al., 2009). Yan and colleagues argue that lexical semantic preview effects in Chinese are present because there is a more direct mapping from the character symbol to its meaning, as many Chinese characters consist of semantic primitives, which are pictographic to a certain extent. This direct mapping from the character symbol to its meaning might facilitate semantic integration of the word in a sentence. Moreover, because most Chinese words consist of only one or two characters and no more than four characters, information is more densely packed than in Western writing systems. Therefore, we expected that if they are not a trivial result of lateral eye movements, the previous N400 findings in Spanish should generalize to Chinese script.

The third motivation for our study was to investigate the brain-electric correlates of the identity preview benefit in Chinese readers. A potential problem in this regard is the discrepancy between the findings of Dimigen et al. (2012) obtained in saccadic reading and those of Barber et al. (2013) obtained with RSVP. Here, we hypothesized that the surprising absence of an identity preview effect in Barber et al. (2013) might be due to the unnaturally short stimulus durations (100 ms) and long blank intervals (ISIs ≥ 150 ms) used in their paradigm. Preview benefits in behavior have been shown to increase with increasing fixation durations on the pre-target word, so short stimulations of just 100 ms may attenuate the effect (Kliegl, Hohenstein, Yan, & McDonald, 2013; Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999; Yan, Risse, et al., 2012). Another possible explanation is that parafoveally obtained information decays over the duration of a long ISI. Pernet, Uusvuori, and Salmelin (2007) showed that parafoveal priming effects can be short-lived: In a lexical decision task, they found identity priming with a 50 ms ISI between parafoveal prime and foveal target; however, with a 100 ms ISI, there was no facilitation. Because saccades during normal reading last only around 30 ms (as in the Dimigen et al., 2012 study), an ISI of 150 ms may already be too long for parafoveal information to affect foveal processing. To explore this possibility, we systematically manipulated the ISI in the present study.

Our design was as follows: In two experiments, we adopted the RSVP-with-flankers paradigm for Chinese readers and presented sentences at a reading-like SOA of 280 ms, very similar to that used by Barber and colleagues in their short SOA condition. For one target word in the sentence, parafoveal and foveal congruency were independently manipulated (see Figure 1). This design allowed us to test the following hypotheses: First, if the parafoveal N400 effect of semantic congruency is genuine, we expected to replicate it under conditions of strict fixation control. As Chinese readers have shown strong parafoveal processing in behavioral studies, we anticipated that this effect should be very robust. Second, based on the FRP results by Dimigen et al. (2012), we hypothesized that in ERP waveforms aligned to the *foveal* target presentation, we should see a distinct influence of the validity of the preview. Specifically, we expected an attenuation of the N1 component evoked by the foveal word (preview positivity). We did not have a strong hypothesis regarding the N400 elicited by the foveal presentation. However, based on the previous research, we assumed that ERP amplitude in this later time window may depend both on the validity of the preview (i.e., late preview effect, as in Dimigen et al., 2012) and the contextual fit of the foveally presented word (i.e., foveal congruency effect, as in Barber et al. 2013).

In Experiment 1, presentation times were adjusted to resemble the average duration of fixations and eye movements during natural reading (see also Dambacher et al., 2012). Words were presented for 250 ms and the ISI was set to a typical saccade duration of 30 ms. In Experiment 2, we kept the SOA constant, but used the same brief presentation durations (100 ms) as Barber and colleagues, resulting in a longer ISI.

Experiment 1

Method

Participants. Sixteen participants (9 women, age 21 to 33, *M* = 25.0 years) with normal or corrected-to-normal vision were recruited from a pool of native Mandarin-speaking Chinese students temporally studying in Germany. According to the questionnaire by Oldfield (1971), all participants were right-handed (laterality quotient: *M* = +100). Participants were naive with regard to the exact purpose of the experiment. Written informed consent, conforming to the declaration of Helsinki (World Medical Association, 2013), was obtained prior to the experiment.

Materials. A total of 152 Chinese sentences with a similar syntactic structure (animate noun + verb + adjective/ quantifier + inanimate noun + …) and with a length between 13 and 19 characters (M = 16) were constructed. The target word was always the inanimate noun in the sentence. It was embedded somewhere between positions 9th to 12th character in the sentence and followed by an additional 4 to 9 characters. The target word was always a one-character word, but the remaining words in the sentence could consist of one, two or three characters. The materials were evaluated with a congruency rating by 16 undergraduate students from South China Normal University, who did not participate in the main experiment. Subjects had to rate the congruency of the sentence up to and including the target word on a 5-point scale (where 1 = highly incongruent; 5 = highly congruent). Mean congruency ratings of congruent and incongruent words were 4.37 (SD = 0.53) and 1.37 (SD = 0.41), respectively. The difference was significant, $b = 1.50$, SE = 0.02, $t = 60.3$. The appearances of the target words in the congruent and incongruent contexts were counterbalanced across participants.

Sentential constraint of the sentences was assessed in two sentence completion studies (cloze procedures) performed by 44 and 36 undergraduate students, respectively, from South China Normal University. In the first study, participants were presented with each sentence up to the target word (but not including it) and told to complete it with “the first word that comes to mind”. In the second study, the same procedure was used, but the sentences were only presented up to the pre-target word (but not including it). Sentential constraint was defined as the cloze probability of the sentence continuation given with the highest probability across participants (Federmeier, Kutas, & Schul, 2010). Sentential constraint ranged from 2.3% (lowest possible value, i.e. each participant guessed a different word) to 100%, with a mean value of 47% at the pre-target word position (SD=23%) and from 2.8% (lowest possible value) to 100%, with a mean value of 48% at the target word position (SD=22%).

Apparatus. A 22-inch Iiyama Vision Master monitor was used to display the stimuli. The monitor was set to a refresh rate of 160 Hz; the resolution was 1024×768 pixels. Eye movements were recorded with an infrared-video based eye tracker (iView X™ Hi-Speed system, SMI GmbH, Germany) at a rate of 240 Hz and an instrument spatial resolution of better than 0.025°. The eye tracker monitored movements of the right eye, although viewing was binocular. Head position was stabilized via the chin and forehead rests of the tracker. Tracking quality was controlled with a fixation check at the onset of each trial (see below). The EEG was recorded from 42 Ag/AgCl scalp electrodes placed at standard positions of the 10-10 electrode system and referenced online against the left mastoid. The setup also included inferior posterior electrodes (PO9, PO10, Iz). EEG electrodes were placed in a textile cap. In addition, the electro-oculogram (EOG) was recorded from four electrodes positioned on the infraorbital ridge and outer canthus of each eye. A ground electrode was placed at FCz. Signals were amplified with Brain Products DC amplifiers, with a time constant of 10 s, and sampled at 500 Hz. Impedances were kept below 5 kΩ. Eye track and EEG were synchronized using the EYE-EEG plug-in (Dimigen et al., 2011; see http://www2.hu-berlin.de/eyetracking-eeg) for EEGLAB (Delorme & Makeig, 2004). Exact synchronization was ensured with shared TTL trigger pulses that were sent from the presentation PC (running Presentation, Neurobehavioral Systems Inc., Albany, CA) to the EEG and eye tracker recording computers on each trial. Offline, the EEG was band-pass filtered from 0.1 to 40 Hz (using EEGLAB's new FIR filter for low-pass filtering and subsequently the ERPLAB Butterworth filter for high-pass filtering) and digitally re-referenced against the mean of all electrodes (average reference).

