Differences in Parafoveal Processing of Inter-word and Intra-word Previews in Chinese: The Role of Inter-word Shadowing

Authors: Guan Yiyun, Song Xini, Zheng Yuwei, Zhang Yingliang, Cui Lei, Cui Lei

Date: 2019-05-07T00:00:00+00:00

Abstract

Abstract: To investigate the influence of lexical property features on parafoveal processing, the experiment manipulated whether the fixated character and the parafoveal character belonged to the same linguistic unit, using eye-tracking methodology combined with the boundary paradigm to explore its effect on preview effects. Additionally, the study examined the influence of word boundary information on parafoveal processing of inter-word and intra-word characters. Results showed that the preview effect for intra-word characters was greater than that for inter-word characters, with no differences in preview effects among the inter-word shadow, non-word shadow, and normal conditions. This indicates that word boundaries provided by shadows had no influence on the parafoveal processing of inter-word and intra-word characters, supporting the lexical segmentation and recognition model whereby lexical segmentation and lexical recognition occur simultaneously.

Full Text

Preview Processing Differences Between Between-Words and Within-Words in Chinese Reading: The Role of Inter-Word Shading

GUAN Yiyun[1], SONG Xini[1], ZHENG Yuwei[2], ZHANG Yingliang[1], CUI Lei[1]
([1] School of Psychology, Shandong Normal University, Jinan 250358, China)
([2] School of Education and Psychological Science, Jinan University, Jinan 250022, China)

Abstract

To investigate how lexical properties influence parafoveal preview processing, this study manipulated whether the fixated character and the preview character belonged to the same linguistic unit. Using eye-tracking methodology combined with the boundary paradigm, we examined how this factor affected preview benefits. Additionally, we investigated the influence of word boundary information on the preview processing of between-words and within-words. Results showed that preview benefits were larger for within-words than for between-words, with no differences in preview effects among normal, inter-word shading, and non-word shading conditions. These findings indicate that word boundaries provided by shading did not affect the preview processing of either between-words or within-words, supporting models of word segmentation and lexical recognition that posit these processes occur simultaneously.

Keywords: word boundary; between-words; within-words; preview processing; word segmentation
Classification Code: B482

1. Problem Statement

Reading is a complex cognitive task during which eye movements consist primarily of fixations and saccades. Readers acquire information mainly during fixations, while saccades serve to bring new information onto the fovea—the central 2° of the visual field. The region surrounding the fovea, extending approximately 2° to 5° from the center of vision, is known as the parafovea, where visual acuity declines significantly. Consequently, readers must move their eyes to project new textual content onto the foveal region (Rayner, 1998, 2009). Numerous studies have demonstrated that readers can obtain information not only from the fovea but also from the parafovea through preview processing, which accelerates reading and is termed the preview benefit (Rayner, 1998).

Preview processing substantially influences reading efficiency. Research shows that when preview words are masked and cannot be processed parafoveally, reading efficiency decreases significantly (Perea, Teiero, & Winskel, 2015; Pollatsek & Rayner, 1982). Accordingly, two major eye movement control models provide detailed descriptions and predictions of preview processing during reading. Serial models propose that attention is distributed sequentially, focusing on only one word at a time, with attention shifting to the next word only after processing of the currently fixated word is complete (Rayner, White, Kambe, Miller, & Liversedge, 2003). In contrast, parallel models argue that attention is spatially distributed in parallel, allowing simultaneous processing of all words within the visual field (Engbert, Nuthmann, Richter, & Kliegl, 2005). The interpretation and predictions of these two eye movement control models regarding preview processing have become a central topic of debate in visual cognitive processing research.

1.1 Preview Processing of Between-Words and Within-Words

Researchers typically employ the boundary paradigm (Rayner, 1975) to investigate preview processing. In this paradigm, an invisible boundary exists between the fixated word and the preview word. Before the eyes cross this boundary, a preview word appears in the target location; once the eyes cross the boundary, the preview word is immediately replaced by the target word. Previous preview processing research has positioned the invisible boundary between words, meaning the preview word and the currently fixated word belong to different words (between-words). Studies have found that preview benefits for English between-words range from approximately 30 to 50 ms (Rayner, 2009). However, recent research suggests that the position of the invisible boundary influences the magnitude of preview effects. Hyönä, Bertram, and Pollatsek (2004) were the first to position the invisible boundary within a word, such that the preview word and the fixated word belonged to the same compound word (within-words). Their results revealed an 80 ms preview benefit when the boundary was located within a compound word—substantially larger than typical findings. The researchers proposed two possible explanations for this enhanced preview effect: First, within-word lexemes lack spaces between them, placing the preview word closer to the fixation point than in between-word conditions and thereby producing larger preview benefits. Second, for within-words, the two morphemes on either side of the invisible boundary belong to the same linguistic unit (a compound word), whereas for between-words, the two words belong to separate linguistic units, potentially leading to larger preview effects for within-words.

To test Hyönä et al.'s (2004) first hypothesis regarding the influence of spaces on preview processing, Juhasz, Pollatsek, Hyönä, Drieghe, and Rayner (2009) selected compound words both with and without spaces (e.g., "basketball" vs. "tennis ball"), where both types had lexemes belonging to the same linguistic unit. Their results showed that preview benefits for unspaced compounds were significantly larger than for spaced compounds, thereby confirming the first hypothesis and demonstrating that spaces increase the distance between the reader' s fixation point and the preview word, resulting in smaller preview benefits for between-words. To test Hyönä et al.' s (2004) second hypothesis, Juhasz et al. (2009) controlled for identical distances between fixated and preview words while manipulating whether the words on either side of the invisible boundary belonged to the same linguistic unit. They used spaced compounds and phrases as target words (e.g., "tennis ball" vs. "filtby ball"). Results showed that although preview benefits for compounds (31 ms) differed from those for phrases (20 ms), this difference did not reach statistical significance, failing to support Hyönä et al.' s second hypothesis. However, the authors cautioned against dismissing the potential influence of linguistic units on preview processing, noting that although phrase constituents belong to different linguistic units, the high syntactic expectancy of nouns following adjectives may lead readers to mistakenly treat adjective-noun phrases as a single linguistic unit, thereby eliminating differences between between-words (phrases) and within-words (compounds).

Consequently, further research is needed to clarify the differences in preview processing between between-words and within-words.

Because English words are separated by spaces, researchers investigating preview processing in between-words and within-words could only examine both word types when spaces were present, making it impossible to explore their preview processing while completely eliminating the influence of spaces. Chinese characters, as logographic units, are composed of strokes arranged in specific configurations, with each character occupying identical width. A single character can constitute a word independently or combine with other characters to form multi-character words. In Chinese text, characters are the basic units, and no explicit boundary information exists between words. This characteristic allows researchers to investigate the influence of linguistic units on preview processing while eliminating the confounding effect of spaces. Cui et al. (2013) were the first to adopt the boundary paradigm to examine how linguistic units affect preview processing in Chinese reading. Their study compared preview benefits for simple words, compound words, and phrases in Chinese text reading. Simple words consist of two characters but contain only one morpheme, with the two characters inseparable (e.g., “玫瑰” *rose*). Compound words comprise two characters that can also function as independent words (e.g., “灯塔” *lighthouse*). Phrases also consist of two characters, both of which are independent words that cannot combine to form a single word (e.g., “斜塔”*leaning tower*). Results showed significant parafoveal-on-foveal effects only during preview processing of simple words, where preview information influenced current reading; however, neither compound words nor phrases exhibited significant parafoveal-on-foveal effects, and no significant difference in preview benefits emerged between compound words and phrases. These findings indicated that although the two characters comprising simple and compound words belong to the same linguistic unit, the first character of simple words has higher expectancy for the second character than in compound words, leading to greater preview processing. For compound words and phrases, despite compound morphemes belonging to the same linguistic unit while phrase constituents do not, preview processing did not differ between them. This result aligns with Juhasz et al. (2009). Cui et al. (2013) noted that due to the uncertainty of word segmentation in Chinese (Hoosain, 1992), many readers cannot clearly distinguish compounds from phrases; in post-experiment segmentation assessments, participants misidentified 45% of phrases as compounds, likely accounting for the lack of significant differences in preview processing between compounds and phrases. Therefore, the present study provides clearer definitions of word boundaries for between-words and within-words to investigate the influence of linguistic units on preview processing.

As noted above, although English research can examine the influence of between-words and within-words on preview processing when spaces are present, it cannot completely eliminate the effect of spaces. In Cui et al.’ s (2013) study, participants did not adequately distinguish between compound and phrase boundaries. To more clearly investigate how linguistic units affect preview processing, the present study provides more explicit definitions of word boundaries for between-words and within-words by manipulating whether the invisible boundary appears before the first morpheme of a compound word or between its first and second morphemes. This manipulation strictly controls whether the target character and the character preceding the boundary belong to the same linguistic unit. Additionally, by manipulating preview type (consistent preview vs. pseudo-character preview), we examine whether preview effects differ between between-words and within-words. We hypothesize that when the influence of spaces is eliminated, within-words will produce larger preview benefits than between-words.

1.2 The Influence of Boundary Information on Preview Processing of Between-Words and Within-Words

Recent research has focused on preview processing in Chinese reading and differences in preview processing between Chinese and alphabetic writing systems. Similar to findings from Indo-European languages, researchers have found that preview processing significantly influences both when and where readers fixate during Chinese text reading (Morris, Rayner, & Pollatsek, 1990; Perea & Acha, 2009; Yen, Radach, Tzeng, Hung, & Tsai, 2009). When the preview word is high-frequency, fixation durations are shorter and skipping rates higher (Liu, Reichle, & Li, 2015; Rayner, Ashby, Pollatsek, & Reichle, 2004), and saccade lengths into and out of high-frequency words are longer than for low-frequency words (Hyönä, 1995; Liu, Reichle, & Li, 2016; Wei, Li, & Pollatsek, 2013). However, researchers have also identified substantial differences in the magnitude and depth of preview processing between Chinese and Indo-European languages. First, preview benefits are larger in Chinese. While English preview effects are approximately 30–50 ms (Rayner, 2009), Chinese preview effects can exceed 60 ms (Cui et al., 2013; Yan, Richter, Shu, & Kliegl, 2009; Yang, Wang, Xu, & Rayner, 2009). A meta-analysis by Vasilev and Angele (2017) confirmed that Chinese preview effects are approximately 10 ms larger than those in alphabetic scripts when measured by gaze duration. Second, Chinese preview processing accesses higher-level information. Most English studies have demonstrated orthographic and phonological preview effects but failed to find semantic-level preview effects (Dimigen, Kliegl, & Sommer, 2012; Rayner, Balota, & Pollatsek, 1986), with semantic preview benefits emerging only under special circumstances (e.g., initial capitalization or high predictability) (Rayner, Schotter, & Drieghe, 2014; Schotter, Lee, Reiderman, & Rayner, 2015). In contrast, numerous studies have identified stable semantic preview effects in Chinese reading. Wang, Tong, Yang, and Leng (2009) manipulated preview characters to create semantically coherent and incoherent sentence conditions, finding significantly longer fixation times on target words in semantically incoherent conditions, demonstrating semantic preview benefits. Subsequent research has consistently supported this finding (Li, Wang, Mo, & Kliegl, 2017; Yan et al., 2009; Yang et al., 2009).

The larger and stronger preview effects in Chinese may stem from unique characteristics of the Chinese language. First, Chinese has more compact visual density (e.g., 住宿 vs. “accommodation”), allowing readers to process more words within the visual span while fixating on the current word. Second, each Chinese character is an independent visual unit containing orthographic, phonological, and likely semantic information. Moreover, the relationship between orthography and semantics in Chinese characters is tighter than in English, enabling earlier access to semantic information during preview processing (Schotter, 2013; Schotter et al., 2015). Third, Yen, Radach, Tzeng, Hung, and Tsai (2009) proposed that larger preview effects in Chinese result from the absence of explicit word boundary information. In Chinese text, words lack clear boundaries, and a character can function as an independent word or combine with other characters to form multi-character words. Consequently, while processing the currently fixated character (or word), readers may need to conduct more extensive preview processing of subsequent characters (or words) to determine word boundaries in advance. According to Yen et al., if the absence of boundary information causes larger preview effects in Chinese, would reducing word boundary information in English increase preview effects? Conversely, would adding word boundary information to Chinese text decrease preview effects? Researchers have yet to reach consistent conclusions.

Some studies indicate that readers obtain larger preview effects when word spaces are removed from English text. Drieghe, Fitzsimmons, and Liversedge (2017) manipulated pre-target word frequency (high vs. low) and sentence type (spaced vs. unspaced) in English text to compare preview effects between spaced and unspaced conditions. Results showed that when the pre-target word was high-frequency, preview effects in unspaced text were significantly larger than in high-frequency spaced, low-frequency spaced, and low-frequency unspaced conditions. The authors argued that when no word boundary information is available and pre-target word processing is easy (high-frequency), readers are more likely to process multiple words simultaneously to determine word boundaries, resulting in increased preview effects. In contrast, when word boundary information is present, readers focus more intensively on currently fixated information, leading to reduced preview effects.

However, other researchers contend that spaces in English text facilitate preview processing. Drieghe, Brysbaert, and Desmet (2005) found that adding a space between word N and word N+1 reduced reading times for word N. Drieghe et al. (2005, 2017) argued that adding boundary information between words reduces inter-word boundary masking (i.e., the lack of explicit boundaries between words), leading to faster word recognition. Additionally, Sheridan, Reichle, and Reingold (2016) compared preview effects between normal English text with inter-word spaces and spaceless English text where boundaries were filled with numbers. Their results showed that preview effects for spaced text were significantly larger than for spaceless text when measured by first fixation duration. They proposed that spaces facilitate word segmentation and help readers target saccades effectively while reducing inter-word boundary masking, thereby promoting preview processing.