Procedure. Prior to the experiment, participants were first given instructions and then randomly assigned to one of four stimulus sets. Before the start of the actual experiment, participants did a small number of practice trials during which visual feedback was given if their gaze left the central fixation area (the screen background flashed red for one display cycle). The experiment started after participants had given the correct behavioral response on four consecutive trials, but after a minimum of eight practice trials.

The trial scheme is shown in Figure 1. Each trial began with a fixation check, followed by the presentation of the sentence. If the fixation check failed, the experimenter decided whether a calibration of the eye-tracker was necessary. Sentences were presented in triads of Chinese characters (foveal target character plus a flanker character on each side) for 250 ms, followed by a 30 ms blank interval. The target word was always a one-character word, so that its congruency could be unambiguously determined during its parafoveal and foveal presentation. As in the Barber et al. (2013) study, a congruent or an incongruent word was shown in the parafoveal and foveal position, yielding four experimental conditions. Please note that this also means that in half of the trials, the target word stimulus changed from its parafoveal to its foveal position. Thus, in two of the four conditions, the word changed (congruent to incongruent, incongruent to congruent) and in two conditions it stayed the same in the parafoveal and foveal location (congruent to congruent and incongruent to incongruent, see below for an example sentence). In the following, these two cases will be referred to as an invalid preview (the word changes) and a valid preview (the word stays the same).

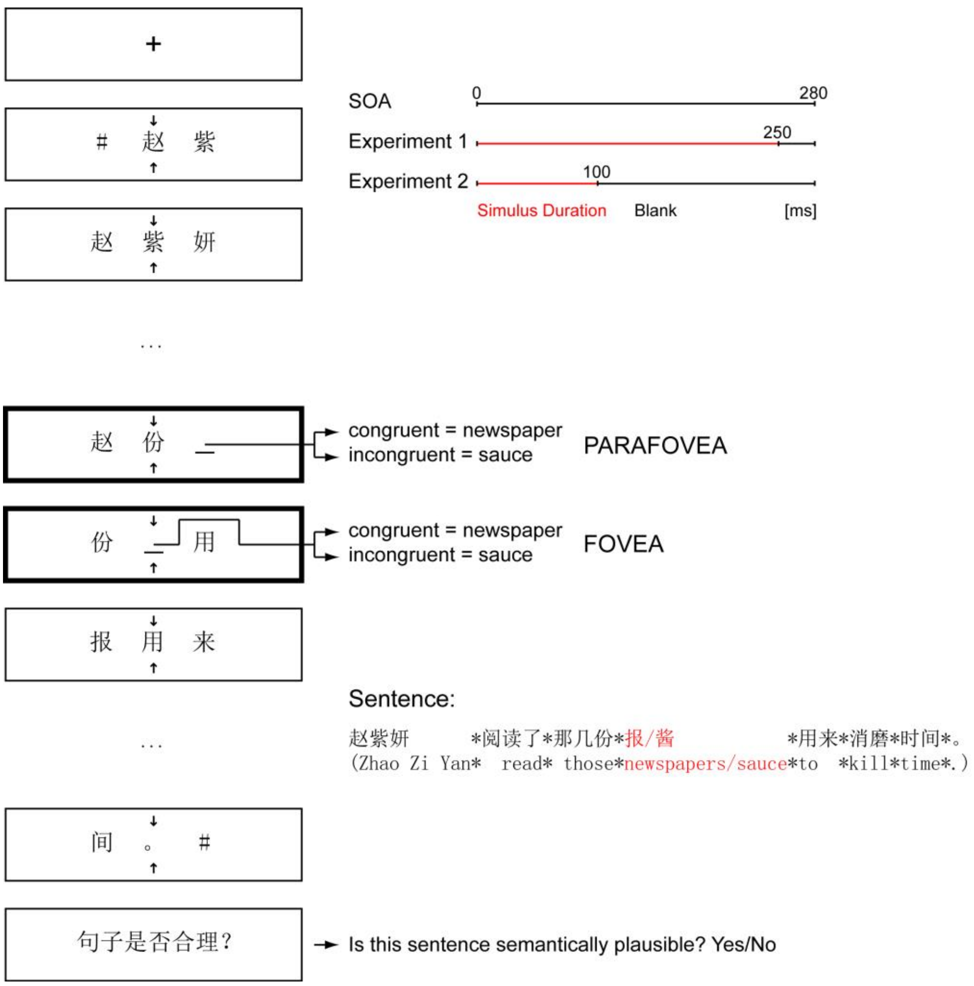

Figure 1. Trial scheme for Experiments 1 and 2. Sentences were presented character-by-character in the center of the screen, flanked by the preceding character to the left and the subsequent character to the right. Following the design by Barber et al. 2013, one target word in the sentence was presented either in its congruent (here: *newspapers*) or its incongruent version (here: *sauce*) while in the parafoveal and while in the foveal position, yielding a 2 × 2 design (*parafoveal congruency* × *foveal congruency*. The target word always consisted of a single Chinese character. Please note that the interaction between parafoveal and foveal congruency also determines the *validity* of the preview, that is, whether or not the reader obtains a parafoveal preview on the word later seen in the foveal position (*preview validity*). Participants were asked to keep their gaze fixated on the central character and to judge the plausibility of the sentence with a button press. The two experiments differed only in terms of the stimulus duration and ISI, while the speed of presentation (SOA) remained constant. The ”…” indicates that other words in the sentence were presented in the same fashion.

Example sentence with the four conditions:

赵紫妍*阅读了*那几份*报/酱*用来*消磨*时间*。

(Zhao Zi Yan* read*those* newspapers/sauce* to* kill* time*.)

(1) parafovea congruent, fovea congruent: newspapers, newspapers [valid preview]

(2) parafovea incongruent, fovea congruent: sauce, newspapers [invalid preview]

(3) parafovea congruent, fovea incongruent: newspapers, sauce [invalid preview]

(4) parafovea incongruent, fovea incongruent: sauce, sauce [valid preview]

Note: the preview validity effect is mathematically identical to the interaction between parafoveal and foveal congruency.

To facilitate central fixation, two small vertical arrows, one above and one below the central word, served as visual markers for fixation (see Figure 1). After each sentence, participants were prompted to decide whether the sentence was semantically congruent or incongruent by pressing the left or right mouse button. The congruency decision was based solely on the foveal status of the target word. That is, if the target word was incongruent during parafoveal presentation and congruent during foveal presentation, the sentence was considered semantically congruent and vice versa. If gaze had left the fixation area during the trial, a screen appeared after the answer, informing participants about this and admonishing to fixate better in the following trials. There was a self-terminated pause after every two sentences. All characters were presented in simple black Kai font on a white background. Each character was about 1.8×1.8 cm in size, and subtended 1.62° of visual angle at the viewing distance of 60 cm. Characters were presented with one empty character space between them. Thus, the left edge of the right parafoveal flanker character was presented 2.43° (1.5 character spaces × 1.62°) away from the screen center.