Cui, Denis, Bai, Yan, and Liversedge (2014) examined the effect of adding inter-word spaces on preview processing in Chinese text. Their results showed that for single-character words, preview effects were larger in inter-word spaced conditions than in normal text conditions, whereas for two-character words, no significant differences emerged between spaced and normal conditions. The authors argued that adding inter-word spaces provides boundary information for single-character words, and after inserting spaces, single-character words are flanked by spaces on both sides, eliminating boundary masking during processing. For two-character words, although spaces also provide boundary information, boundary masking persists between the two morphemes, with each morpheme adjacent to a space on only one side. Therefore, they concluded that the increased preview effect for single-character words in inter-word spaced conditions results from reduced boundary masking rather than from boundary information itself.

The aforementioned studies demonstrate that using spaces as word boundary information affects boundary masking differently for different word types. Therefore, the present study employs inter-word shading as word boundary information to investigate its influence on preview processing. Using shading allows examination of word boundary effects on preview benefits while controlling sentence length (Bai, Yan, Liversedge, Zang, & Rayner, 2008) and ensures that boundary masking is affected equally across different word types. Furthermore, Rayner and Schotter (2014) found that preview effects were significantly larger when preview words were presented in uppercase letters than in lowercase, suggesting that this salient feature increased reader attention and thereby enhanced preview effects. Consequently, the present study uses yellow shading to mark boundary information, maximizing the potential for boundary information to influence preview processing.

In summary, this study manipulates whether the invisible boundary is positioned before the first morpheme of a compound word or between its first and second morphemes, thereby controlling whether the target character and the character preceding the boundary belong to the same linguistic unit, to examine whether preview effects differ between between-words and within-words. We hypothesize that within-words will produce larger preview effects than between-words when the influence of spaces is eliminated. Additionally, we use normal, inter-word shading, and non-word shading text conditions to investigate how word boundary information affects the preview processing of between-words and within-words. We propose that word boundary information will have a greater influence on between-words than within-words because, when processing the first morpheme of between-words, inter-word shading provides boundary information that facilitates word segmentation, thereby producing larger preview effects in the inter-word shading condition than in normal or non-word shading conditions. In contrast, when processing within-words (i.e., the second morpheme of a two-character word), readers have already completed segmentation of the current word, so word boundary information should have minimal impact on preview effects for within-words.

2.1 Participants

Participants were 104 undergraduate students from Tianjin Normal University. All had normal or corrected-to-normal vision, were native Chinese speakers without reading disabilities, and were unfamiliar with the experimental purpose. Participants received a small gift as compensation upon completion.

2.2 Apparatus

We used an SR Research Eyelink 2000 eye tracker (SR Research, Canada) with a sampling rate of 1000 Hz to record participants' right eye movements. Experimental materials were presented in Song font, black characters on a white background. Each character measured 2.1 cm × 2.1 cm on the screen. Participants viewed the screen from a distance of 60 cm, resulting in each character subtending approximately 1° of visual angle, ensuring that when participants fixated on the character preceding the target, the target character fell within the parafoveal processing region.

2.3 Design and Materials

The experiment employed a three-factor within-subjects design. Factor 1 was word type with two levels: between-words and within-words. Factor 2 was sentence type with three levels: normal presentation, inter-word shading, and non-word shading. Factor 3 was preview type with two levels: consistent and inconsistent.

We selected 120 pairs of between-words and within-words. As described above, between-words positioned the invisible boundary before the first morpheme of the target word, meaning the fixated word and preview word belonged to different words. Within-words positioned the invisible boundary between the two morphemes of the target word, meaning the fixated word and preview word belonged to the same compound word. Therefore, we controlled for the first character after the invisible boundary (the preview character) being identical across each between-word pair (e.g., "纸盒" *paper box*) and within-word pair (e.g., "稿纸" *draft paper*). Whole-word frequency (based on the People' s Daily word frequency corpus, 1998) and stroke count did not differ significantly between word types ($ts < 1.39$). To avoid potential effects of positional frequency of target characters as either the first or second morpheme of two-character words, we also controlled for positional frequency differences ($t(119) = 0.84$, $p > 0.05$). Detailed lexical properties are presented in Table 1 .

In the inconsistent preview condition, the preview character was a pseudo-character constructed using TrueType font software by combining Chinese character components. These pseudo-characters shared the same structural features as real characters but did not exist in Chinese and possessed no phonological or semantic information.

Experimental sentences shared identical structures for within-word and between-word pairs, meaning that except for the target word, all other sentence content was identical. All sentences were semantically plausible, and target words could not combine with preceding or following characters to form new words. Sentences contained a maximum of 17 characters and were presented on a single line on the computer screen. Target characters never appeared in the first three or last position of sentences to avoid special position effects such as sentence-final integration effects. We used the boundary paradigm (Rayner, 1975): before the eyes crossed the invisible boundary, the preview position displayed a character that was either identical to or inconsistent with the target character; once the eyes crossed the boundary, the preview character was replaced by the target character. Examples of experimental materials and the paradigm are shown in Figure 1 [Figure 1: see original paper].

In Figure 1, target words are bolded for illustration: within-words (“稿纸”) and between-words (“纸盒”). Note that target words were not bolded during the actual experiment. Invisible boundary positions are marked with vertical lines (|). For within-words, the invisible boundary fell between the two morphemes of the two-character word; for between-words, it fell to the left of the first character. Before crossing the invisible boundary, the preview position displayed either a character identical to the target character (“纸”) or a pseudo-character inconsistent with the target character (” “). When participants’ eyes crossed the invisible boundary, the preview character was replaced by the target character (” 纸”).

2.4 Material Evaluation

Sentence Plausibility Rating. To ensure semantic coherence and equivalent plausibility between sentences containing within-words and between-words, we conducted a sentence plausibility rating. Twenty-four university students who did not participate in the main experiment rated sentences using a five-point scale. For half the participants, 1 indicated “very implausible” and 5 indicated “very plausible,” while this scale was reversed for the other half. In addition to 120 experimental sentences, we included 36 semantically implausible sentences. Results showed no significant difference in plausibility between within-word sentences ($M = 1.84$, $SD = 0.22$) and between-word sentences ($M = 1.79$, $SD = 0.31$; $t < 1$), meeting our research requirements.

Sentence Segmentation Rating. Two groups of 12 participants each (total $N = 24$) performed word boundary segmentation on sentences containing between-words and within-words, respectively, including all 120 experimental sentences. Participants were instructed to read each sentence carefully and mark word boundaries with vertical lines (Yan et al., 2010). Results showed segmentation consistency rates of 85.9% for between-words and 87.2% for within-words, with no significant difference between them, satisfying our experimental requirements.