Data analysis. Trials in which participants gave a wrong answer and/or their gaze left the fixation area during the pre-target or target presentation were excluded from ERP analyses. Epochs were cut around the parafoveal target word onset, from -200 to 1100 ms, and baseline-corrected by subtracting a 100 ms pre-stimulus baseline. Following Barber et al. (2013), the same baseline interval (before the parafoveal presentation) was also used to analyze ERPs evoked by the foveal presentation. This ensures that the earlier N400 effects of parafoveal congruency do not influence the baseline correction for the analysis of the following foveal ERP. Epochs in which the maximum voltage difference at any channel exceeded 120 μV were excluded. Overall 19% of the data were excluded. As we had specific a priori hypotheses about the timing and locus of the effects, we analyzed two regions of interest (ROIs), each consisting of four electrodes. A ROI of four central-posterior locations was used to investigate N400 effects (Cz, Pz, CP1, CP2) and one of four occipito-temporal sites was used to analyze a possible preview positivity (PO9, PO7, PO8, PO10). The time windows used for analysis were 300-500 ms after parafoveal and foveal word onset respectively for the congruency effects and 200-300 ms after foveal word onset for the preview positivity.

We performed linear mixed effect model (LMM) analyses using the *lmer* function of the *lme4* package (Bates, Maechler, & Dai, 2008), supplied in the R system for statistical computing (version 3.1.1, R Development Core Team, 2010). Linear mixed models provide two major advantages over traditional ANOVA approaches that make them well suited for our needs: First, they allow specifying participants and items as crossed random factors, which allows using only one model instead of the traditional F1/F2 ANOVA approach (Forster & Dickinson, 1976). Second, LMMs can handle imbalanced datasets with differing amount of data in the design cells without losing statistical power (Pinheiro & Bates, 2000). This is advantageous, since data loss due to eye movement and EEG rejection criteria might not be equally distributed over the design cells.

We report fixed effect regression weights (*b*'s), as well as standard errors of these estimates (SE). For statistical testing, we report only the *t*-values and no associated *p*-values, as LMMs do not have a clear way of assessing the degrees of freedom. There are still several ways of assessing the significance of an effect. For large datasets like ours it is common to interpret absolute *t*-values larger than 2.0 as significant as the *t*-distribution approaches the normal distribution for a large amount of observations (Baayen, 2008)

We modeled mean ERP amplitudes in the specified ROI's and time windows. We defined two fixed factors (parafoveal congruency, foveal congruency). The interaction of these two factors is mathematically equivalent to a main effect of preview validity, which can be seen by looking at the contrast vectors for the four conditions. The parafoveal congruency effect contrasts conditions (1) and (3) with conditions (2) and (4). Accordingly, the contrast vector looks like this: (1, -1, 1, -1). For the foveal congruency contrast, which compares the conditions (1) and (2) vs. (3) and (4), the vector is: (1, 1, -1, -1). Multiplying these two vectors gives their interaction: (1, -1, -1, 1). Conditions (1) and (4) are the ones with valid preview, while (2) and (3) are the ones with invalid preview.

Thus, we have a 2 × 2 design, which can be reconceptualized, on theoretical grounds, as three orthogonal main effects. These main effects all have the same statistical power as they involve contrasting two conditions with the two remaining conditions (see Shaffer, 1977 and Brauer & Judd, 2000, for the reconceptualization of interactions as main effects; see Kliegl, Mayr, Junker, & Fanselow, 1999 and Risse & Kliegl, 2014 for applications of this logic).

We included participants and items as crossed random factors. Following recent suggestions by Barr et al. (2013) we chose a maximum random effects structure and allowed intercepts and slopes to vary over those random factors. If certain random effects could not be identified they were removed from the model. In those cases likelihood ratio tests justified the simpler model even at a liberal alpha level (.20). If the maximum random effects model did not converge, we resorted to a model without random effect correlations. The *lmer* specification of the model was as follows:

```
lmer(ERP ~ Parafoveal_Congruency * Foveal_Congruency + (
Parafoveal_Congruency * Foveal_Congruency |Subject) +
(Parafoveal_Congruency * Foveal_Congruency |Item)).
```

Note that a statistical term for the congruency of the foveal word was also included in the model for the N400 elicited by the parafoveal word presentation. The reason for including this term is that the foveal word is already visible during the N400 time window following the parafoveal presentation (i.e., 300-500 ms after parafoveal word onset corresponds to 20-220 ms after foveal word onset). It is therefore theoretically possible that the status of the foveal word affects the amplitude in the parafoveal N400 window. Although this was not the case (see *Results*), we included all factors in all models.

We analyzed the manual response at the end of each trial in terms of accuracy. In addition, we computed the proportion of trials in which the participant's gaze deviated more than ±1.4° from central fixation. These bad fixation trials were excluded from all ERP analyses. Accuracy (correct/incorrect) and bad fixation trials (good fixation/bad fixation) were coded as binary response variables with a binomial distribution and entered into generalized LMMs. We report the log odds ratio estimate (logit) z-values and corresponding p-values.

Results

Behavioral results. On average, participants correctly answered the plausibility question after 90% of the trials. Response accuracy was significantly lower in the parafoveal incongruent - foveal congruent condition (79%) than in the other three conditions ($logit$ = -1.46, z = -9.7, $p < 0.001$). Six percent of all trials were excluded due to gaze being outside of the foveal area.

ERPs: Parafoveal and foveal congruency effects. There was a significant N400 congruency effect of the parafoveal target word in the central-posterior ROI between 300–500 ms after the parafoveal target word onset (Figure 2, top left). Incongruent parafoveal words elicited a more negative waveform than congruent parafoveal words, b = 0.89, SE =0.20, t = 4.3. There were no effects of foveal congruency or preview validity in this early time window.