2.5 Procedure

Upon entering the laboratory, the experimenter explained the instructions to participants and answered any questions to ensure correct understanding of the procedure. The eye tracker was then adjusted and a three-point calibration was performed. Before each sentence or question, a drift correction was conducted to ensure recording accuracy: a calibration point appeared on the left side of the screen center and disappeared only when the participant' s fixation aligned with it, after which the next screen was presented. If fixation drift occurred during the experiment, recalibration was performed immediately.

After calibration, the experiment began with a practice block consisting of six sentences and two questions to familiarize participants with the procedure, followed by the formal experiment. Reading materials were presented sentence-by-sentence at the center of the screen. Participants were instructed to read each sentence at their own pace and press a key when finished, proceeding to the next sentence. After every three sentences, participants answered a comprehension question based on the content they had read. A Latin square design ensured that each participant read only one condition of each target sentence. Following the experiment, participants were asked whether they noticed any unusual phenomena. Some participants reported occasional character changes but could not clearly report the specific nature of these changes.

Each participant read 156 sentences total: 120 experimental sentences presented in random order, 30 filler sentences, and 6 practice sentences. The entire experiment lasted approximately 35 minutes, including 5 minutes for calibration.

3. Results and Analysis

We excluded data from four participants with comprehension question accuracy below 75% (Rayner, 1998) and four participants who detected boundary changes in more than 25% of trials. Effective eye movement data were obtained from 96 participants. Fixation durations shorter than 60 ms or longer than 600 ms were excluded from analysis, resulting in the removal of approximately 11.2% of data.

Based on our research objectives, we report the following eye movement measures that primarily reflect early cognitive processing: first fixation duration (the duration of the first fixation on an interest area), gaze duration (the total duration from first entering an interest area until first exiting it), regression path duration (the sum of all fixation durations from first entering an interest area until fixating to its right), and skipping rate (the probability of skipping an interest area during first-pass reading) (Yan et al., 2013). F values represent analyses with participants as the random factor, while F values represent analyses with items as the random factor.

Figure 2 [Figure 2: see original paper] illustrates the interest area units used for data analysis. Because word type was manipulated by positioning the target character in between-words versus within-words, data analysis focused primarily on the target character. We report eye movement measures for the target region, with interest area units shown in Figure 2 (vertical lines indicate boundary positions). Specific eye movement data are presented in Table 2 .

Table 2 Means and standard deviations of eye movement measures on target characters

Note: Units are milliseconds; I = consistent preview; D = inconsistent preview; PB = preview benefit (preview benefit = inconsistent preview condition - consistent preview condition)

First Fixation Duration

For first fixation duration, the main effect of word type was significant, F (1, 94) = 39.77, $p < 0.001$, $^{2} = 0.30$; F (1, 113) = 28.56, $p < 0.001$, $^{2} = 0.20$, with longer first fixation durations for within-words than between-words. The main effect of preview type was significant, F (1, 94) = 102.83, $p < 0.001$, $^{2} = 0.52$; F (1, 113) = 243.26, $p < 0.001$, $^{2} = 0.68$, with significantly longer first fixation durations for inconsistent than consistent preview conditions. The main effect of sentence type was not significant, F (2, 118) = 1.59, $p > 0.1$; F (2, 226) = 2.25, $p > 0.1$, with no significant differences among normal, inter-word shading, and non-word shading conditions. The interaction between word type and sentence type was significant, F (2, 188) = 5.75, $p < 0.01$, $^{2} = 0.06$; F (2, 226) = 4.27, $p < 0.05$, $^{2} = 0.04$. Simple effects tests revealed no significant difference between within-words (245 ms) and between-words (241 ms) in the non-word shading condition, but within-words produced significantly longer first fixation durations than between-words in both normal presentation and inter-word shading conditions. The interaction between word type and preview type was significant, F (1, 94) = 11.63, $p < 0.01$, $^{2} = 0.11$; F (1, 113) = 8.23, $p < 0.01$, $^{2} = 0.07$, with significantly larger preview benefits for within-words (42 ms) than between-words (24 ms). No other interactions were significant (Fs < 2.25).

Gaze Duration

For gaze duration, the main effect of word type was significant, F (1, 94) = 49.68, $p < 0.001$, $^{2} = 0.35$; F (1, 113) = 41.94, $p < 0.001$, $^{2} = 0.27$, with longer gaze durations for within-words than between-words. The main effect of preview type was significant, F (1, 94) = 144.06, $p < 0.001$, $^{2} = 0.61$; F (1, 113) = 405.04, $p < 0.001$, $^{2} = 0.78$, with significantly longer gaze durations for inconsistent than consistent preview conditions. The main effect of sentence type was not significant, F (2, 188) = 0.48, $p > 0.1$; F (2, 226) = 0.76, $p > 0.1$, with no significant differences among normal presentation, inter-word shading, and non-word shading conditions. The interaction between word type and sentence type was significant, F (2, 188) = 5.42, $p < 0.01$, $^{2} = 0.05$; F (2, 226) = 4.74, $p < 0.05$, $^{2} = 0.04$. Simple effects tests revealed no significant difference between within-words and between-words in the non-word shading condition, but within-words produced significantly longer gaze durations than between-words in both normal presentation and inter-word shading conditions. The interaction between word type and preview type was significant, F (1, 94) = 15.06, $p < 0.001$, $^2 = 0.14$; F (1, 113) = 11.79, $p < 0.01$, $^2 = 0.09$, with significantly larger preview benefits for within-words (60 ms) than between-words (36 ms). No other interactions were significant (Fs < 2.10).

Regression Path Duration

For regression path duration, the main effect of word type was significant, F (1, 92) = 60.35, $p < 0.001$, $^2 = 0.40$; F (1, 113) = 42.23, $p < 0.001$, $^2 = 0.27$, with longer regression path durations for within-words than between-words. The main effect of preview type was significant, F (1, 92) = 113.84, $p < 0.001$, $^2 = 0.55$; F (1, 113) = 275.34, $p < 0.001$, $^2 = 0.71$, with significantly longer regression path durations for inconsistent than consistent preview conditions. The main effect of sentence type was significant in the item analysis, F (2, 184) = 2.06, $p > 0.1$; F (2, 226) = 7.02, $p < 0.01$, $^2 = 0.06$, with significantly shorter regression path durations for inter-word shading (318 ms) than for normal presentation (334 ms) and non-word shading (338 ms) conditions, though normal and non-word shading did not differ significantly. The interaction between word type and preview type was significant, F (1, 92) = 34.99, $p < 0.001$, $^2 = 0.28$; F (1, 113) = 27.61, $p < 0.001$, $^2 = 0.20$, with significantly larger preview benefits for within-words (60 ms) than between-words (36 ms). The interaction between sentence type and preview type was significant, F (2, 184) = 4.62, $p < 0.05$, $^2 = 0.05$; F (2, 226) = 3.62, $p < 0.05$, $^2 = 0.03$, with smaller preview benefits for inter-word shading (57 ms) than for normal presentation (86 ms) and non-word shading (95 ms) conditions, and smaller preview benefits for normal than non-word shading. No other interactions were significant (Fs < 2.10).