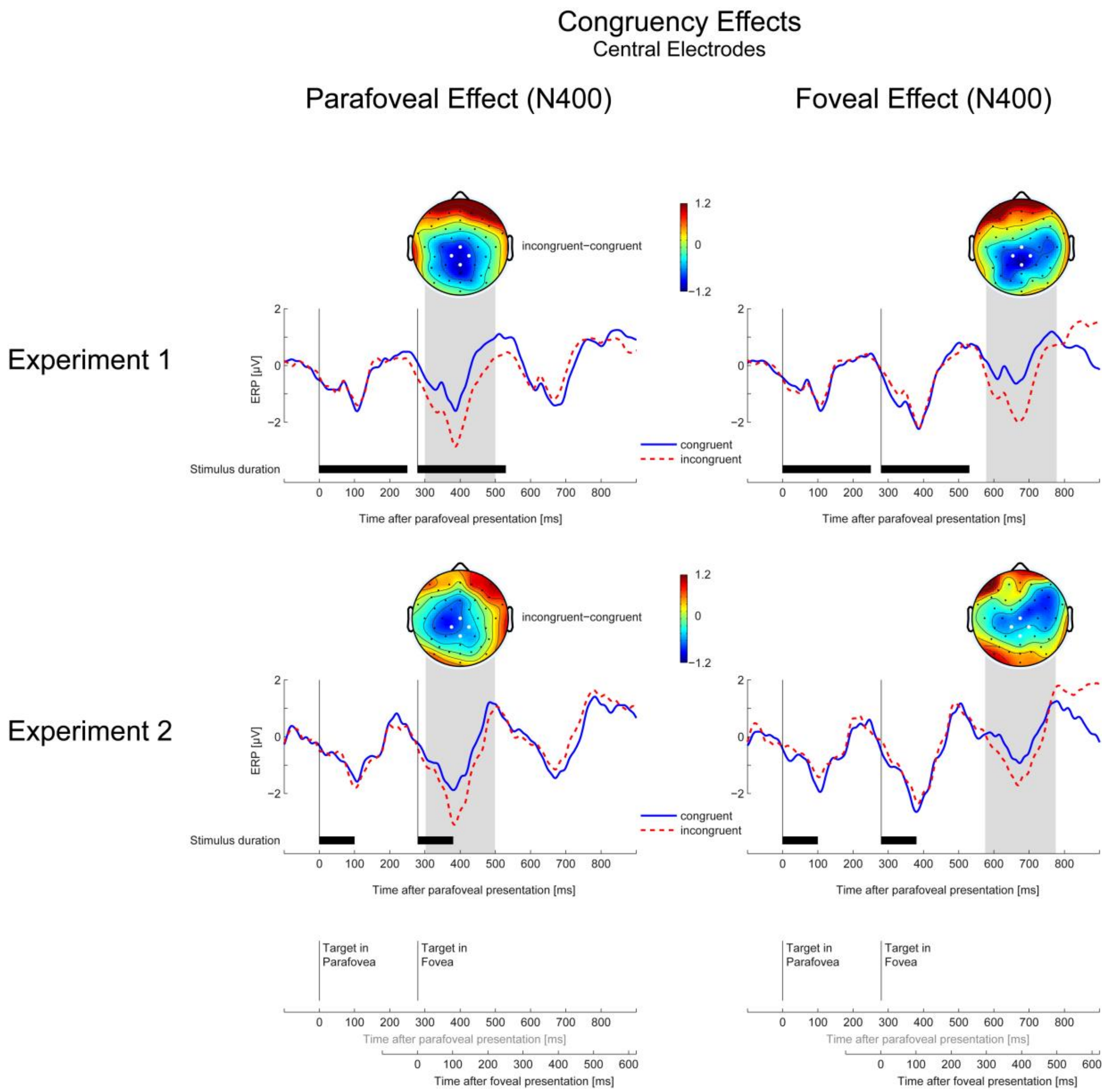

Figure 2. Congruency effects on the N400 component in Experiment 1 and 2. Time 0 on the time axis marks the onset of the character triad containing the target word character in the right parafoveal position. Lines show the average activity of a ROI of four electrodes, highlighted in white in the topographic scalp maps. Scalp maps show the difference in ERP activity (incongruent minus congruent) in the time window of 300-500 ms after the onset of the word in the parafoveal and foveal position. Black rectangles above the time axis mark the stimulus durations of the character triads. The vertical line marks the onset of the triad containing the target word in the fovea.

Similarly, there was also a main effect of congruency of the foveal target word at central-posterior electrode sites 300-500 ms after the foveal onset of the target word, $b = 1.03$, $SE = 0.23$, $t = 4.5$ (Figure 2, top right).

ERPs: Preview validity effects. As shown in Figure 3 (middle left panel), for the occipito-temporal ROI 200-300 ms after the foveal target word onset, we found a preview validity effect. At these electrodes, an invalid parafoveal preview (incongruent after congruent, or congruent after incongruent) elicited a more negative N1 component than a valid, identical preview word ($b = 1.53$, $SE = 0.54$, $t = 2.8$).

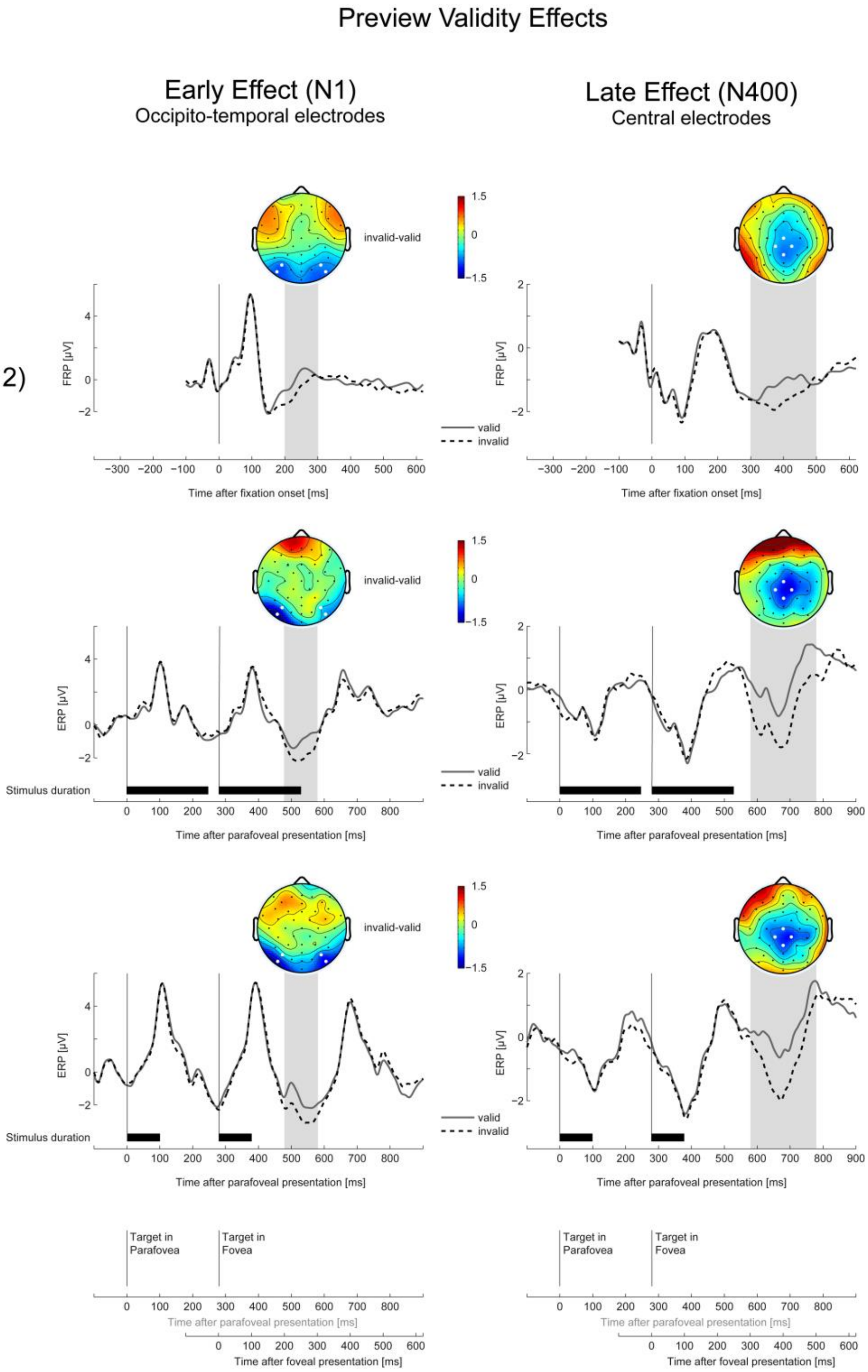

Figure 3. Preview validity effects in the ERP of Experiment 1 and 2 in comparison to prior findings in fixation-related potentials (FRPs). *Top panel*: Preview validity effects in FRPs from Dimigen et al. (2012). Time 0 marks the onset of the first fixation on a target word that was either visible (identical preview) or masked by an unrelated word (invalid preview) during the preceding fixation. The four occipito-temporal electrodes contributing to the depicted average in the time series are highlighted in white in the topographies. The topographical map shows the difference in the time window indicated in grey (200-300 ms for the early effect and 300-500 ms for the late effect). *Middle and bottom panel*. Preview validity effects in Experiment 1 and 2. Black rectangles indicate the presentation durations of the word triads. The target character appeared in the parafoveal position at 0 ms and in the foveal position at 280 ms. Black rectangles above the time axis mark the stimulus durations of the character triads. The vertical lines mark the onset

of the triad containing the target word in the parafovea and fovea respectively. In both experiments, the N1 of the ERP evoked by the foveal target presentation was markedly attenuated if the same character had been previously seen in the parafovea. As in FRPs, this global *preview positivity* occurred 200-300 ms after the target word entered foveal vision (480-580 ms relative to the parafoveal presentation, see grey area). The late preview validity effect occurred 300-500 ms after the target word entered foveal vision (580-780 ms relative to the parafoveal presentation).