Skipping Rate

For skipping rate, the main effect of word type was significant, F (1, 95) = 28.36, $p < 0.001$, $^2 = 0.23$; F (1, 119) = 13.99, $p < 0.001$, $^2 = 0.11$, with lower skipping rates for within-words than between-words. The main effect of preview type was not significant, F (1, 94) = 0.84, $p > 0.1$; F (1, 113) = 0.99, $p > 0.1$, with no significant difference in skipping rates between consistent and inconsistent preview conditions. The main effect of sentence type was significant, F (2, 190) = 3.96, $p < 0.05$, $^2 = 0.04$; F (2, 238) = 6.00, $p < 0.01$, $^2 = 0.05$, with significantly higher skipping rates for inter-word shading and non-word shading than for normal presentation, though inter-word shading and non-word shading did not differ significantly. The three-way interaction among word type, sentence type, and preview type was significant in the item analysis, F (2, 190) = 1.90, $p > 0.1$; F (2, 238) = 13.22, $p < 0.05$, $^2 = 0.03$. No other interactions were significant (Fs < 2.10).

Because the three-way interaction was significant in the item analysis, we conducted simple effects analyses. For within-words, the interaction between preview type and sentence type was significant, F (2, 190) = 3.55, $p < 0.05$, ^{2} = 0.04; F (2, 238) = 4.76, $p < 0.01$, ^{2} = 0.04, with no significant difference between consistent and inconsistent preview in normal presentation and non-word shading conditions, but significantly higher skipping rates for consistent than inconsistent preview in the inter-word shading condition. Main effects of sentence type and preview type were not significant (Fs < 1.8). For between-words, the main effect of sentence type was significant in the item analysis, F (2, 190) = 0.89, $p > 0.1$; F (2, 238) = 4.66, $p = 0.01$, ^{2} = 0.04, with significantly lower skipping rates for normal presentation and non-word shading than for inter-word shading, though normal and non-word shading did not differ significantly. Main effects of preview type and the interaction between preview type and sentence type were not significant (Fs < 1.7).

For normal presentation, the main effect of word type was significant in the participant analysis, F (1, 95) = 4.21, $p < 0.05$, ^{2} = 0.04; F (1, 119) = 2.48, $p > 0.1$, with higher skipping rates for within-words than between-words. Main effects of preview type and the interaction between word type and preview type were not significant (Fs < 1.4). For inter-word shading, the main effect of word type was significant, F (1, 95) = 10.61, $p < 0.01$, ^{2} = 0.10; F (1, 119) = 8.22, $p < 0.01$, ^{2} = 0.07, with higher skipping rates for within-words than between-words. The main effect of preview type was marginally significant in the participant analysis, F (1, 95) = 3.54, $p < 0.1$, ^{2} = 0.04; F (1, 119) = 2.53, $p > 0.1$, with higher skipping rates for consistent than inconsistent preview. The interaction between word type and preview type was marginally significant in the participant analysis and significant in the item analysis, F (1, 95) = 3.70, $p < 0.1$, ^{2} = 0.04; F (1, 119) = 4.95, $p < 0.05$, ^{2} = 0.04, with higher skipping rates for consistent than inconsistent preview for within-words but no difference for between-words. For non-word shading, the main effect of word type was significant in the participant analysis, F (1, 95) = 11.13, $p < 0.01$, ^{2} = 0.11; F (1, 119) = 5.36, $p < 0.05$, ^{2} = 0.04, with higher skipping rates for within-words than between-words. Main effects of preview type and the interaction between word type and preview type were not significant (Fs < 1.0).

Additionally, because the invisible boundary was positioned between the two morphemes of within-words and before the two morphemes of between-words, we compared the second character of within-words with the whole two-character between-word as the interest area for preview processing. The interest area units for this analysis are shown in Figure 3 [Figure 3: see original paper] (vertical lines indicate boundary positions), with specific eye movement data presented in Table 3 .

Table 3 Means and standard deviations of eye movement measures on target regions

Note: Units are milliseconds; I = consistent preview; D = inconsistent preview; PB = preview benefit (preview benefit = inconsistent preview condition -

consistent preview condition)

First Fixation Duration For first fixation duration, the main effect of word type was significant, F (1, 94) = 43.29, $p < 0.001$, $^2 = 0.32$; F (1, 117) = 28.05, $p < 0.001$, $^2 = 0.19$, with longer first fixation durations for within-words than between-words. The main effect of preview type was significant, F (1, 94) = 99.27, $p < 0.001$, $^2 = 0.51$; F (1, 117) = 193.64, $p < 0.001$, $^2 = 0.62$, with significantly longer first fixation durations for inconsistent than consistent preview conditions. The main effect of sentence type was not significant, F (2, 188) = 0.76, $p > 0.1$; F (2, 234) = 1.34, $p > 0.1$, with no significant differences among normal presentation, inter-word shading, and non-word shading conditions. The interaction between word type and sentence type was marginally significant in the item analysis, F (2, 188) = 2.15, $p > 0.1$; F (2, 234) = 2.67, $p = 0.07$, $^2 = 0.05$. Simple effects tests revealed no significant difference between within-words and between-words in the non-word shading condition (difference = 7 ms), but within-words produced significantly longer first fixation durations than between-words in both normal presentation (difference = 14 ms) and inter-word shading (difference = 18 ms) conditions. The interaction between word type and preview type was significant, F (1, 94) = 28.82, $p < 0.001$, $^2 = 0.24$; F (1, 117) = 32.65, $p < 0.001$, $^2 = 0.22$, with significantly larger preview benefits for within-words (42 ms) than between-words (18 ms). No other interactions were significant (Fs < 1.54).

Gaze Duration For gaze duration, the main effect of word type was significant, F (1, 94) = 274.12, $p < 0.001$, $^2 = 0.75$; F (1, 117) = 420.70, $p < 0.001$, $^2 = 0.78$, with shorter gaze durations for within-words than between-words. The main effect of preview type was significant, F (1, 94) = 136.70, $p < 0.001$, $^2 = 0.59$; F (1, 117) = 262.47, $p < 0.001$, $^2 = 0.69$, with significantly longer gaze durations for inconsistent than consistent preview conditions. The main effect of sentence type was significant, F (2, 188) = 10.16, $p < 0.001$, $^2 = 0.10$; F (2, 234) = 15.76, $p < 0.001$, $^2 = 0.12$, with significantly shorter gaze durations for normal presentation (313 ms) and inter-word shading (312 ms) than for non-word shading (329 ms), though normal presentation and inter-word shading did not differ significantly. The interaction between word type and sentence type was significant, F (2, 188) = 22.16, $p < 0.001$, $^2 = 0.19$; F (2, 234) = 23.61, $p < 0.001$, $^2 = 0.17$. Simple effects tests revealed that within-words produced significantly shorter gaze durations than between-words across all sentence types, but the difference was largest in the non-word shading condition (difference = 120 ms), followed by normal presentation (difference = 80 ms), and smallest in the inter-word shading condition (difference = 70 ms). No other interactions were significant (Fs < 1.74).