Finally, we observed a main effect of preview validity ($b = 2.28$, $SE = 0.46$, $t = 5.0$) in the foveal N400 time window. Conditions with invalid preview yielded more negative amplitudes than conditions with valid preview (Figure 3, middle right panel).

Discussion

Experiment 1 yielded several main results. The first was a robust N400 effect for parafoveally presented incongruous words. This finding replicates Barber et al. (2013) and generalizes their findings to Chinese script reading at a natural, reading-like stimulus timing. Importantly, this parafoveal effect was observed despite rigorous fixation control, showing that N400 effect for parafoveal words are not due to undetected gaze shifts. More generally, our finding that participants left the central fixation region in only a small fraction (6%) of the trials shows that they were able to maintain a precise central fixation even with longer word presentations of 250 ms that resemble the average duration of eye fixations during natural reading. It is important to note, however, that participants received fixation training with immediate feedback before the experiment and delayed feedback after each trial. This might have reduced the amount of lateral eye movements.

In the ERP time-locked to the foveal target presentation, we observed two robust effects of preview validity: Replicating findings by Dimigen et al. (2012), we found that between 200-300 ms, the N1 component evoked by the foveal word presentation was significantly weaker (more positive) after a valid preview. It is important to note that Dimigen et al. used a more natural reading situation with eye movements and measured fixation-related potentials instead of ERPs, but also used a somewhat more unnatural list reading task. The current study is therefore the first to demonstrate a preview validity effect on the wave shape of traditional stimulus-locked ERPs during sentence reading. This result suggests that the preview positivity effect in the EEG generalizes across paradigms (natural reading vs. RSVP) and writing systems (German vs. Chinese). Moreover, we found that the amplitude of the N400 was significantly smaller (more positive) for words after a valid than after an invalid preview. As shown in Figure 3, this effect resembled the late preview benefit reported in Dimigen et al. (2012). However, while the effect was only marginally significant in their study, it was clearly significant here and of similar amplitude as the effect of foveal congruency in the same N400 time window.

While these results fit nicely to those obtained under very different reading conditions by Dimigen et al. (2012), they raise the question why Barber and colleagues did not observe similar influences of preview validity, despite many similarities in the experimental design. One possible explanation is that Chinese readers take up more information from parafoveal words. A second possibility is that the more natural word presentation durations in our experiment (250 ms instead of 100 ms) afforded a more thorough preview on the upcoming word. A third and closely related explanation concerns the resulting difference in the ISI. The ISI was very short in the present experiment (30 ms), but much longer (150 ms) in Barber et al. (2013). Dimigen et al. (2012) proposed that the preview positivity reflects a form of partial repetition priming that is primarily based on abstract orthographic information retrieved from the initial letters of the upcoming word. These priming effects in the EEG may simply decay during an unnaturally long ISI. This hypothesis was tested in Experiment 2.

Experiment 2

Experiment 2 tested whether the influences of preview validity would be modulated by a difference in stimulus timing in Chinese sentence reading. For this purpose, we repeated Experiment 1 but adapted stimulus durations to the 100 ms used by Barber and colleagues. This also resulted in a longer ISI of 180 ms. The SOA was kept constant. We hypothesized that effects of preview validity on the ERP time-locked to the foveal presentation (preview positivity and late preview effect) should be absent with a long ISI.

Method

Participants. Sixteen Mandarin-speaking Chinese exchange students (8 women, age 19 to 30; M = 24.0 years), different from those in Experiment 1, participated in Experiment 2. All had normal or corrected to normal vision and were naive with regard to the purpose of the experiment. Written informed consent was obtained from all participants. Fifteen participants were right-handed and one was left-handed (laterality quotient: M = +86). The data of one participant was removed from analysis due to low response accuracy (60%).

Materials and apparatus. Material and apparatus were identical to Experiment 1.

Procedure. The procedure was identical to Experiment 1 except that character triads were presented for 100 ms, with an intervening blank interval of 180 ms.

Analysis. Data analyses were identical to Experiment 1, but we performed additional statistical tests to compare the size and duration of effects in Experiments 1 and 2. We ran an LMM on the joint data of both experiments and included *Experiment* (2) and all its second order interactions with experimental factors as predictors. The LMMs were otherwise identical to the ones described above. *Experiment* was not included in the random effects structure, as it was a between-subjects factor and there were no theoretical grounds to expect any variation by item.

To test for possible changes in onset latency and effect duration, we conducted an additional fine-grained analysis: ERP averages were computed at the two ROIs in consecutive 30 ms time windows between 200 ms and 830 ms after the parafoveal onset of the target. This yielded 21 windows which covered all relevant effects. For each time window, we then ran a repeated-measures ANOVA on the factors *Parafoveal Congruency* (2) × *Foveal Congruency* (2) × *Experiment* (2, between-subjects factor). Degrees of freedom were adjusted by multiplication with Huynh-Feldt's epsilon (Huynh & Feldt, 1976). To correct for multiple comparisons across time windows, the resulting *p*-values were further corrected by applying the false discovery rate procedure by Benjamini and Hochberg (1995) as implemented by Groppe, Urbach, and Kutas (2011).

Results

Behavioral results. The mean percentage of correct answers was 93%. Again, response accuracy in the incongruent – congruent condition (89%) was lower than in the other three conditions (*logit* = -0.81, $z = -4.6$, $p < 0.001$). Eight percent of all trials were excluded due to gaze recordings outside the fixation area, but the occurrence of lateral eye movements did not differ between conditions. Despite the change in stimulus duration, neither the accuracy of manual responses nor the amount of bad gaze trials differed significantly between the two experiments.

ERPs: Parafoveal and foveal congruency effects. As in Experiment 1, there was a significant main effect of parafoveal congruency on the N400 time-locked to the parafoveal target presentation (Figure 2, bottom left). Incongruent parafoveal words elicited a more negative response than congruent words ($b = 0.57$, $SE = 0.23$, $t = 2.5$). As in Experiment 1, foveal congruency and preview validity did not influence centroparietal ERP amplitude in this parafoveal N400 window. In addition, as in Experiment 1, we also observed a main effect of foveal congruency on the N400 time-locked to the foveal presentation of the target word ($b = 0.70$, $SE = 0.30$, $t = 2.4$; Figure 2, bottom right).

ERPs: Preview validity effects. Once the target word entered foveal vision, previewed words elicited a more positive N1 between 200-300 ms than words after an invalid preview ($b = 1.34$, $SE = 0.63$, $t = 2.1$). Amplitude and scalp distribution of this preview positivity were similar to Experiment 1 (Figure 3, bottom left). As in Experiment 1, there was also a significant late preview validity effect on the N400 at central-posterior electrodes ($b = 2.08$, $SE = 0.59$, $t = 3.5$, Figure 3, bottom right).