Regression Path Duration For regression path duration, the main effect of word type was significant, F (1, 94) = 213.48, $p < 0.001$, $^2 = 0.69$; F (1, 117) = 201.41, $p < 0.001$, $^2 = 0.63$, with shorter regression path durations for within-words than between-words. The main effect of preview type was significant, F (1, 94) = 152.50, $p < 0.001$, $^2 = 0.62$; F (1, 117) = 311.66, $p < 0.001$, $^2 = 0.73$, with significantly longer regression path durations for inconsistent than consistent preview conditions. The main effect of sentence type was significant in both analyses, F (1, 94) = 10.97, $p < 0.001$, $^2 = 0.11$; F (1, 117) = 14.15, $p < 0.001$, $^2 = 0.11$, with significantly shorter regression path durations for normal presentation (398 ms) and inter-word shading (387 ms) than for non-word shading (419 ms), though normal presentation and inter-word shading did not differ significantly. The interaction between word type and sentence type was significant in the participant analysis and marginally significant in the item analysis, F (2, 188) = 3.54, $p < 0.05$, $^2 = 0.04$; F (2, 234) = 2.91, $p = 0.06$, $^2 = 0.02$. Simple effects tests revealed that within-words produced significantly shorter regression path durations than between-words across all sentence types, but the difference was largest in the non-word shading condition (difference = 126 ms), followed by normal presentation (difference = 98 ms), and smallest in the inter-word shading condition (difference = 91 ms). The interaction between word type and preview type was significant in the participant analysis and marginally significant in the item analysis, F (1, 94) = 5.15, $p < 0.05$, $^2 = 0.05$; F (1, 117) = 3.80, $p = 0.05$, $^2 = 0.03$, with significantly larger preview benefits for within-words (108 ms) than between-words (85 ms). The interaction between sentence type and preview type was marginally significant in the participant analysis and significant in the item analysis, F (2, 188) = 2.86, $p = 0.06$, $^2 = 0.03$; F (2, 234) = 3.07, $p < 0.05$, $^2 = 0.03$, with significantly smaller preview benefits for inter-word shading (75 ms) than for normal presentation (110 ms) and non-word shading (105 ms). No other interactions were significant (Fs < 2.27).

Skipping Rate For skipping rate, the main effect of word type was significant, F (1, 95) = 187.45, $p < 0.001$, $^2 = 0.66$; F (1, 113) = 13.99, $p < 0.001$, $^2 = 0.11$, with higher skipping rates for within-words than between-words. The main effect of preview type was not significant, F (1, 95) = 0.84, $p > 0.1$; F (1, 119) = 0.99, $p > 0.1$, with no significant difference in skipping rates between consistent and inconsistent preview conditions. The main effect of sentence type was significant in the item analysis but not in the participant analysis, F (2, 190) = 1.82, $p > 0.1$; F (2, 238) = 6.00, $p < 0.01$, $^2 = 0.05$, with higher skipping rates for inter-word shading and non-word shading than for normal presentation, though inter-word shading and non-word shading did not differ significantly. The three-way interaction among word type, sentence type, and preview type was significant in the item analysis, F (2, 190) = 2.07, $p < 0.01$, $^2 = 0.04$; F (2, 238) = 3.20, $p < 0.05$, $^2 = 0.03$. No other interactions were significant (Fs < 2.08).

Because the three-way interaction was significant in the item analysis, we conducted simple effects analyses. For within-words, the interaction between preview type and sentence type was significant, F (2, 190) = 3.50, $p < 0.05$, $^2 = 0.04$; F (2, 238) = 4.76, $p < 0.01$, $^2 = 0.04$, with no significant difference between consistent and inconsistent preview in normal presentation and non-word shading conditions, but significantly higher skipping rates for consistent than inconsistent preview in the inter-word shading condition. Main effects of sentence type and preview type were not significant (Fs < 1.7). For between-words, the main effect of sentence type was significant in the item analysis, F (2, 190) = 0.89, $p > 0.1$; F (2, 238) = 4.66, $p = 0.01$, $^2 = 0.04$, with lower skipping rates for normal presentation and non-word shading than for inter-word shading, though normal and non-word shading did not differ significantly. Main effects of preview type and the interaction between preview type and sentence type were not significant (Fs < 1.7).

For normal presentation, the main effect of word type was significant in the participant analysis, F (1, 95) = 83.54, $p < 0.001$, $^2 = 0.47$; F (1, 119) = 2.48, $p > 0.1$, with higher skipping rates for within-words than between-words. Main effects of preview type and the interaction between word type and preview type were not significant (Fs < 1.4). For inter-word shading, the main effect of word type was significant, F (1, 95) = 59.35, $p < 0.001$, $^2 = 0.39$; F (1, 119) = 8.22, $p < 0.01$, $^2 = 0.07$, with higher skipping rates for within-words than between-words. The main effect of preview type was significant in the participant analysis, F (1, 95) = 8.39, $p < 0.01$, $^2 = 0.08$; F (1, 119) = 2.53, $p > 0.1$, with higher skipping rates for consistent than inconsistent preview. The interaction between word type and preview type was significant, F (1, 95) = 4.92, $p < 0.05$, $^2 = 0.05$; F (1, 119) = 4.95, $p < 0.05$, $^2 = 0.04$, with higher skipping rates for consistent than inconsistent preview for within-words but no difference for between-words. For non-word shading, the main effect of word type was significant in the participant analysis, F (1, 95) = 98.48, $p < 0.001$, $^2 = 0.51$; F (1, 119) = 5.36, $p < 0.05$, $^2 = 0.04$, with higher skipping rates for within-words than between-words. Main effects of preview type and the interaction between word type and preview type were not significant (Fs < 1.7).

General Discussion

This study used eye-tracking methodology to investigate differences in preview processing between between-words and within-words by manipulating whether the invisible boundary was positioned before the first morpheme of a compound word or between its first and second morphemes, thereby strictly controlling whether the target character and the character preceding the boundary belonged to the same linguistic unit. Additionally, to examine the influence of word boundary information on preview processing in Chinese text, we introduced inter-word shading as a boundary marker to explore the mechanism and impact of boundary information on preview effects, providing possible explanations for differences in preview effects between Chinese and alphabetic writing systems and offering insights into word segmentation and dynamic processing during Chinese reading, thereby furnishing new perspectives and empirical foundations for eye movement control models.

Results showed that when the target character served as the interest area, no significant differences emerged among normal presentation, inter-word shading, and non-word shading conditions for first fixation duration, gaze duration, and skipping rate, consistent with Bai et al. (2008). This may indicate that word boundary information marked by shading does not affect reading processes. Alternatively, as Bai et al. suggested, readers are familiar with unmarked text from daily reading, so the null effect between normal and boundary-marked text may actually reflect a facilitative effect of the unfamiliar shaded condition, as readers showed no reading impairment despite the unusual format. Moreover, we found significantly shorter reading times for inter-word shading than for normal text in regression path duration, suggesting that facilitative effects of inter-word shading may indeed exist, though further research is needed for confirmation.

When analyzing data with the second character of within-words and the whole two-character between-word as the interest area, we found significantly shorter reading times for normal text and inter-word shading than for non-word shading in gaze duration and regression path duration, consistent with Bai et al. (2008) and demonstrating the role of words in Chinese reading processing. Furthermore, because word boundary information marked by shading did not produce facilitative effects in early measures (first fixation duration, gaze duration) but only in later measures (regression path duration), these results suggest that word boundary information may facilitate reading during later rather than early stages (Bai et al., 2008). Understanding when word segmentation occurs is important for comprehending Chinese reading processes, though the time course of its effects requires further investigation.

Second, we found larger preview benefits for within-words than between-words in first fixation duration, gaze duration, and regression path duration, indicating that preview effects are larger when the preview character and currently fixated character belong to the same linguistic unit than when they belong to different units. This confirms Hyönä et al.' s (2004) hypothesis that lexical properties affect preview processing, with larger preview benefits when morphemes on either side of a boundary belong to the same linguistic unit. These results differ from previous studies examining preview effects for within-words and between-words (Cui et al., 2013; Juhasz et al., 2009). However, as the authors noted, word expectancy and segmentation uncertainty are important factors affecting results. By strictly controlling whether the target character and the character preceding the boundary belonged to the same linguistic unit, the present study demonstrates that linguistic unit status is indeed an important factor contributing to larger preview effects for within-words than between-words.

These results also indicate that Chinese readers can flexibly segment words based on target character position. Zang et al. (2016) investigated whether the positional frequency of a pre-target character as either the first morpheme of a two-character word or as a single-character word affects preview processing. They found larger preview effects when the pre-target character more frequently appeared as the first morpheme than as a single character, demonstrating that readers can adjust preview processing based on positional frequency. The present study manipulated whether the invisible boundary fell between the first and second morphemes or before the first morpheme of a compound word, positioning the target character within or between words. Results showed larger preview effects when the target character was the second morpheme of a two-character word, indicating that Chinese readers can rapidly assess the linguistic properties of characters based on their position and adjust preview processing accordingly.

Third, we found no difference in preview effects between inter-word shading and normal presentation conditions in first fixation duration, gaze duration, and skipping rate. This result aligns with Cui et al. (2014), showing that neither spaces nor shading affects preview processing of compound words. However, Cui et al. found that spaces facilitated preview processing of single-character words, whereas the present study found no difference in preview effects for target characters between inter-word shading and normal presentation conditions. The researchers suggested that although inter-word shading and spaces both provide word boundary information, their mechanisms may differ. This indicates that when investigating word boundary information, it is necessary to distinguish whether differences result from added boundary information or from the specific marking method used, as these involve different cognitive processes. Furthermore, the lack of difference in preview effects between inter-word shading and normal presentation suggests that word boundary information is not the fundamental cause of differences in preview processing between Chinese and English, requiring deeper analysis of underlying factors affecting processing mechanisms across writing systems. As noted previously, inter-word spaces not only provide boundary information but also reduce or eliminate inter-word boundary masking. Therefore, differences between studies may result from reduced boundary masking in the inter-word space condition, confirming Cui et al.' s speculation about their findings. This result also indicates that word boundary information does not facilitate preview processing during early stages, suggesting that word segmentation in Chinese reading is an intelligent process that does not require boundary information provided by inter-word shading to determine word boundaries. Li, Rayner, and Cave' s (2009) model of word segmentation and lexical recognition proposes that readers automatically and efficiently segment words using pre-existing lexical representations during reading. Consequently, inter-word shading did not facilitate preview processing, indicating that word boundary information is not the cause of differences in preview processing between Chinese and English, and that deeper theoretical investigation of underlying factors is needed.

We found that preview effects for inter-word shading were significantly smaller than for normal presentation only in regression path duration. However, Drieghe et al. (2017) found that unspaced text produced larger preview effects than spaced text only in first fixation duration and single fixation duration, and only when the pre-target word was high-frequency; no differences emerged for low-frequency pre-target words. The discrepancy between studies may result from differences in processing difficulty of pre-boundary words. When pre-boundary word processing is easy, readers are more likely to process preview information early, leading to increased preview effects. Therefore, the relationship between word segmentation and preview processing requires more fine-grained analysis from a dynamic perspective, examining the temporal dynamics of sequential words.

Finally, we found no differential effects of inter-word shading on preview processing of between-words and within-words in first fixation duration, gaze duration, or regression path duration. Two possible explanations exist: First, word boundary information may affect preview effects for between-words more than within-words. When processing the first morpheme of between-words, inter-word shading provides boundary information that facilitates word segmentation, potentially producing larger preview effects for between-words in the inter-word shading condition than in normal presentation. For within-words (the second morpheme of two-character words), readers have already completed segmentation of the current word, so differences between inter-word shading and normal presentation should be smaller. However, because inter-word shading provides incorrect word length information for within-words during pseudo-character preview (Cui et al., 2014)—readers are more likely to treat the pseudo-character as a single-character word—this may increase reading times for pseudo-character previews, thereby increasing the difference in preview effects for within-words between inter-word shading and normal presentation conditions. For between-words (the first morpheme of two-character words), the boundary information provided by the pseudo-character is consistent with that provided by shading, so the impact of pseudo-words on boundary information provided by shading is minimal. When the effect of shading-provided boundaries equals the effect of incorrect word length information provided by pseudo-characters, no significant difference emerges in how inter-word shading affects preview processing of between-words versus within-words.

Second, word segmentation and lexical recognition may occur simultaneously. Li et al. (2009) proposed a model of word segmentation and lexical recognition in which readers first process visual features of characters in parallel, then transmit this information to word processing units that activate corresponding words. In the word recognition unit, related words compete, and through multiple cycles, a single word emerges as the winner. When a word wins this competition, it is recognized and segmentation is complete. According to Li et al., target characters are processed simultaneously regardless of whether they occupy the first or second morpheme position, and when the target word is recognized, word boundary information is also identified, resulting in no differential effects of boundary information on preview processing of between-words and within-words. This again suggests that previous explanations attributing larger Chinese preview effects to the lack of word boundaries require reconsideration and necessitate deeper theoretical investigation.

In summary, by manipulating whether the invisible boundary fell before the first morpheme of a compound word or between its first and second morphemes, this study strictly controlled whether the target character and the character preceding the boundary belonged to the same linguistic unit. Results demonstrated significantly larger preview effects for within-words than between-words, indicating that linguistic properties affect preview processing when the influence of spaces is eliminated. Additionally, by adding word boundary information (inter-word shading) to Chinese text, we examined the influence of boundary information on preview processing. Results showed that shading-provided boundary information did not significantly affect preview processing of either between-words or within-words, suggesting that word segmentation and lexical recognition may occur simultaneously, consistent with Li et al.' s (2009) model of word segmentation and lexical recognition.