Comparison to Experiment 1. LMM analysis revealed no significant effects of the factor *Experiment*. There was no interaction of the parafoveal congruency effect with experiment ($b = 0.16$, $SE = 0.15$, $t = 1.1$). The foveal congruency effect also did not interact with experiment in the foveal N400 window ($b = 0.18$, $SE = 0.16$, $t = 1.1$). The preview effect also showed no interaction with experiment, neither for the early effect at PO9 ($b = 0.07$, $SE = 0.19$, $t = 0.4$) nor for the late effect at Cz ($b = 0.12$, $SE = 0.16$, $t = 0.8$). No other interaction terms were significant (all t-values < 1.7).

An additional fine-grained, window-wise analysis with ANOVAs also produced no significant ERP differences between experiments, neither in terms of effect amplitude nor in terms of effect duration. Although both the parafoveal and the foveal N400 effects and their interactions appear numerically larger in Experiment 1 (compare Figure 4), the corresponding statistical trends did not survive correction for multiple comparisons (all *p*-values > 0.64).

Figure 4. Comparison of effects in Experiment 1 and 2. *Top panels*: Congruency effects. Left: Parafoveal congruency effect. Right: Foveal congruency effect. *Bottom panels*: Preview validity effects. Left: Early preview validity effect. Right: Late preview validity effect. Shown are the difference curves (incongruent minus congruent/ invalid minus valid) of the effects. Time 0 marks the onset the character triad containing the target word character in the parafoveal position. Electrodes plotted are the same averages as in Figure 2 (Cz, CP1, CP2, Pz) for the congruency effects and the late preview effect and the same as in Figure 3 on the left side (PO9, PO10, PO7, PO8) for the early preview effect. The boxes above the time axis indicate the stimulus presentation duration and ISI in each experiment, color-coded by experiment. Vertical lines mark the onset of the target word in the parafoveal and foveal location, respectively.

Discussion

We replicated a very robust parafoveal N400 effect in Chinese readers. Importantly, simultaneous eye tracking excludes the possibility that this effect is a consequence of fixations outside the central screen area. However, even though Experiment 2 closely resembled Barber et al.'s design, the results concerning the ERP effects of preview validity were still different: Neither the early preview positivity nor the late preview effect vanished at the longer ISI. Instead, statistical analyses comparing the two experiments suggest that the effects were not even attenuated relative to Experiment 1. While this result speaks to the robustness of the preview effects, it leaves us with the question why Barber et al. did not observe a similar pattern. We will return to this question in the *General Discussion*.

Our results also suggest that the brief word presentations of only 100 ms were not necessary from a methodological point of view, because the proportion of trials with bad fixations did not differ between experiments. We believe that this is due to the fact that our participants received fixation training and feedback during the experiment. We will expand on this point below and advocate for the use of thorough fixation control in the RSVP-with-flanker paradigm.

General Discussion

We report data from two experiments, which combined the RSVP-with-flankers paradigm with concurrent eye-tracking. Participants read Chinese sentences containing a one-character target word that was presented either in its congruent or incongruent version when it first appeared in the parafovea and when it moved into fovea vision. In Experiment 1, we used presentation times that resembled the durations of fixations and saccades during normal reading. In Experiment 2, we kept the overall presentation speed (SOA) constant, but presented the words for the same short durations as Barber et al. (2013) with longer blank intervals in-between.

Our first finding was a robust N400 congruency effect for parafoveally presented Chinese words, irrespective of stimulus timing. Second, this result was unrelated to lateral eye movements. Third, the ERP waveforms evoked by the subsequent foveal target presentation were not only impacted by the word's congruency with the sentence, but also by the validity of the preview, leading to a markedly attenuation of the occipito-temporal N1 and the following N400 component after valid previews. In the following these results are discussed in turn.

A robust parafoveal N400 effect in Chinese

Our finding of a parafoveally-triggered N400 effect is in line with three previous studies, which have reported such an effect (Barber et al. 2010; 2013; Kretzschmar et al., 2009). The current results demonstrate that parafoveal N400 effects generalize from alphabetic languages to the logographic Chinese script. The fact that the congruency of a not-yet-fixated word modulates the N400 shows that readers retrieved sufficient information about the upcoming word to determine its semantic fit into the sentence. Based on a model of spreading activation, Kretzschmar et al. (2009) proposed that high predictability sentences allow the reader to determine the contextual fit of an upcoming word based on (partial) visual-orthographic information (e.g. the word-initial letters). This crucial role of contextual constraint for the occurrence of parafoveal N400 effects at normal reading speeds is also suggested by the results of Barber et al. (2013), who did not observe a parafoveal N400 at a short SOA with sentences of low contextual constraint. If parafoveal N400 effects really depend on a matching process between expected words and partial orthographic information obtained from the parafovea, this might also explain why Dimigen et al. (2012) found no EEG effect of semantically associated previews in a reading situation in which readers were presented with a context-free lists of words that afforded no predictions about the upcoming word.

However, it is of course also possible that the processes triggering a parafoveal N400 are different in Chinese readers, given that they seem to process parafoveal information to a larger extent than readers of Western script (Yan et al., 2009; Yan, Risse, et al., 2012; Yan, Zhou, et al., 2012). It is therefore possible that Chinese readers can process parafoveal incongruency even in the absence of contextual predictability. However, we did not include a low-constraint condition in the present study and the range of contextual predictabilities in our sentences was higher than that in the low-constraint condition of Barber et al. (2013). Future studies using the RSVP-with-flankers paradigm could investigate whether readers of Chinese can access the meaning of upcoming words in sentences without any contextual constraint.

Control of lateral eye movements

We used an eye-tracker to ensure that the parafoveal N400 is not a trivial consequence of gaze shifts towards the parafoveal location. Results clearly demonstrate that this was not the case. Irrespective of stimulus duration (250 ms or 100 ms), we detected only relatively few lateral eye movements while the target was on the screen. However, since our participants were trained in precise fixation and also received feedback during the experiment, we cannot infer from these observations on the frequency of eye movements in previous studies without eye-tracking. Importantly, short stimulus durations only prevent eye movements if a single word-triad is presented on each trial. If several triads are presented in succession – as in sentence reading – participants can move their gaze rightwards at an earlier point during trial (possibly with several small, involuntary saccades), and continue reading from there. For this reason, we believe it is necessary to control for eye movements in this paradigm. One option to monitor absolute gaze position are carefully calibrated EOG recordings with a DC amplifier. However, with recent technical and methodological advances in co-registering EEG and eye-tracking (e.g. see Dimigen et al., 2011 for a review), this combination provides a better control.

A preview effect on the early ERP waveform: the preview positivity

While the general identity preview benefit is firmly established in the eye-movement literature (Schotter, Angele & Rayner, 2012), it is almost unexplored in the EEG. In both experiments, we found a marked attenuation of the late parts of the N1 component after valid previews (preview positivity). This effect was first shown by Dimigen et al. (2012) in fixation-related potentials using the boundary paradigm during saccadic reading. However, one potential problem with the boundary paradigm is that it involves a physical change of the display during the saccade only in the invalid preview conditions (when the invalid preview word changes to the correct foveal word), but not in the valid preview condition. Therefore, differences in the FRP might be, partly or purely, a low-level artifact of the stimulation technique. This study demonstrates and replicates a preview validity effect also in stimulus-locked ERPs. In the RSVP-with-flankers paradigm, the display changes both in the valid and the invalid preview condition, because the target always moves from the parafoveal to the foveal location. Thus, the effect is not a low-level artifact, but really driven by parafoveal processing.

The exact nature of the underlying facilitation is not yet clear. Based on the eye-tracking literature, Dimigen et al. (2012) proposed that the attenuation of the N1 component by preview reflects priming at the level of abstract (i.e. letter case-independent) orthographic representations and/or phonological representations with an especially important role of the initial letters/phoneme of the upcoming word (at least in alphabetic languages). The preprocessing of lower-level visual features may also contribute to the effect. Dipole modeling indicates that the basis for the N1 effect is an activation difference in occipito-temporal cortex for previewed words (Dimigen et al., 2012). These explanations are of course tentative and more sophisticated preview manipulations (e.g. phonologically related but orthographically different previews, previews with a change of letter case during the saccade) are necessary to disentangle the contributing factors.

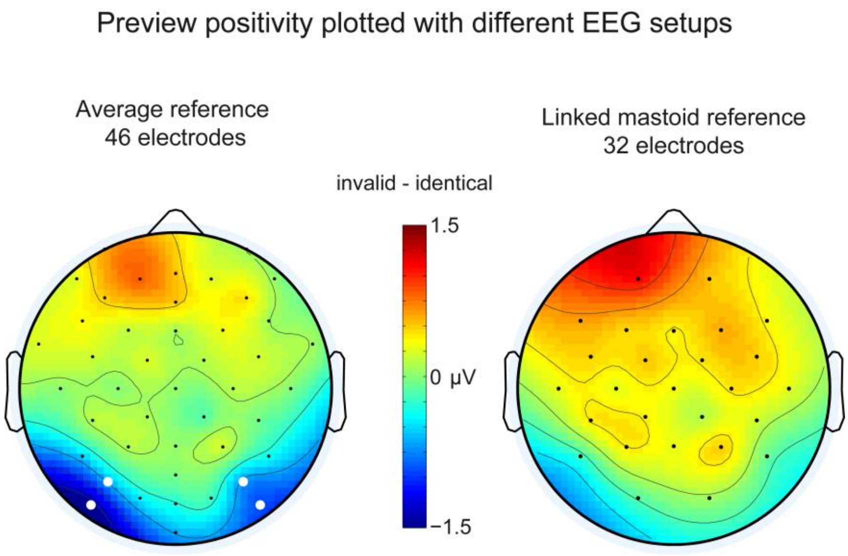

Figure 5. Effect of electrode setup and reference electrode on the measurement of the preview positivity. *Left*: Preview positivity (invalid minus identical preview, averaged across the data of Experiment 1 and 2), plotted with 46 electrodes and an average reference montage. Electrodes highlighted in white indicate the ROI used in the present study. *Right*: Same data, but now plotted at the 32 scalp electrodes used by Barber et al. (2013) and with a linked mastoid reference. With the latter setup, the effect is diffusely spread out to frontal electrodes and virtually absent over occipito-temporal brain regions.

Why was a preview positivity observed in none of the experiments by Barber and colleagues (2010, 2013)? Since we found the effect in both experiments, we can rule out that the difference is explained by stimulus timing and the duration of the inter-stimulus interval. Instead, we resorted to a methodological comparison. In the current study, we used a 46 electrode montage, which is a superset of the 32 electrodes used by Barber et al. This montage also includes sensors over inferior occipito-temporal regions that are located near the scalp projections of the preview positivity's equivalent dipoles (as estimated by Dimigen et al., 2012). Furthermore, we used an average reference whereas Barber and colleagues referenced against linked mastoids. Because the mastoid bones are located close to occipito-temporal sites, electrodes at these locations will capture the preview positivity to some extent, thereby attenuating it and spreading it with opposite polarity to all other electrodes. To test whether differences in electrode montage can explain the divergent results, we reanalyzed our data using 32 electrodes and a linked mastoid reference. The resulting scalp topographies, shown in Figure 5, strongly suggest that without coverage of the entire scalp and with a mastoid reference, focal effects at occipito-temporal sensors, like the preview positivity, may be missed. We believe that these results exemplify the importance of including inferior electrodes in studies on visual word recognition, which are located closer to ventral-stream areas implicated in early visual-orthographic processing (Cohen et al., 2000; McCandliss, Cohen, & Dehaene, 2003).

Foveal N400 amplitude reflects semantic congruency and preview validity

Results were more complicated for the N400 component following the foveal target presentation. Here, we observed both the classic main effect of foveal congruency (Kutas & Hillyard, 1980) and a significant effect of preview validity. This result is compatible with the notion that foveal congruency and preview validity contribute together to the measured N400 amplitude. According to this interpretation, the late preview effect (observed as a trend in Dimigen et al., 2012) would most likely reflect a form of repetition priming, which has been shown to produce similar N400 modulations (e.g., Holcomb, & Grainger, 2006; 2007). Unrepeated words (analogous to our invalid preview condition) yield more negative amplitudes than repeated words (analogous to our valid preview condition). Based on the eye-tracking literature and the similarity of our findings to those reported by Dimigen et al. (see Figure 3), we believe that this is the most parsimonious interpretation of the results.

However, as explained in the *Introduction*, in Barber et al.'s design, the factor *preview validity* is implicitly manipulated and mathematically equivalent to the interaction between the congruency of the parafoveal word and the foveal word (see section *Data analysis* in the *Methods*). Using this design, we therefore cannot fully exclude the possibility that this statistical interaction in the foveal N400 window is due to foveal congruency processing being generally different after the processing of a congruent versus an incongruent word in the parafovea. For example, the incongruent word in the parafovea might cause difficulty for the subsequent processing of a foveally congruent word, while a congruent word in the parafovea might not cause difficulties during the subsequent processing of a foveally incongruent word. However, we did not find empirical evidence from previous research in support of such an interaction. In contrast, preview validity effects have been reported reliably in numerous eye movement studies and the pattern of preview validity effects found in ERPs in this study was highly consistent with that reported by Dimigen et al. (2012) in FRPs. We therefore tend to interpret the observed effects in the foveal N400 window as an effect of preview validity, just like the earlier N1 effect. Clearly, future studies are needed to further examine the interaction between parafoveal and foveal congruency at the level of the N400.

Regardless of the theoretical interpretation, the question remains why such late preview validity effect showed up in our study but was absent in Barber et al. (2013). In contrast to the preview positivity discussed above, this is not easily explained by differences in the electrode montage. Current studies on parafoveal processing differ not only in terms of the language used, but also with regard to the paradigm (RSVP-with-flankers, FRPs), reading mode (sentence reading, list reading), task (comprehension questions in Barber et al. vs. plausibility judgments in the present study), SOA, stimulus durations, and electrode setup and it is necessary to explore the impact of each of these design choices.

Conclusions

By combining the RSVP-with-flankers paradigm with rigorous fixation control we were able to show strong and consistent parafoveal N400 effects of semantic congruency in Chinese readers independently of stimulus timing. Following up on recent FRP results, we also demonstrate that the wave shape of stimulus-locked ERPs differs under reading conditions that afford a normal preview on the upcoming word. Once a previewed word reaches foveal vision, this preprocessing attenuates the N1 component and likely also the N400 component of the ERP. Together, these results underline the importance of considering parafoveal vision in electrophysiological studies on reading.