The present study yields the following conclusions: (1) Preview effects are larger for within-words than between-words, indicating that linguistic units affect preview processing and that readers can flexibly segment words and conduct preview processing based on character position information. (2) Inter-word shading did not significantly affect preview processing of either between-words or within-words, suggesting that word segmentation and lexical recognition may occur simultaneously, consistent with Li et al.' s (2009) model of word segmentation and lexical recognition.

References

Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1277-1287.

Chinese Daily Word Frequency Dictionary. (1998). Beijing, China: Beijing Language and Culture University. [人民日报词频表. (1998). 北京, 中国: 北京语言大学信息处理研究所.]

Cui, L., Drieghe, D., Bai, X., Yan, G. & Liversedge, S. P. (2014). Parafoveal preview benefit in unspaced and spaced Chinese reading. *Quarterly Journal of Experimental Psychology, 67*(11), 1550-1561.

Cui, L., Drieghe, D., Yan, G., Bai, X., Chi, H. & Liversedge, S. P. (2013). Parafoveal processing across different lexical constituents in Chinese reading. *Quarterly Journal of Experimental Psychology, 66*, 403-416.

Drieghe, D., Brysbaert, M., & Desmet, T. (2005). Parafoveal-on-foveal effects on eye movements in text reading: Does an extra space make a difference? *Vision Research, 45*(13), 1693-1706.

Drieghe, D., Cui, L., Yan, G., Bai, X., Chi, H., & Liversedge, S. P. (2017). The morphosyntactic structure of compound words influences parafoveal processing in Chinese reading. *Quarterly Journal of Experimental Psychology, 71*(1), 1-28.

Drieghe, D., Fitzsimmons, G., & Liversedge, S. P. (2017). Parafoveal preview effects in reading unspaced text. *Journal of Experimental Psychology Human Perception & Performance, 43*(10), 1701-1716.

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.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*, 777-813.

Hoosain, R. (1992). Psychological reality of the word in Chinese. In H.-C. Chen & O.J.L. Tzeng (Eds.), *Language processing in Chinese* (pp. 111-130). Amsterdam, The Netherlands: North-Holland.

Hyönä, J. (1995). Do irregular letter combinations attract readers' attention? Evidence from fixation locations in words. *Journal of Experimental Psychology Human Perception & Performance, 21*(1), 68-81.

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement-contingent display change study. *Memory & Cognition, 32*, 523-532.

Juhasz, B. J., Pollatsek, A., Hyönä, J., Drieghe, D., & Rayner, K. (2009). Parafoveal processing within and between words. *Quarterly Journal of Experimental Psychology, 62*, 1356-1376.

Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology, 58*, 525-552.

Li, N., Wang, S., Mo, L., & Kliegl, R. (2017). Contextual constraint and preview time modulate the semantic preview effect: Evidence from Chinese sentence reading. *Quarterly Journal of Experimental Psychology, 71*(1), 1-32.

Liu, Y., Reichle, E. D., & Li, X. (2015). Parafoveal processing affects outgoing saccade length during the reading of Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 1229-1236.

Liu, Y., Reichle, E. D., & Li, X. (2016). The Effect of Word Frequency and Parafoveal Preview on Saccade Length During the Reading of Chinese. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 1008-1025.

Morris, R. K., Rayner, K., & Pollatsek, A. (1990). Eye movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 268-281.

Perea, M., & Acha, J. (2009). Space information is important for reading. *Vision Research, 49*, 1994-2003.

Perea, M., Tejero, P., & Winskel, H. (2015). Can colours be used to segment words when reading? *Acta Psychologica, 159*, 8-13.

Pollatsek, A., & Rayner, K. (1982). Eye movement control in reading: The role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 817-833.

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

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Rayner, K. (2009). The thirty fifth Sir Frederick Bartlett lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*, 1457-1506.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 720-732.

Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology, 40*, 473-483.

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

Rayner, K., & Schotter, E. R. (2014). Semantic preview benefit in reading English: The effect of initial letter capitalization. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 1617-1628.

Rayner, K., White, S. J., Kambe, G., Miller, B., & Liversedge, S. P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *Cognitive and applied aspects of eye movement research* (pp. 213-234). Amsterdam: Elsevier.

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

Schotter, E. R., Lee, M., Reiderman, M., & Rayner, K. (2015). The effect of contextual constraint on parafoveal processing in reading. *Journal of Memory and Language, 83*, 118-139.

Sheridan, H., Reichle, E. D., & Reingold, E. M. (2016). Why does removing inter-word spaces produce reading deficits? The role of parafoveal processing. *Psychonomic Bulletin and Review, 23*(5), 1543-1552.

Vasilev, M. R., & Angele, B. (2017). Parafoveal preview effects from word N + 1 and word N + 2 during reading: A critical review and Bayesian meta-analysis.

Psychonomic Bulletin & Review, 24(1), 84-100.

Wang, S., Tong, X., Yang, J., & Leng, Y. (2009). Semantic codes are obtained before word fixation in Chinese sentence reading: Evidence from eye-movements (in Chinese). *Acta Psychologica Sinica, 41*, 220—232. [王穗苹, 佟秀红, 杨锦绵, 冷英. (2009). 中文句子阅读中语义信息对眼动预视效应的影响. 心理学报, 41(3), 220-232.]

Wei, W., Li, X., & Pollatsek, A. (2013). Word properties of a fixated region affect outgoing saccade length in Chinese reading. *Vision Research, 80*, 1-6.

Yan, M., Kliegl, R., Richter, E., Nuthmann, A. & Shu, H. (2010). Flexible saccade-target selection in Chinese reading. *Quarterly Journal of Experimental Psychology, 63*, 705-725.

Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Readers of Chinese extract semantic information from parafoveal words. *Psychonomic Bulletin & Review, 16*, 561-566.

Yan, G., Xiong, J., Zang, C., Yu, L., Cui, L., & Bai, X. (2013). The advance of Parafoveal Preview Effects in Chinese Reading. *Advances in Psychological Science, 21*(4), 589-605. [闫国利, 熊建萍, 臧传丽, 余莉莉, 崔磊, 白学军. (2013). 阅读研究中的主要眼动指标评述. 心理科学进展, 21(4), 589-605.]

Yang, J., Wang, S., Xu, Y., & Rayner, K. (2009). Do Chinese readers obtain preview benefit from character n+2? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1192-1204.

Yen, M., Radach, R., Tzeng, O. J., Hung, D. L., & Tsai, J. (2009). Early parafoveal processing in reading Chinese sentences. *Acta Psychologica, 131*, 24-33.

Zang, C., Wang, Y., Bai, X., Yan, G., Drieghe, D., & Liversedge, S. P. (2016). The use of probabilistic lexicality cues for word segmentation in Chinese reading. *Quarterly Journal of Experimental Psychology, 69*, 548-560.

Note: Figure translations are in progress. See original paper for figures.