References

Altarriba, J., Kambe, G., Pollatsek, A., & Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish–English bilinguals. *Perception & Psychophysics, 63*, 875–890. doi: 10.3758/BF03194444

Baccino, T. & Manunta, Y. (2005). Eye-fixation-related potentials: Insight into parafoveal processing. *Journal of Psychophysiology, 19,* 204-215.

Baayen, R. H. (2008). *Analyzing linguistic data* (Vol. 505). Cambridge, UK: Cambridge University Press

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17*(3), 364-390. doi: 10.1016/0010-0285(85)90013-1

Barber, H. A., Doñamayor, N., Kutas, M., & Münte, T. (2010). Parafoveal N400 effect during sentence reading. *Neuroscience Letters*, *479*, 152–156. doi: 10.1016/j.neulet.2010.05.053

Barber, H. A., van der Meij, M., & Kutas, M. (2013). An electrophysiological analysis of contextual and temporal constraints on parafoveal word processing. *Psychophysiology, 50*(1), 48-59. doi: 10.1111/j.1469-8986.2012.01489.x

Bates, D., Maechler, M., & Dai, B. (2008). The lme4 package. *Computer software manual. Retrieved from http://cran.r-project.org/web/packages/lme4/*

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological, 57*(1), 289-300. doi: 10.2307/2346101

Brauer, M. & Judd, C.M. (2000). Defining variables in relationship to other variables: When interactions suddenly turn out to be main effects. *Journal of Experimental Social Psychology*, 36, 410-423

Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275-307). New York: Elsevier.

Chen, H.-C., Song, H., Lau, W. Y., Wong, K. F. E., & Tang, S. L. (2003). Developmental characteristics of eye movements in reading Chinese. In C. McBride-Chang & H.-C. Chen (Eds.), *Reading development in Chinese children* (pp. 157-170). London: Praeger Publisher.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., & Michel, F. (2000). The visual word form area - Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123*, 291-307. doi: 10.1093/brain/123.2.291

Dambacher, M., Dimigen, O., Braun, M., Wille, K., Jacobs, A. M., & Kliegl, R. (2012). Stimulus onset asynchrony and the timeline of word recognition: Event-related potentials during sentence reading. *Neuropsychologia, 50*(8), 1852-1870. doi: 10.1016/j.neuropsychologia.2012.04.011

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9-21. doi: 10.1016/j.jneumeth.2003.10.009

Dimigen, O., Kliegl, R., & Sommer, W. (2012). Trans-saccadic parafoveal preview benefits in fluent reading: A study with fixation-related brain potentials. *Neuroimage, 62*(1), 381-393. doi: 10.1016/j.neuroimage.2012.04.006

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of Eye Movements and EEG in Natural Reading: Analyses and Review. *Journal of Experimental Psychology-General, 140*(4), 552-572. doi: 10.1037/a0023885

Dimigen, O., Valsecchi, M., Sommer, W., & Kliegl, R. (2009). Human microsaccade-related visual brain responses. *Journal of Neuroscience*, 29, 12321-12331.

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language, 115*(3), 149-161. doi: 10.1016/j.bandl.2010.07.006

Forster, K. I., & Dickinson, R. (1976). More on the language-as-fixed-effect fallacy: Monte Carlo estimates of error rates for *F1*, *F2*, *F'*, and *min F'*. *Journal of Verbal Learning and Verbal Behavior, 15*(2), 135-142. doi: 10.1016/0022-5371(76)90014-1

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711-1725. doi: 10.1111/j.1469-8986.2011.01273.x

Hohenstein, S., & Kliegl, R. (2013). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 166-190

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic Preview Benefit in Eye Movements During Reading: A Parafoveal Fast-Priming Study. *Journal of Experimental Psychology-Learning Memory and Cognition, 36*(5), 1150-1170. doi: 10.1037/a0020233

Holcomb, P. J., & Grainger, J. (2006). On the time-course of visual word recognition: An ERP investigation

using masked repetition priming. *Journal of Cognitive Neuroscience*, *18*, 1631–1643. doi: 10.1162/jocn.2006.18.10.1631

Holcomb, P. J., & Grainger, J. (2007). Exploring the temporal dynamics of visual word recognition in the

masked repetition priming paradigm using event-related potentials. *Brain Research*, *1180*, 39–58. doi:10.1016/j.brainres.2007.06.110

Huynh, H., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational and Behavioral Statistics, 1*(1), 69-82. doi: 10.3102/10769986001001069

Inhoff, A. W., & Liu, W. M. (1998). The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology-Human Perception and Performance, 24*(1), 20-34. doi: 10.1037/0096-1523.24.1.20

Keren, A., Yuval-Greenberg, S., Deouell, L.Y. Saccadic spike potentials in gamma-band EEG: characterization, detection and suppression. Neuroimage, 49, 2248-2263. doi: 10.1016/j.neuroimage.2009.10.057

Kliegl, R., Hohenstein, S., Yan, M., & McDonald, S. A. (2013). How preview space/time translates into preview cost/benefit for fixation durations during reading. *Quarterly Journal of Experimental Psychology, 66*(3), 581-600. doi: 10.1080/17470218.2012.658073

Kliegl, R., Mayr, U., Junker, M. & Fanselow, G. (1999). Testing age invariance in language processes. In S. Kemper & R. Kliegl (Eds.), *Constraints on language: Aging, grammar,and memory* (pp. 137-167). Springer: New York. doi: 10.1007/0-306-46902-2_6.

Kretzschmar, F., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). Parafoveal versus foveal N400s dissociate spreading activation from contextual fit. *NeuroReport, 20*(18), 1613-1618.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences - brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205. doi: 10.1126/science.7350657

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences, 7*(7), 293-299. doi: 10.1016/s1364-6613(03)00134-7

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*, 578–587. doi: 10.3758/BF03203972

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113. doi: 10.1016/0028-3932(71)90067-4

Pernet, C., Uusvuori, J., & Salmelin, R. (2007). Parafoveal-on-foveal and foveal word priming are different processes: Behavioral and neurophysiological evidence. *Neuroimage, 38*(2), 321-330. doi: 10.1016/j.neuroimage.2007.07.035

Pinheiro, J., & Bates, D. M. (2000). *Mixed effects models in S and S-PLUS*. New York: Springer.

Rayner. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*(1), 65-81. doi: 10.1016/0010-0285(75)90005-5

Rayner. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin, 124*(3), 372.

Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology-Revue Canadienne De Psychologie, 40*(4), 473-483. doi: 10.1037/h0080111

Rayner, K., & Duffy, S. A. (1988). On-line comprehension processes and eye movements in reading. In M. Daneman, G. E. MacKinnon & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (pp. 13-66). New York: Academic Press.

Rayner, K., Schotter, E. & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit in reading revisited. *Psychonomic Bulletin & Review*, *21*(4), 1067-1072

Risse, S. & Kliegl, R. (2014). Dissociating Preview Validity and Preview Difficulty in Parafoveal Processing of Word n+1 during Reading. *Journal of Experimental Psychology: Human Perception and Performance*. 40 (2), 653-668. doi: 10.1037/a0034997

Schotter, E. R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language, 69*(4), 619-633. doi: 10.1016/j.jml.2013.09.002

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention Perception & Psychophysics, 74*(1), 5-35. doi: 10.3758/s13414-011-0219-2

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *Quarterly Journal of Experimental Psychology Section A Human Experimental Psychology, 52*(4), 1021-1046. doi: 10.1080/027249899390909